CONSUMER CREDIT RISK MODEL WITH HIERARCHICAL DATA STRUCTURE

GAMINI PREMARATNE
DU JUN

Abstract. Credit risk researchers have mainly focused on area of wholesale and corporate loans. The objective of this paper is to contribute the literature by introducing a model that is accordant with consumer credit risk characteristics. Unprecedentedly, we bring credit card industry practitioners' concepts, such as *buckets* and *credit card account structures* that are utilized in practitioners' routine risk control procedures, into credit risk modeling. Leveraging on the hierarchical models from biomedical and educational literature, we proposed a new model that can account for multi-layer of associations among credit card accounts. Applying to a real credit card portfolio, the proposed model shows great advantages over other four reference models discussed in the paper. The paper highlights that without proper assumptions on the various correlations sourced from data structures, credit scoring models may be less relevant and sometimes misleading.

Introduction

The consumer lending market covers a large number of products, including mortgages, loans, credit cards, debit cards, etc. Among these products, credit card market is one of the most rapidly growing consumer financial products in Singapore, both in terms of monetary significance and simply the number of people involved. From Table 1, it's shown that Singapore credit card annual transaction amount increase on average by 16% in past five years and reached Singapore Dollar (S$) 25.66 billion in 2008. The number of people who have access to credit card facilities also increased dramatically. Number of credit card facilities grows to 6.6 million (includes supplementary cards), higher by 67.8% in last five years, in comparison of a 19.7% incremental in total population during the same period. Another significant characteristic of the credit card market is the increased competition. Imagine each local resident of Singapore on average has 1.3 credit card plastics in his/her wallet, this implies the credit card issuers, competing for larger market shares, are inevitably expanding their new acquisitions toward population with lower income and less credibility under an unpredictable world wide economic climate. With incontestable financial significance and increasingly complicated market, it is not surprising that the retail banking industry is paying more and more attention towards credit risk associated with credit card business. More desirable than ever, sophisticated mathematical and statistical models are sought by card issuers in order to properly analyze their credit card portfolio and to accurately predict their cardholders' performance.

Key words and phrases. consumer credit risk, hierarchical model, credit card.
[1]Source of data: Monetary Authority of Singapore, Monthly Statistical Bulletin

Table 1. Singapore Credit Card Statistics

Period	Number of Cards		Total Card	Rollover	Charge-off [1]
	Main	Supplementary	Billings (S$Mil.)	Balance (S$Mil.)	Rates (%)
2004	2,985,973	946,784	14,047	2,641	7.6
2005	3,415,507	1,026,516	16,073	2,842	5.0
2006	3,968,044	1,121,932	18,640	2,822	4.4
2007	4,472,124	1,173,988	22,640	2,979	3.9
2008	5,056,450	1,221,333	25,662	3,379	3.6
2009 Jan	5,078,770	1,226,593	1,989	3,440	
2009 Feb	5,097,939	1,231,046	1,784	3,508	
2009 Mar	5,123,931	1,234,111	2,038	3,340	4.6
2009 Apr	5,156,519	1,230,370	1,953	3,398	
2009 May	5,186,062	1,242,738	1,999	3,455	
2009 Jun	5,230,856	1,249,206	2,214	3,499	5.8
2009 Jul	5,273,102	1,258,057	2,133	3,443	
2009 Aug	5,297,630	1,261,049	2,170	3,499	
2009 Sep	5,353,195	1,268,326	2,287	3,618	5.7
2009 Oct	5,332,990	1,264,523	2,343	3,605	

As of recent, consumer credit risk literature is broadly classified into two groups. The first group is largely motivated and driven by the release by the Basel Committee on Banking Supervision (BCBS henceforth) in 1999 of a consultative paper on a New Basel Capital Accord (NBCA, see BCBS (1999)[18]) for internal banks capital ratio benchmarks. Intrinsically, the main focus of the new Basel proposal is not for consumer lending portfolio but rather mainly for wholesale and corporate portfolios. In addition, popular modern commercial techniques, like CreditMetrics [14], CreditRisk+ [19] and Moody's KMV [16], are not specially customized for retail portfolio as well. This has raised concerns to many practitioners and researchers in consumer lending area. And their concerns were voiced out in the Journal of Banking & Finance (2004) special issue [8]. Separately, Andrade and Thomas (2004) [7] proposed a theory for consumer default. Using certain assumptions, they provided a option-based reasoning for process of default in consumer credit. Andrade and Sicsú (2007) [6] proposed a model to generate the distribution of a portfolio default rate by Monte Carlo simulation.

The second set of literature on consumer credit risk focuses on developing statistical tools to monitor customer performance, characterize behavior patterns and identify bad risks. A thorough review on these statistical methods can be found in Hand (2001) [15]. In existing statistical methods, basic unit of analysis is usually considered as individual loan account. Thus, association among accounts owned by single customer is ignored. Moreover, these methods do not monitor customers' evolvement over time. So valuable information such as inter-account correlation, autocorrelated behavior or macro-economic impact on a cohort is washed out. In this paper, we engrafted the logic that credit card issuing banks used to build data structure for credit card management and the hierarchical models popular in biomedical and educational studies to develop a new credit risk model that can account for multiple layers of association among credit card loans.

Credit cards have evolved over the last forty years as one of the most accepted, convenient, and profitable financial products and play an important role in the strategic plans of many banks. This paper is inspired by credit control procedures adopted by these banks. Their way of designing products, payment structures and portfolio segmentation haven't been put under scrutiny and thus interesting finding may lie inside. On the business requirement side, a nicely wrapped-up, sensitive and relevant credit risk model will come in handy for banks' self assessment and strategic planing.

We established a new credit risk model with hierarchical generalized linear model structure, which borrows statistical methods from other research areas such as educational and biomedical studies and provided for credit risk research a new direction for further growth. In most existing consumer credit risk models, basic unit of analysis is usually individual loan account. Thus, association among accounts owned by single customer is ignored. In addition, these methods do not monitor customers' evolvement over time. So valuable information such as inter-account correlation, autocorrelated behavior or macro-economic impact on a cohort is ignored. In contrast, the proposed model is intended to analyze underlying credit worthiness of credit card cardholders measured by ordinal responses to single or multiple items, where items represent one or multiple financial decision problems facing by a credit card holder on a particular time-point (in month or quarter). Items are nested within a time-point and time-points in turn are nested within subject (credit card holder). In this way, very flexible structure of association among observations can be assumed. On top of this merit, the special design of the model can allow us handle missing observations at time- and item-level, which is a common problem in longitudinal analysis. The model was implemented on a dataset sampled from credit card cardholders' population of a consumer bank employing Bayesian approach via Markov chain Monte Carlo (MCMC) methods.

The remainder of this paper is organized as follows. In section 1, we give a brief description on the credit card accounts structures and the hierarchical model, which is popular in other research areas. Based on it, in section 2 we describe the data that will be used in this study. In section 3 we develop a new credit risk model that can leverage on the credit card accounts structure to analyze the various correlations within the credit card portfolio. In addition in section 4, the proposed model is fit to the data and is compared with other four reference models. Our conclusion for the paper is presented in section 5.

1. Introduction to Credit Card Account Structure and Hierarchical Models

1.1. **Credit Card Account Structure.** Credit cards are not only a tool used by consumers to make retail transaction convenient but also a cost saving channel widely utilized in consumers' daily life: credit card issuers will typically offer special offers, like discount, rewards points cumulation or cash rebate, target to different types of credit cards segments. Thus there is incentive for consumers to hold

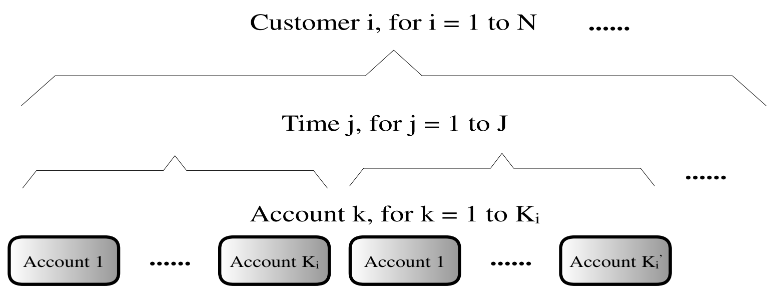


FIGURE 1. Credit Card Accounts Structure

multiple credit card accounts for different spending purpose. In some countries, e.g. in Singapore, credit card market supervision enforce a combined credit limit policy which requires multiple credit card account under one customer's name have to share one credit limit. Under this scenario, it is important for card issuers to identify and analyze credit risk at customer level rather than account level, because risk indicators emerged from a single account can not fully describe the true risk level intrinsic to a cardholder who own multiple credit card accounts.

However, surprisingly many credit card issuers still mainly base their risk mitigation operations on account level risk analysis. This is not due to limitation of information acquisition – equipped with modern credit card account management systems, credit cards issuers can produce and access huge volume of customer information on daily basis. The roadblock of fully utilizing customer information is that there is little choice of statistical models that can describe association of individual accounts under same customer. For instance, if a customer has two credit cards accounts, one is 30 days past due and the other is a prompt payment account, how the correlation act on customer's behavior on each of the two remains an unsolved puzzle to credit card issuers. Similarly, in literature, research papers on credit card delinquency analysis (Gross and Souleles [12] and Zhao, Zhao and Song [26]) generally use credit card account as main unit of analysis, rather than individual or household.

In card issuers' daily operations, information of credit card applicants, for example demographic data, are collected at the time of card application, which is usually static information. After accepted by card issuer as customer, cardholders accounts are reviewed and evaluated at different points of time. The recorded information is usually called as snapshots of the cardholder and is dynamic. As mention above, it's popular that a customer would like to apply for multiple credit card accounts to meet his/her different type of demands. One customer may have multiple credit card accounts, which are usually tracked and recorded separately. Thus we can imagine credit cards portfolio of a issuer as data organized in hierarchy. We refer to a hierarchy as consisting of units grouped at different levels. The fact that snapshot of account information are nested under a snapshot of customer, which in turn are grouped under each individual customers, can be depicted by Figure 1.

Many card issuers use segmentation approach to manage and analyze credit card portfolios, which may includes millions of customers. As summarized in Comptroller's Handbook of Credit Lending [17], portfolio segmentation usually considers some aspect of portfolio delinquency, and generally divides the portfolio into various degrees of delinquency, or buckets, such as: current bucket (M0 bucket); 0-29 days past due (M1 bucket); 30-59 days past due (M2 bucket); 60-89 days past due (M3 bucket), etc. With the whole portfolio being segmented into various delinquency buckets, the degree of migration from one bucket to the next is measured over time; i.e. the bank's credit control management department tracks the volume of loans which roll from, say, the M1 bucket to the M2 bucket, and measure this volume through a roll-forward rate. Application of the roll-forward rates to the volume of loans in each bucket will provide some estimation of losses in the existing portfolio.

A credit card issuer will send customer a monthly statement detailing all the purchases the customer has made with the card during the month. Every statement explicitly indicates a date by which the customer should pay his credit card bill without incurring a late payment charge (payment due date). The customer has the option to pay the bill in full, or make a partial payment subject to a minimum sum. As a practice, card issuers generally do not charge interest if the outstanding balance shown on the monthly statement is paid up in full by the payment due date. However if the outstanding balance is not paid fully by the due date, interest charges will incur. A credit card holder can pay a minimum due, which is usually 2% of balance as the interest charge plus 1% of the principal. If full or partial payments is not made by the due date, a late payment charge would be levied by the card issuer and the customer is considered to be a M1 delinquent customer. In the next month a M1 delinquent account is required to at least pay the minimum due as of last month plus 3% of the latest outstanding balance such that it can be regularized. Now the customer has three alternative choices: if the customer is able to pay total due amount, his account is regularized; if the customer is able to pay only one bucket of payment, the account will stay in M1 bucket; if he chooses to default again, then the account will flow into M2 bucket. If the customer continuously default such that his account flow over M6 (180+ DPD), the card issuer will contractually write the total balance of the account off from its book and admit it as credit losses. Under certain circumstance, card issuer need to report credit losses from an account earlier than 180 DPD. This kind of write-off is called *Non-contractual write-off* (NCWO). Most common examples of NCWO are customer bankruptcy. Hence, for accounts in any bucket there is possibility of write-off. Figure 2 illustrate how credit card accounts migrate among buckets through different paths.

From the flow chart, we can find that

(1) An account need at least 7 months to flow from current bucket to contractual charge-off (CWO).
(2) For each month customers in $M(x)$ bucket will face a $x+3$ choices problem; For example, for an account currently reside in $M1$ bucket, there are four outcomes can happen for the next month, namely (1) flows back to $M0$; (2) stays in $M1$; (3) roll forward into $M2$; (4) or is written off (NCWO in this scenario).
(3) Up to payment arrangement, different delinquency path can be chosen by customer and it will cause different economic effects on the card issuer's balance sheet.

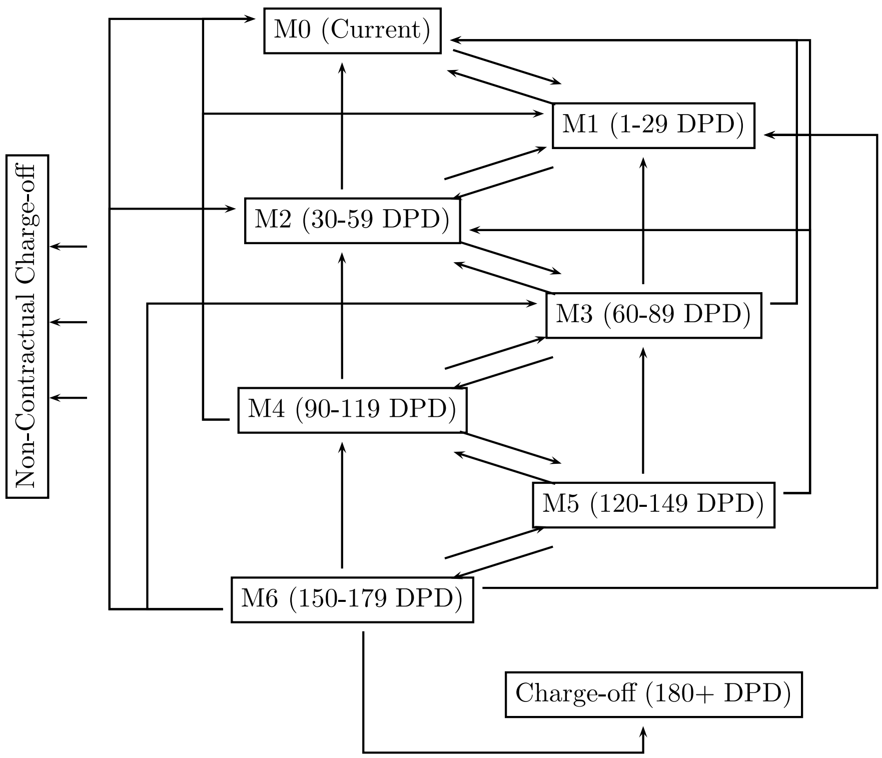


Figure 2. Flow Chart of Credit Card Accounts

(4) Unlike accounts in all the buckets, written-off accounts (CWO and NCWO) are unable to move anywhere.
(5) The total credit losses in a time horizon is lump sum of losses caused by all written-off accounts.

1.2. **Application of Hierarchical Models.** This hierarchical data structures, though not familiar to retail credit risk researchers, increase in popularity in social and behavioral science and clinical trials studies. For example, offspring from the same parents tend to be more alike in their physical and mental characteristics than individuals chosen at random from the population at large. Many designed experiments also create data hierarchies, for example clinical trials carried out in several randomly chosen centers or groups of individuals.

Singer 1998 [22] introduced a flexible method for fitting multilevel models, hierarchical models, and individual growth model. In her paper, she took classical two-level school effects model as example. In the two-level school effects example, she described a general situations in which data are nested at two levels. It's an organizational hierarchy such as students grouped within classes and then classes nested within schools. And the goal is to examine the behavior of a level-1 outcome (student performance) as a function of both level-1 (class effects) and level-2 (school effects) predictors. The ideas presented by the example can be easily extended to three-level or multilevel models.

In the area of survey or experiment analysis, measurements are usually repeated on the same subjects for different time-points. Such data are often referred to as 'longitudinal' (or panel in economics context) as opposed to 'cross-sectional' where each subject is measured only once. Longitudinal data can be involved in hierarchical models. Consider for example a data set consisting of repeated measurements of the heights of a random sample of children. Then the linear model between height (y) and age (x) can be expressed as

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + e_{ij} \tag{1}$$

Moreover, equation (1) can be easily extended to include further explanatory variables, measured either at the time level, such as household average income of year, or at the subject level such as birthweight. For example the equations for the discussed model can be formed as below (see Goldstein 2002 [11] for more details)

$$\begin{aligned} &y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 s_i + \beta_2 w_j + e_{ij} \ where \ e_{ij} \sim N(0, \sigma^2) \\ &s_i = \alpha_0 + \alpha_1 BirthWeight_i + r_i \ where \ r_i \sim N(0, \tau^2) \\ &w_j = \delta_0 + \delta_1 AvgIncome_j + u_j \ where \ u_j \sim N(0, \upsilon^2), \end{aligned}$$

where response-level explanatory variables s_i and w_j themselves rely on subject-level variable $BirthWieght$ and time-level variable $AvgIncome$.

An hierarchical model can also be extended to include autoregressive structure. Wang et al. 2007 [25] presented a multilevel survival frailty model for analyzing clustered data and recurrent disease among aged population. Beside the cross-sectional correlations of observations due to similar residential environments considered in their model, the subject level serial dependence is assumed between recurrent events recorded on the same individual. Therefore, their method incorporated random components into the usual survival frailty model to account for the autoregressive structure.

Given the similar data structure between credit card accounts (as depicted in Section 3.1.1) and above-mentioned social behavioral science and clinical trials studies, we developed a credit risk model as a special three-level hierarchical generalized linear model that can be applied to ordinal data with complex data structure. The aim is to comprehensively model credit card accounts data with the following features: (a) velocity of deterioration along increasing delinquency bucket, (b) correlations of accounts at different buckets but under same cardholder, (c) impact of macro-economic environment on cardholders decision, (d) dynamic structure of cardholder's credit worthiness.

2. The Data

Credit card account records, collected and organized by credit card issuers, usually consist of a wide range of details about each accounts. Following Avery et al. 2004 [2], the data can be classified into five broad categories: Account identification (eg account number and customer number), account dates (eg account open date and write-off date), account balances, account description, and payment performance.

The credit account records used in this study is provided by a consumer bank in Singapore. The sample includes information from monthly statements of 189 credit card accounts (143 cardholders) from December 2002 until June 2004. All the credit card accounts in our sample are active at the the beginning of our sample period (excluding charged off accounts and closed inactive accounts). These accounts are, at least once, 30 days past due in customer level repayment history [2] and are randomly selected from the bank's credit card customer database. [3] The performance of these accounts are tracked through 2003 Quarter 1 to 2004 Quarter 2. This time period was selected for several reasons: (1) due to business confidentiality and agreement with the bank, very limited time horizon is allowed to be used in this study. For randomly-selected sample size in this study, only 19-month continuous track on restricted variables is available; (2) To demonstrate applicability of the proposed model under various timing environments, the sample period managed to include seasonality effects i.e. Jan-Feb holiday season, extreme economic downturn in Q3 Q4 of Singapore economy due to SARS; (3) SARS period was included as SARS crisis broke out in Singapore in March 2003, which essentially hurt Singapore economic performance at 2003 Q3 and Q4, and the impact quickly faded out after that. How Singapore credit card holders responded to this short term crisis is an interesting question that can be answered by the proposed model. The data provides card holder identifications[4], delinquency history, credit limit, open dates, account status. In addition, the data set also provides information on consumers' demographics such as age, gender and nationality.

We present summary statistics of cardholders' performance and demographic variables in Table 2. There are 142 cardholders 188 active accounts prior beginning of 2003 Q1 in the sample. The sample size reduced to 99 cardholders and 129 accounts at 2004 Q2 beginning, which is due to 59 accounts were written off during 2003 Q1 - 2004 Q1. Account deterioration is defined as the situation where next quarter worst bucket is worse than current quarter worst bucket. *No. Acct Deteriorate* reports this quantity in terms of number of accounts cross sub-sample periods. *No. Acct Woff* reports this number of accounts written off within each sub-sample period. The corresponding percentage ratio of these two metrics are presented in the the 6th row and 7th row respectively. It can be observed that there is a significant increase of deterioration rate and written-off rate in 2003 Q3. Singapore unemployment rates [5] are given in the last row of Table 2, which supports a positive correlation of unemployment rate and credit card deteriorate/default rate. *Avg Balance*, *Avg Dlq Mth Lst 12 mth* and *Avg MOB* give the average balance size (in SG$), average number of delinquent months per cardholder in last 12 month and average months on book respectively. Cardholders' average age and percentage of foreigner accounts in the portfolio are reported in the table as well.

---

[2]This study focuses on delinquent cardholders' behavior and thus only cardholders with stain on their repayment history are selected.

[3]Banks usually assign a random number to an account upon the time of booking. Utilizing the random digit, x% of the bank's customer base was randomly picked up for the study.

[4]Cardholders' identification information used in this study was converted from original customer numbers into descending serial numbers, and thus can only be used for tagging. In this way no customer sensitive information was released by the bank.

[5]Seasonally Adjusted Unemployment Rate. Obtained from Minster of ManPower website: http://www.mom.gov.sg/publish/momportal/en/communities/others/mrsd/statistics/Unemployment.html

TABLE 2. Summary Statistics of Performance and Demographic Variables

	2003 Q1	**2003 Q2**	**2003 Q3**	**2003 Q4**	**2004 Q1**
No. Cardholders	142	137	124	112	99
No. Acct Rglr / Imprv	102	110	78	80	84
No. Acct Deteriorated	81	57	70	49	35
No. Acct Woff	5	16	19	19	10
% Acct Deteriorate	43.1%	31.1%	41.9%	33.1%	27.1%
% Acct Woff	2.7%	8.7%	11.4%	12.8%	7.8%
Avg Balance	$ 4,277	$ 4,230	$ 4,287	$ 4,525	$ 3,993
Avg Dlq Mth Lst 12 mth	7.2	8.2	9.0	8.8	8.3
Avg MOB	60.6	63.3	67.3	71.1	72.4
Age	40	40	39	39	40
No. Foreign Accts	5	5	3	3	2
Unemply Rate	3.7%	3.6%	4.8%	3.9%	3.7%

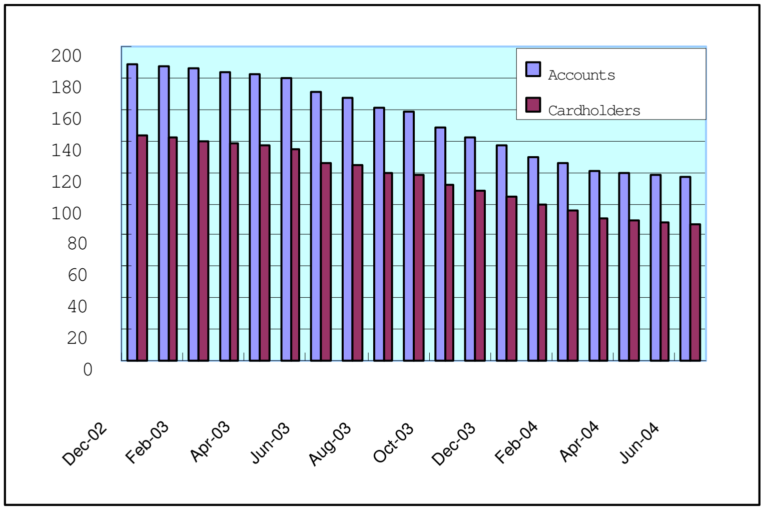


FIGURE 3. Dropout of Accounts & Customers over sample period

Figure 3 shows numbers of "survived" accounts and "survived" cardholders at each month and how the sample size shrinks over time. The dropout accounts/cardholders are those accounts written off from bank's financial book and are removed from our sample. Similar to other longitudinal studies, data used in this research requires handling missing of records. In this study it is assumed that the missing mechanism is *ignorable* and thus inference is base on the observed-data posterior of all parameters, ignoring the missing data mechanism. The ignorability condition, first described by Rubin 1976 [20]), can be used to facilitate posterior inference without having to specify the missing data mechanism. In later sections, detailed conditions of ignorable missing data mechanism are introduced. In literature, data sets in which time points may vary among individuals are quoted as imbalanced data. We use the following notations for missing data at the time-level.

N_i is the total number of observations for cardholder i in the sample, and subscription "$_{ij}$" denotes a time point for the j-th observation among the observations for cardholder i. As the only cause of missing in this study is the write-off process of issuing bank, which is a non-reversible process, the observation number j is identical as time point j for all cardholders (as there is no interruption among observations of a cardholder).

Data grouping and categorizing are used to simplify the question studied but without limiting the generality of the model. M2+ (30+ days past due (DPD)) accounts are grouped together as one item in cardholder's decision making process. The grouping takes assumption that accounts with DPD over 30 days are homogeneous in characteristics. In this model we define a cardholder CM_i maximally have three items at a time point j

$$CM_{ij}\ with \left\langle \begin{array}{lll} Item\ 1: & \text{if} & account(s)\ in\ M0, \\ Item\ 2: & \text{if} & account(s)\ in\ M1, \\ Item\ 3: & \text{if} & account(s)\ in\ M2+. \end{array} \right.$$

For missing observations at the item level, K_{ij} denotes a set of items that are observed for cardholder i at the j-th time point. For example, if the first and the third items are observed at the $j$ith time point for cardholder i, then $K_{ij} = \{1, 3\}$.

Cardholders are measured in two dimensions, time and item. Let y_{ijk} denote a response to the k-th item at the j-th time point for cardholder i. The category of response to item k is denoted as c, where c can take value from 1 to C. Although number of categories can be different over items, for simplicity we make it consistent cross all items that each item can have three categories, improved/stay flat, worsening and written-off, as specified in below equation.

$$y_{ijk} = \begin{cases} 1 & \text{if} \quad account(s)\ performance\ improve/stayflat, \\ 2 & \text{if} \quad account(s)\ performance\ worsen, \\ 3 & \text{if} \quad account(s)\ written-off. \end{cases}$$

3. The Model

In this section, the corresponding model of the above-mentioned data structure is described. Since credit card holders are measured in two dimensions, time and item, we present a three-level model, item level (within-time), time level (within-cardholder), and between-cardholder level. A similar three-level structure can be found in Segawa (2005) [21], where a multi-indictor growth model were formulated to analyze growth of trait latent variable measured by ordinal items.

3.1. Item Level (Within-time Level) Model.

The item level ordinal response is modeled by standard ordinal logit model. We denote the "true" value of the ith cardholder's latent capability of repayment debt for item k on time point j as y^*_{ijk}. y^*_{ijk} is a continuous latent variable which determine the observed response y_{ijk},

$$y_{ijk} = \begin{cases} 1 & \text{if} \quad -\infty < y^*_{ijk} \leq \alpha_{k1}, \\ 2 & \text{if} \quad \alpha_{k1} < y^*_{ijk} \leq \alpha_{k2}, \\ 3 & \text{if} \quad \alpha_{k2} < y^*_{ijk} \leq \infty. \end{cases} \tag{2}$$

where α_{kc} is a threshold parameter of the upper cutoff of category c of item k ($-\infty$ and ∞ are actually α_{k0} and α_{kC} respectively), satisfying the condition $\alpha_{k0} < \alpha_{k1} < \alpha_{k2} < \alpha_{kC}$.

In responding to item k at time point j, cardholder is assumed to make a decision based on the latent variable

$$y^*_{ijk} = \beta_{0k} + \beta_{1k}\theta_{ij} + \epsilon_{ijk} + \epsilon^*_{ijk} \tag{3}$$

where ϵ^*_{ijk} is independently and identically distributed as the standard logistic for all i, j, and k with $E(\epsilon^*_{ijk}) = 0$, and $var(\epsilon^*_{ijk}) = \sigma^2_{e^*} = \frac{\pi^2}{3}$ (location and scale parameters are 0 and 1 respectively). The second error term ϵ_{ijk} is normally distributed random effect with $var(\epsilon_{ijk}) = \sigma^2_e$. For identifiability, the mean random effect is assumed to be 0 and the variance of the random effect σ^2_e is estimated from the data. θ_{ij} are latent trait representing underlying credit worthiness of cardholder i at time j. β_{0k} and β_{1k} represent intercept and slope parameters at item level, respectively.

Following the ordinal logit setup, denote the cumulative probability of cardholder i responding to item k at category c and below at j-th time point as $Q_{ijk,c}$. We can model it without explicitly introducing latent variable y^*_{ijk} as below

$$logit\ Q_{ijk,c} = \alpha_{kc} - (\beta_{0k} + \beta_{1k}\theta_{ij}). \tag{4}$$

Then the probability of cardholder i responding to item k at j-th observation with category c is given by $p_{ijk,c} = Q_{ijk,c} - Q_{ijk,c-1}$. The likelihood function for item level parameters β and α follow directly from equation 4, that is,

$$L(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \prod_{i=1}^{N}\prod_{j=1}^{T}\prod_{k=1}^{K_{ij}}[F(\alpha_{y_{ijk}} - (\beta_{0k} + \beta_{1k}\theta_{ij})) - F(\alpha_{y_{ijk}-1} - (\beta_{0k} + \beta_{1k}\theta_{ij}))], \tag{5}$$

where y_{ijk} is the observed value of category c for subscription $_{ijk}$.

If we parameterize the model by using the latent data $\mathbf{y}^*$ along with the parameters β and α, then the likelihood may be written in terms of the latent variable $\mathbf{y}^*$ as

$$L(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{y}^*) = \prod_{i=1}^{N}\prod_{j=1}^{T}\prod_{k=1}^{K_{ij}} f(y^*_{ijk} - (\beta_{0k} + \beta_{1k}\theta_{ij}))I(\alpha_{y_{ijk}-1} \le y^*_{ijk} < \alpha_{y_{ijk}}), \tag{6}$$

where $I(\cdot)$ indicates the indicator function. Note that the latent variables y^*_{ijk} may be integrated out of equation (6) to obtain equation (5). This representation of likelihood suggests a simple Gibbs sampling approach for simulating from the joint posterior of $(\mathbf{y}^*, \beta, \alpha)$, which will be elaborated with multilevel setup in later section.

3.2. **Time Level (Within-cardholder Level) Model.** The within-cardholder level model is used to specify the variation of the latent trait of a cardholder over time periods on top of item level specification. The within-cardholder model is of the form

$$\theta_{ij} = \mathbf{z}_{ij}\mathbf{u}_i + \epsilon_{\theta_{ij}}, \tag{7}$$

where $\mathbf{z}_{ij} = [Z_{ij,0} Z_{ij,1} \cdots Z_{ij,P}]$ represent P explanatory variables at time level (within-cardholder level), and $\mathbf{u}_i = [u_{i0} u_{i1} \cdots u_{iP}]$ are corresponding regression coefficient vector.

Depending on characteristics of the data, the residuals of cardholder traits $\epsilon_{\theta_{ij}}$ are able to be specified as either *i.i.d.* or autoregressive process. For *i.i.d.*, the residuals of traits are given as

$$\epsilon_{\theta_{ij}} \sim N(0, \sigma_\theta^2). \tag{8}$$

For autoregressive process setup, for convenience we consider $AR(1)$ only in this paper. In this case two requirements need to be satisfied: (1) There should be no interruption between observations for each cardholder. For example, the observed time points should not be 1, 2, 4. (2) The time interval of observations should be equally spaced. The requirements are satisfied as the missing data mechanism in this study is driven by card issuers' write-off process which is not reversible; cardholders' performances are all measured on quarterly basis. The $AR(1)$ trait residual $\epsilon_{\theta_{ij}}$ is modeled as

$$\epsilon_{\theta_{ij}} = \phi \epsilon_{\theta_{i,j-1}} + \epsilon^*_{\theta_{i,j}}, \ |\phi| < 1 \tag{9}$$

where ϕ is the autoregressive coefficient. For all $i = 1 \dots N$ and $j = 2 \dots J$, $\epsilon^*_{\theta_{i,j}}$ are independent and identically distributed as $N(0, \sigma_\theta^2)$. Assuming stationarity, for all cardholder at $j > 1$, the residual should be specified as

$$\epsilon_{\theta_{ij}} \sim N(0, \frac{\sigma_\theta^2}{1 - \phi^2}). \tag{10}$$

For time level explanatory variables, the key performance indicator, *customer level number of months account delinquent in last 12 months* (CMX12), is picked up to dimension cardholder performance in recent period. The raw data of this variable (with mean ranged from 7.2 to 9.0 as shown in Table 2) is demeaned and then scaled down by 10. *Unemployment rate* is selected to reflect the macro-economic effects. In 2004 conference on validation of consumer credit risk model held by Federal Reserve Bank of Philadelphia and Wharton School, one of the key topics discussed by consumer credit risk researchers is that credit issuing banks should incorporate market and economic variables in their internal scoring and loss forecasting models to cater for factors influencing loan repayment that are outside of an individual's control (Burns and Ody (2004) [4]). A preliminary observation from the Table 2 support a positive correlation between unemployment rate and credit card repayment deterioration rate and write-off rate. In this study, *unemployment rates* (UE) enter in the model as a proxy of social and economic environment and are identical for cardholders' at the same time point.

3.3. **Between-cardholder level Model.** The underlying variation of the regression coefficients, $\mathbf{u}_i$, over cardholders is further modeled by a linear regression equation involving cardholder-level covariates, $\mathbf{s}_i$

$$\mathbf{u}_i = \mathbf{s}_i \mathbf{v} + \boldsymbol{\epsilon}_{u_i} \tag{11}$$

and

$$\boldsymbol{\epsilon}_{u_i} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_u), \tag{12}$$

where $\boldsymbol{\epsilon}_{u_i}$ is *i.i.d.* for all i. And assume that ϵ^*_{ijk}, $\epsilon_{\theta^*_{ij}}$, and ϵ_{u_i} are mutually independent. For cardholder level variable, *Months on Book* (MOB) is picked up. The quantity is measured as number of months since the date credit cardholder was approved for a credit card by the issuing bank to the beginning of the sample period. This variable is commonly used by the industry and credit risk researchers as an effective predictor for consumer credit risk. Comparing those with shorter period of relationship with bank, cardholder with years of relationship with bank is more likely to survive repayment difficulty and continue the borrowing relationship with the bank. Preliminary observation from the Table 2 suggests that this is the case for our sample portfolio too. The portfolio average MOB stands at 60.6 months at 2003 Q1, which increases over time and ends at 72.4 months at 2004 Q1, suggesting more 'young' accounts are purged from sample due to bank's write-off process.

4. Model Estimation and Results

4.1. **Markov chain Monte Carlo Estimation Approach for Multilevel Ordinal Response Model.** Markov chain Monte Carlo (MCMC) methods are increasingly being used for the estimation of multilevel model due to its ability to be flexibly extended to complex models structures where maximum likelihood method may not be able to be implemented easily. Other than that, MCMC methods can help us to generate the exact posterior distribution of any function of the unknown parameters and are able to incorporate prior beliefs into model estimation. The multilevel ordinal model is estimated using MCMC methods. Goldstein et al. (2007) [9] gives a detailed description on estimation of the factor model for mixed binary and ordinal responses. Albert and Chib (1993) [1] proposed an MCMC algorithm for estimation of a single-level probit model for binary and polychotomous data, which was later extended by Goldstein and Browne (2005) [10] to estimate mixed binary responses.

In this study, the data consist of: 1) the item level responses y_{ijk}; 2) the values of the time level and cardholder level explanatory variables, denoted by $\mathbf{Z}$ and $\mathbf{S}$ respectively; 3) parameters $\boldsymbol{\beta}$, $\mathbf{u}$ and $\mathbf{v}$ representing regression coefficients for item, time and between-cardholder level respectively; 4) variance components σ^2_e, σ^2_θ and $\boldsymbol{\Sigma}_u$; 5) autoregressive coefficient ϕ at time level; 6) cardholder credit worthiness parameter $\boldsymbol{\theta}$ and item level latent trait $\mathbf{y}^*$. As a result, the full posterior distribution of the parameters given the data is derived from equation (6) as

$$
\begin{aligned}
f(\mathbf{y}^*, \boldsymbol{\theta}, \quad \phi \quad &, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{u}, \boldsymbol{v}, \sigma^2_e, \sigma^2_\theta, \boldsymbol{\Sigma_u} | \mathbf{Y}, \mathbf{Z}, \mathbf{S}) \propto \\
&\prod_{j=1}^{J_i} \prod_{i}^{N} \left(\left(\prod_{k=1}^{K_{ij}} f(y^*_{ijk} | \theta_{ij}, \boldsymbol{\beta_k}, \boldsymbol{\alpha_k}, \sigma^2_e, y_{ijk}) \right) f(\theta_{ij} | \mathbf{u_i}, \sigma^2_\theta, \phi, \boldsymbol{z_{ij}}) \right) \\
&f(\mathbf{u} | \boldsymbol{v}, \boldsymbol{\Sigma_u}, \boldsymbol{s}) f(\boldsymbol{v} | \Sigma_u) \\
&f(\boldsymbol{\beta}) f(\boldsymbol{\alpha}) f(\phi) f(\sigma^2_e) f(\sigma^2_\theta) f(\boldsymbol{\Sigma_u}), \qquad (13)
\end{aligned}
$$

where $f(\boldsymbol{\beta})$ is defined as a normal prior distribution for item level regression coefficients; $f(\boldsymbol{\alpha})$ denotes prior beliefs of category cutoffs; $f(\phi)$ represents a uniform prior ranging from -1 to 1; item level and time level variance coefficient $f(\sigma^2_e)$ and $f(\sigma^2_\theta)$ both follow inverse gamma as literature convention; and an inverse Wishart (Wishart is essentially a multivariate gamma) prior is assumed for $f(\boldsymbol{\Sigma_u})$. We selected very vague prior distribution in this study such that the prior will have minimal impact relative to the data.

As shown above, the simultaneous posterior distribution of all model parameters is quite complicated and has no close form. Therefore, the complete set of parameters is segmented into a number of subsets in such a way that the conditional posterior distribution of every subset given all other parameters has a tractable form and can be easily sampled. A MCMC procedure will be used for drawing samples from the conditional posterior distributions. For simplicity of illustration, let's assume the link function is a standard normal distribution (ordinal probit model). Then conditional on current values of $\theta_{ij}, \boldsymbol{\beta_k}, \boldsymbol{\alpha_k}, \sigma_e^2, y_{ijk}$ the posterior distribution of y^*_{ijk} result in

$$f(y^*_{ijk}|\theta_{ij}, \boldsymbol{\beta_k}, \boldsymbol{\alpha_k}, \sigma_e^2, y_{ijk}) \propto N(y^*_{ijk}; \beta 0_k + \beta 1_k \theta_{ij}, \sigma_e^2) I(\alpha_{y_{ijk}-1} \le y^* < \alpha_{y_{ijk}}), \tag{14}$$

where $N(\cdot; a, b)$ denotes a normal density with mean a and variance b and $I(\cdot)$ is the indicator function. Equation (14) suggests a truncated normal distribution we can sample y^*_{ijk} from. Conditional posterior for the rest of parameters can be derived in a similar way. The detailed sampling steps can be found in Goldstein et al. (2007) [9].

With a very flexible set of multilevel models it should be noted that in order for such models to be identifiable some constraints must be given to the parameters. These will consist of fixing the values of some parameters. For this purpose, $\beta 0_k = 0$ and $\beta 1_k = 1$ is held as constant for $k = 1$, and $\alpha_k = 0$ are imposed for $k = 1, \ldots, K$.

In our data, not all accounts survive till the end of the sample period. It is assumed that the missing mechanism is ignorable. To define missing data assumption, it is useful to expression missing data mechanism as

$$p(\boldsymbol{r}|\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{\psi}) = p(\boldsymbol{r}|\boldsymbol{y_{obs}}, \boldsymbol{y_{miss}}, \boldsymbol{x}, \boldsymbol{\psi}), \tag{15}$$

where $\boldsymbol{r}$ denote missing data indicators. $\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{\psi}$ are respectively responses, independent variables and model coefficients. Missing data mechanism is called as "missing at random" (MAR) if, for all $\boldsymbol{y_{obs}}$, $\boldsymbol{x}$ and $\boldsymbol{\psi}$,

$$p(\boldsymbol{r}|\boldsymbol{y_{obs}}, \boldsymbol{y_{miss}}, \boldsymbol{x}, \boldsymbol{\psi}) = p(\boldsymbol{r}|\boldsymbol{y_{obs}}, \boldsymbol{x}, \boldsymbol{\psi}); \tag{16}$$

i.e., if $p(\boldsymbol{r}|\boldsymbol{y_{obs}}, \boldsymbol{y_{miss}}, \boldsymbol{x}, \boldsymbol{\psi})$ is a constant function of $\boldsymbol{y_{miss}}$ (Daniels and Hogan (2008) [5]). This is exactly the case in our data: the only dropout reason in this study is the write-off process of credit card issuing bank, denoted as $y_{iwk} = 3$, where w is the write-off time. That is, missing indicator, $r_{i,w+m,k}$, at time $w + m$ equals to 1 for all $m > 0$ irrespective of $y_{i,w+m,k}$.

With MAR assumption, missing mechanism is ignorable and thus inference is based on the observed-data posterior of all parameters, $\boldsymbol{\psi}$. There are two methods to sample from observed-data posterior, *augmented Gibbs* (Tanner and Wong (1987) [24]) and *nonaugmented Gibbs* (Segawa (2005) [21]). We take the *nonaugmented Gibbs* method in this study as the computational disadvantages associated with *augmented Gibbs.*

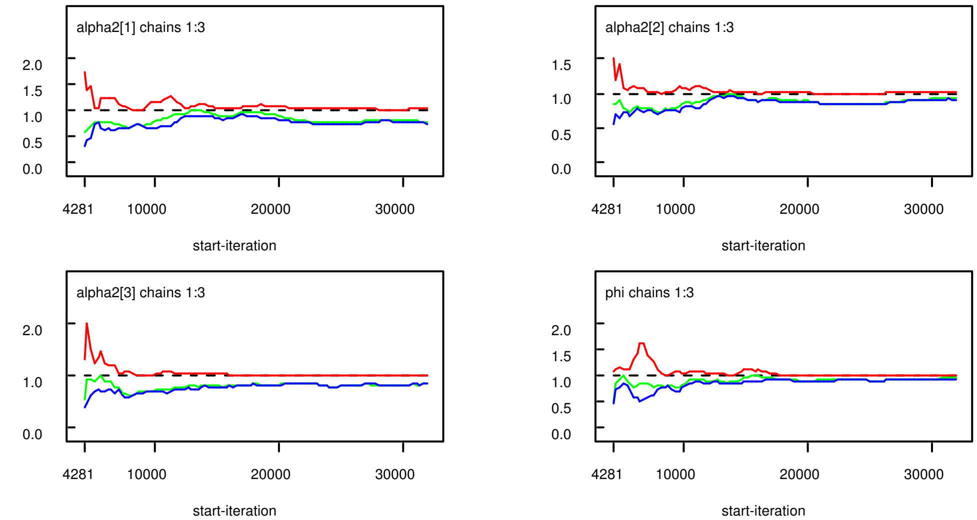


FIGURE 4. BGR statistic for convergence checking

4.2. **Convergency Check, Model Justification and Results.** MCMC is introduced as above to create a Markov chain whose stationary distribution is the same as the posterior distribution of target parameters. If a lot of samples are extracted from the chain, they should converge to the posterior distribution. In this study we simultaneously run three chains with dispersed initial values. The total number of samples is 180,000 ($= [50000 + 10000] \times 3$) for each analysis. Here the length of burn-in is taken as 50,000 for each chain, which can be shorten as analysis in this study finds convergence is usually reached around 10,000 iterations.

4.2.1. *Convergency Check.* In literature, there are a few methods to decide if convergence has been reached. Gelman-Rubin convergence statistic (BGR thereafter), as modified by Brooks and Gelman (1998) [3], is used in this study. BGR statistics are plotted in Figure 4. This BGR method calculate the width of the central 80% interval of the pooled runs, shown in green line, and average width of the 80% intervals within the individual runs, shown as blue line. Their ratio R (=pooled/within) is represented as red line. For plotting purpose the pooled and within interval width are normalized to have an overall maximum of one. The statistics are calculated in bins of length 50: R would generally be expected to be greater than 1 if the starting value are suitably over-dispersed. Brooks and Gelman emphasize that one should be concerned both with convergence of R to 1, and with convergence of both the pooled and within interval widths to stability. For all 25 parameters estimated in the model, BGR plots show R converges to 1, and green and blue lines stabilize after 20,000 iterations. BGR plots for only four parameters are reported in Figure 4 for the sake of conciseness.

4.2.2. *Model Selection.* To illustrate the supereminence of the model employed in this study, we present four reference models with increasing complexity, starting from a basic linear regression model. Model comparisons and selections are based on the Deviance Information Criteria (DIC) (see Spiegelhalter et al. (2002) [23]

for details). DIC is intended as a generalization of Akaike Information Criterion (AIC) for hierarchical models. Let's assume $\mathbf{y}$ and $\boldsymbol{\psi}$ represent the whole collection of data and model parameters, respectively. Spiegelhalter et al. defined *effective* number of parameters in a model, p_D, as

$$p_D = E_{\boldsymbol{\psi}|\mathbf{y}}[D(\boldsymbol{\psi})] - D(E_{\boldsymbol{\psi}|\mathbf{y}}[\boldsymbol{\psi}]) = \overline{D} - D(\overline{\boldsymbol{\psi}}), \tag{17}$$

where $D(\boldsymbol{\psi})$ is the deviance function, $D(\boldsymbol{\psi}) = -2log\ f(\mathbf{y}|\boldsymbol{\psi}) + 2log\ h(\mathbf{y})$. Here, $f(\mathbf{y}|\boldsymbol{\psi})$ is the likelihood function of the data and $h(\mathbf{y})$ is the likelihood function evaluated at the observed proportions. To see how p_D represents number of effective parameters, expand $D(\boldsymbol{\psi})$ around estimates $\overline{\boldsymbol{\psi}}$

$$\begin{aligned} D(\boldsymbol{\psi}) &\approx D(\overline{\boldsymbol{\psi}}) - (\boldsymbol{\psi} - \overline{\boldsymbol{\psi}})^T L''_{\boldsymbol{\psi}}(\boldsymbol{\psi} - \overline{\boldsymbol{\psi}}) \\ &\approx D(\overline{\boldsymbol{\psi}}) + \chi^2_p, \end{aligned} \tag{18}$$

where the second step is based on Bayesian Central Limit Theorem. Taking expectations with respect to the posterior distribution of $\boldsymbol{\psi}$ for equation (18) and substituting into equation (17) result in

$$p_D = E_{\boldsymbol{\psi}|\mathbf{y}}[D(\boldsymbol{\psi})] - D(\overline{\boldsymbol{\psi}}) \approx p. \tag{19}$$

The above shows the p_D will be approximately the true number of parameters.

Posterior expectation of deviance function, $\overline{D} = E_{\boldsymbol{\psi}|\mathbf{y}}[D(\boldsymbol{\psi})]$, can be taken as indicator of the fit of a model. Similar to AIC, DIC statistic summarizes goodness of fit and parsimoniousness of models within Bayesian context by setting

$$DIC = \overline{D} + p_D. \tag{20}$$

The models with small DIC are preferred models (smaller the better). Table 3 reports the $\overline{D}$, p_D, and DIC score for the four reference model and the proposed three-level hierarchical generalized linear model. For all the models, our results are based on three parallel MCMC sampling chains of 10,000 iterations each, following a 50,000-iteration burin-in period.

I begin with a simple generalized linear model (Model I), which turns out to have a large total DIC score, mainly contributed by poor fit of the model. In Model II, the model is extended to a two-level model: item level (M0 accounts, M1 accounts and M2+ accounts) latent traits are explained by cardholders' credit worthiness at a time point in level-1 model; and in turn cardholders' credit worthiness at a time point is attributed to time-level KPIs cardholder-level KPIs together in level-2 model. The model is essentially a generalized linear model with random effect coming from cardholder-level and time-level variations. From Table 3, we can find that the Model II improves model fit significantly ($\overline{D}$ drops from 1464.26 in Model I to 671.01 in Model II) with acceptable cost on model parsimonious. Model III extends Model II by incorporating autoregressive error term on level 2, which is not successful as $\overline{D}$ higher from 671.01 in Model II to 843.84. Model IV reserves AR(1) feature but introducing the third level structure with KPI MOB shifting from level 2 to level 3 as factor effecting level 2 parameters. However, without level 3 error term, we can substitute Model IV level 3 equation into level 2 equation and reach a 2-level model. Model IV provides a very small total DIC score (602.74) in comparison to the first three models, suggesting 3-level structure is preferred to other simpler structures. Model V, the proposed model, is fully parameterized for

[6] In Table 3 we omit subscription i, j, k for the sake of brevity.

TABLE 3. Model Selection for the Credit Card Accounts Data

Model	Model Specification [6]	$\overline{D}$	P_D	DIC
Model I				
- Level 1	$y^* = \beta_0 + \beta_1 CMX12 + \beta_2 UE + \beta_3 MOB + \epsilon^*$			
- Level 2	N.A.			
- Level 3	N.A.	1464.26	4.90	1469.2
Model II				
- Level 1	$y^* = \beta_0 + \beta_1 \theta + \epsilon^*$			
- Level 2	$\theta = u_0 + u_1 CMX12 + u_2 UE + u_3 MOB + \epsilon_\theta$			
- Level 3	N.A.	671.01	342.13	1013.14
Model III				
- Level 1	$y^* = \beta_0 + \beta_1 \theta + \epsilon^*$			
- Level 2	$\theta = u_0 + u_1 CMX12 + u_2 UE + u_3 MOB + \epsilon_\theta$			
	$\epsilon_\theta \sim AR(1)$			
- Level 3	N.A.	843.84	217.74	1061.58
Model IV				
- Level 1	$y^* = \beta_0 + \beta_1 \theta + \epsilon + \epsilon^*$			
- Level 2	$\theta = u_0 + u_1 CMX12 + u_2 UE + \epsilon_\theta$			
	$\epsilon_\theta \sim AR(1)$			
- Level 3	$\mathbf{u} = \mathbf{v_0} + \mathbf{v_1} MOB$	310.03	292.70	602.74
Model V				
- Level 1	$y* = \beta_0 + \beta_1 * \theta + \epsilon + \epsilon*$			
- Level 2	$\theta = u_0 + u_1 * CMX12 + u_2 * UE + \epsilon_\theta$			
	$\epsilon_\theta \sim AR(1)$			
- Level 3	$\mathbf{u} = \mathbf{v_0} + \mathbf{v_1} * MOB + \boldsymbol{\epsilon}_u$			
	$\boldsymbol{\epsilon}_u \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{u})$			
	$\boldsymbol{\Sigma}_\mathbf{u}{}^{-1} \sim Wishart(\boldsymbol{\Omega}, 3)$	188.63	184.48	373.11

3-level structure and is shown as most efficient model among the five (with total DIC of 373.11).

4.2.3. *Estimation Results.* The estimation results for the proposed model are presented in Table 4 (on the left-hand-side under panel "Model V"). For comparison, the estimated results for the simple linear model are also presented in the same table (on the right-hand-side under panel "Model I"). It reports posterior mean, standard deviation and 95% confidence interval for each parameter in the model. We use horizontal solid line to segment parameters into three sections: from item level parameters in the top section to between-subject level parameters in the bottom section. Only item level section is shown for simple generalized linear model.

For both models, most item level parameters, αs and βs are statistically significant. However, the interpretations of these parameters are totally different in the two models. In the simple linear model, $\beta_1, \beta_2, \beta_3$ are risk loadings for *customer*

TABLE 4. Simple Generalized Linear and Multilevel Analysis for the Credit Card Accounts Data

Model V	Mean	SD	2.50%	97.50%	Model I	Mean	SD	2.50%	97.50%
$\alpha_{k=1,c=2}$	29.26	5.33	19.50	40.08	α_2	2.46	0.13	2.21	2.73
$\alpha_{k=2,c=2}$	24.19	6.10	13.47	36.27	β_0	-0.12	0.07	-0.25	0.02
$\alpha_{k=3,c=2}$	22.04	6.27	12.45	36.51	β_1	0.23	0.14	-0.05	0.49
$\beta_{0,k=2}$	-6.13	1.99	-10.46	-2.93	β_2	0.35	0.15	0.05	0.65
$\beta_{0,k=3}$	-10.01	4.05	-19.14	-3.05	β_3	0.01	0.02	-0.03	0.05
$\beta_{1,k=2}$	0.77	0.55	0.09	2.25					
$\beta_{1,k=3}$	3.59	1.16	1.44	5.72					
τ_ϵ $(1/\sigma_\epsilon^2)$	0.01	0.01	0.00	0.03					
ϕ	0.61	0.18	0.26	0.90					
τ_θ $(1/\sigma_\theta^2)$	0.08	0.07	0.01	0.28					
v_{00}	2.42	1.71	-1.02	5.98					
v_{01}	1.24	1.52	-2.18	4.17					
v_{02}	1.63	1.63	-1.66	4.91					
v_{10}	0.05	0.17	-0.30	0.38					
v_{11}	-0.04	0.25	-0.51	0.47					
v_{12}	-0.52	0.31	-1.21	0.03					
$\Sigma_u[11]$	3.13	2.48	0.28	9.60					
$\Sigma_u[12]$	-0.09	2.06	-4.44	4.12					
$\Sigma_u[13]$	-0.07	2.04	-4.38	4.19					
$\Sigma_u[21]$	-0.09	2.06	-4.44	4.12					
$\Sigma_u[22]$	4.16	3.37	0.34	12.78					
$\Sigma_u[23]$	2.01	2.62	-1.75	8.54					
$\Sigma_u[31]$	-0.07	2.04	-4.38	4.19					
$\Sigma_u[32]$	2.01	2.62	-1.75	8.54					
$\Sigma_u[33]$	4.14	3.30	0.36	12.68					

level number of months accounts delinquent in last 12 months (CMX12), *unemployment rate* (UE) and *months on book* (MOB) (after demeaning and scaling down), respectively. The estimation results show that higher CMX12 and UE will significantly increase the probability of deterioration or write-off. The other variable MOB on the other hand has no clear influence on deterioration rate and write-off rate, with mean value 0.01 and standard deviation 0.02. This is somewhat not consistent with our preliminary observation from Table 2, recall that average MOB increased significantly around economic downturn. We will discuss in details on this point along with our findings from Model V later.

For Model V, β_{0k}s and β_{1k}s are all statistically significant, except $\beta_{1,k=2}$. With $\beta_{1,k=1}$ fixed at 1, a value of 3.59 on $\beta_{1,k=3}$ suggesting, for M2+ accounts ($k = 3$), probability of deterioration increase much faster when the underlying credit worthiness of cardholder worsen. On the other hand, estimated value 0.77 and standard deviation 0.55 for $\beta_{1,k=2}$ implies that the chance of deterioration for M1 accounts is not linked to cardholder latent trait as close as M0 and M2+ accounts. This could be attributed to the collection efforts (i.e. dunning SMSs and calls which are proved very effective for accounts in early bucket) put in by credit operations department of the issuing bank. Unfortunately due to the very restricted availability of bank's data, the collection productivity data is not factored in the model, which surely be an interesting area of extension for the proposed model in this study.

The precision parameter τ_ϵ, is very low, only 0.01, suggesting a high σ_ϵ^2 that represents variations from random effects. Low precision parameter on item level indicate that the cardholder latent trait estimated based on only tow variables (CMX12 and UE) are not sufficient to represent cardholders' risk levels. This can definitely be improved assuming more key performance indicators are available to the study, however, for this very new attempt on hierarchical generalized linear credit risk model, we focus more on the benefits brought by the model structural advantages.

In the time level model, autoregressive parameter ϕ is estimated as 0.61 with standard deviation 0.18. It appears that a cardholder's latent creditworthiness θ_{ij} is related to two characteristics driving the cardholder's longitudinal data pattern, namely the last-period latent trait $\theta_{i,j-1}$ and the rate of deterioration on time-level KPIs $CMX12$ and UE. To see this clearly, we can take expectation on and rewrite equation (7) and (9) into

$$\begin{aligned} E(\theta_{ij}) = \phi\theta_{i,j-1} &+ u_{0i}(1-\phi) + u_{1i}(CMX12_{ij} - \phi CMX12_{i,j-1}) \\ &+ u_{2i}(UE_j - \phi UE_{j-1}). \end{aligned} \tag{21}$$

This finding provides an explanations to the pattern of *straight rollers* observed from credit operations of credit card issuers, a situation where cardholders newly flow into delinquency and deteriorate with increasing speed towards write-off, which makes a very high proportion of card issuers' total credit loss.

Similar to random effects variance σ_ϵ^2 in item level equation, σ_θ^2 is very high as reflected by a small estimated value 0.08 on precision parameter τ_θ. We interpret it as that unknown between-cardholder level KPIs excluded from time level mean equation enter the error term. Note that very limited KPIs are available, we are confident the model, with wider selection of KPIs, can perform even better in terms of model fit and providing theoretic support to realities.

For between-cardholder level parameters $\boldsymbol{\upsilon}$, all parameters except υ_{12} are statistically insignificant. As introduced in equation (11), between-cardholder model is specified as

$$\begin{aligned} u_{0i} &= \upsilon_{00} + \upsilon_{10} MOB_i + \epsilon_{u_{0i}} \\ u_{1i} &= \upsilon_{01} + \upsilon_{11} MOB_i + \epsilon_{u_{1i}} \\ u_{2i} &= \upsilon_{02} + \upsilon_{12} MOB_i + \epsilon_{u_{2i}} \end{aligned} \tag{22}$$

Parameter υ_{12} is estimated as -0.52 with standard deviation 0.31, suggesting high value of MOB suppress u_{2i}, which is in turn the regression coefficient of risk factor UE_j in equation (7). To highlight the importance of υ_{12}, we discuss the estimated posterior results of two cardholders (55 and 111) at 2003 Q3 in the following paragraphs.

Table 5. Characteristics of Cardholder 55 and 111

Variables [7]	Cardholder 55	Cardholder 111
Time	2003 Q3	2003 Q3
Item	M2+	M2+
$CMX12_{ij}$	14	14
UE_j	4.80%	4.80%
MOB_i	78	9

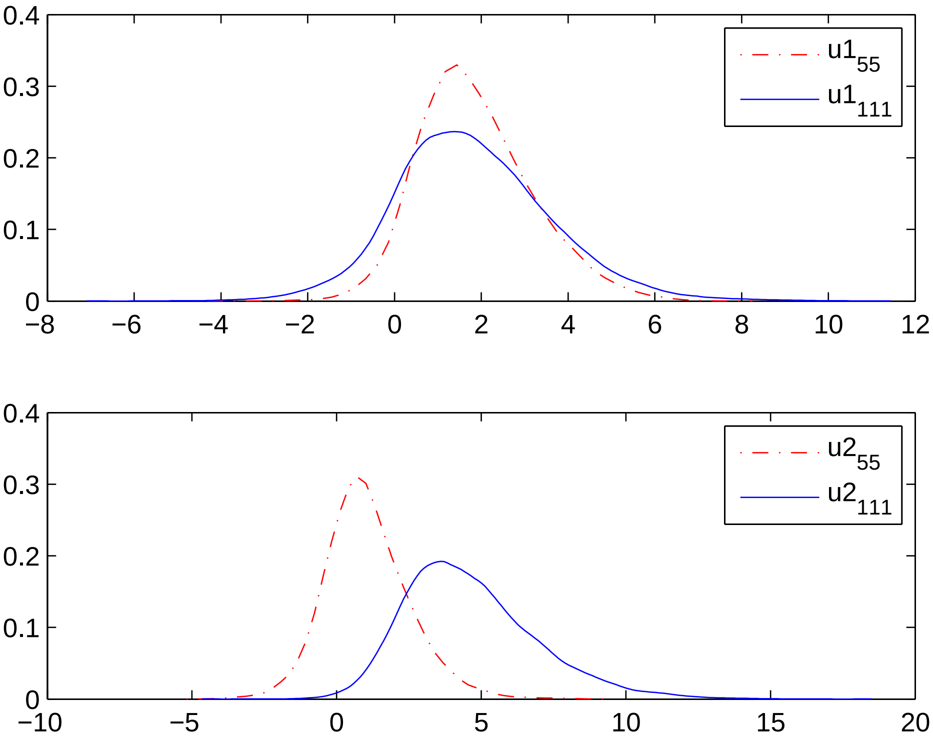
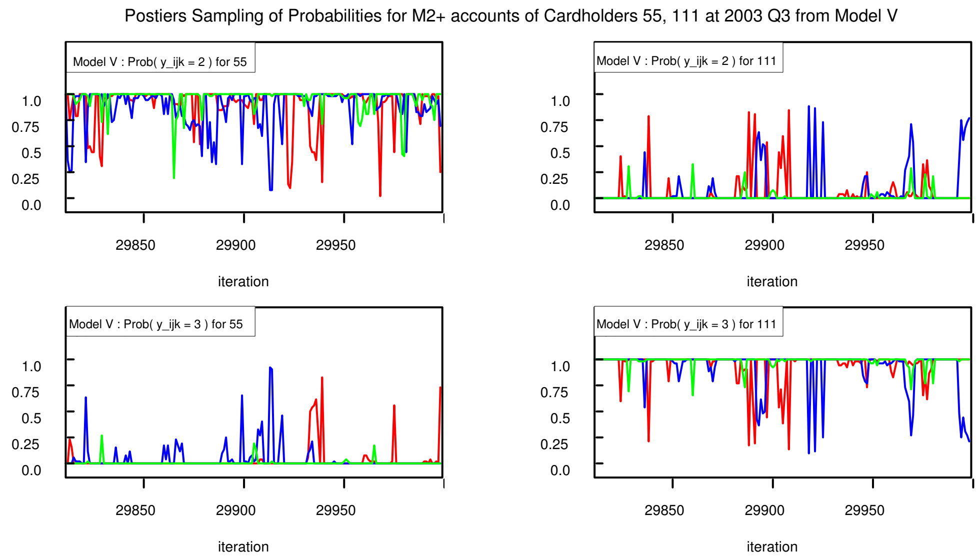


Figure 5. Posteriors densities of u_1 and u_2 for cardholder 55 and 111

To investigate the actual effect of υ_{12} on $u_2 i$, the posterior densities of u_1 and u_2 for cardholder 55 and 111 are plotted. The posterior densities of u_1, coefficients for $CMX12_{ij}$ in time level model, have similar mean except the density plot for cardholder 111 ($MOB = 9$) is with fatter tail. This suggests credit worthiness changes along with $CMX12_{ij}$ to almost same extent, irrespectively cardholders' MOB. On the other hand, for u_2 the density plot of cardholder 111 shift towards

[7]In Table 5 variables $CMX12$, UE and MOB are reported without demeaning and scaling though we did so in actual estimation.



Estimated Postiers Probabilities for M2+ accounts of Cardholders 55, 111 at 2003 Q3 (Model V vs Model I)

Model V	Mean	SD	2.50%	97.50%	Model I	Mean	SD	2.50%	97.50%
P(y_ijk=1) for 55	5.6%	0.16	0.00	0.63	P(y_ijk=1) for 55	41.2%	0.04	0.33	0.50
P(y_ijk=2) for 55	90.6%	0.19	0.25	1.00	P(y_ijk=2) for 55	47.7%	0.03	0.41	0.54
P(y_ijk=3) for 55	3.8%	0.13	0.00	0.51	P(y_ijk=3) for 55	11.1%	0.02	0.07	0.16
P(y_ijk=1) for 111	0.0%	0.00	0.00	0.00	P(y_ijk=1) for 111	42.3%	0.05	0.33	0.52
P(y_ijk=2) for 111	4.7%	0.15	0.00	0.60	P(y_ijk=2) for 111	47.1%	0.03	0.40	0.54
P(y_ijk=3) for 111	95.3%	0.15	0.40	1.00	P(y_ijk=3) for 111	10.7%	0.02	0.07	0.16

FIGURE 6. Posteriors Probabilities of cardholder 55 and 111 at 2003 Q3 responding to M2+ accounts

right as compare with cardholder 55. In addition the density plot for cardholder 111 apparently skews to right while cardholder has a almost symmetric curve. The finding says that the cardholders with high MOB are more robust to social and economic shocks like high unemployment rate, whereas length of vintage is not able to help distinguish bad/good cardholder based on past 12-month performance.

The finding can be used to demonstrate the advantage of hierarchical credit risk model over the conventional linear models. Upper panel of Figure 6 plots the posterior sampling of $Prob(y_{ijk} = 2)$ and $Prob(y_{ijk} = 3)$ from Model V for both cardholder 55 and 111. We can find that Model V assign very high probability to $y_{ijk} = 2$ for cardholder 55, but evaluate highest probability for $y_{ijk} = 3$ rather than $y_{ijk} = 2$ for cardholder 111. For comparison, the estimated posterior probabilities for both Model V and Model I are reported in the lower panel of Figure 6. It's clearly shown that Model I failed to recognize cardholder 111 as a high-risk customer and can not distinguish him from cardholders with longer relationship with the bank, as Model I assigned almost identical probabilities to both cardholders. The failure is expected as we have seen β_3 of Model I in Table 4 is not statistically different

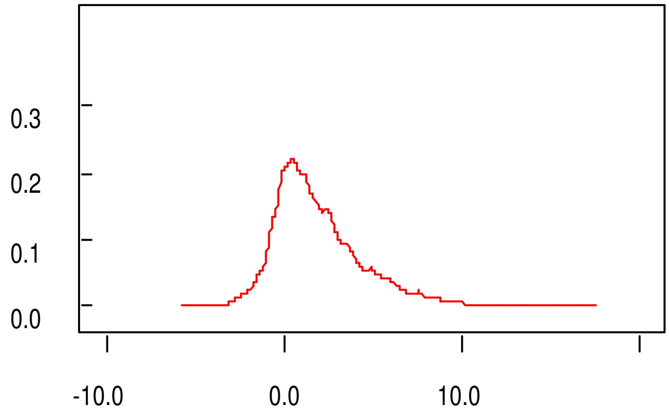


FIGURE 7. Posterior density of parameter $\Sigma_u[23]$

from 0. However, Model V successfully separate cardholder 111 out by tagging him with a high chance of write-off ($Prob(y_{ijk} = 3) = 95.3\%$).

Covariance matrix Σ_u of Model V are reported in 4 as well. Interestingly we find that $\Sigma_u[23]$ is asymmetric and skewed to positive domain, as shown in Figure 7. We can interpret this as a sign of positive correlation of regression coefficients u_1 and u_2. This finding is consistent with intuition as in that cardholder with personal financial issue already are those who are more vulnerable to social and economic shocks. The asymmetry of the posteriors in the figure also suggests traditional statistical inference based on asymptotic normality and approximate standard errors will not be very accurate.

5. Conclusion

In this paper, we have presented a general credit risk model that can be applied to complex data structure like hierarchical and longitudinal data. In consumer credit risk management area, hierarchical structure is a nature way of organizing and presenting credit card account information, however to the best of our knowledge no research has been done to introduce the hierarchical model into credit risk area though such models are popular in educational and biomedical studies. The paper provides opportunities to relook at credit risk model from an innovative perspective.

Although with many restriction in data availability and usage, the proposed model outperform four reference models with incremental complexity, starting from conventional linear ordinal logit model. Especially, the proposed model manifestly revealed the disadvantages of traditional credit risk models that do not properly account for (1) analytic structures of latent trait and explanatory variables; (2) multiple layers of associations between observations; and (3) random variable distributions.

Possible extension of the proposed model can be done in a few ways. With Bayesion MCMC estimation approach, more complicated model can be fit without huge difficulties faced by traditional maximum likelihood method. For example, Guo and Carlin 2004 [13] jointly model longitudinal and survival data from AIDS clinical trial comparing two different treatments. This can also be introduced into credit risk model area, given the similarity in the data features. Further, seasonality and financial crisis analysis can be accommodated into this model.

Acknowledgment. A lot of appreciation goes to Manu Singhal, T.P. Arvind Kumar from Citibank Singapore Limited. They expressed interest in our work and allowed me to use samples from the bank's credit card portfolio, which gave me an opportunity to examine our models in the real credit scenario. As agreed, the opinions expressed in the paper are not those of the Citibank Singapore Limited's. Neither the models discussed here nor the data used here should be taken as Citibank Singapore Limited's business practice and actual portfolio.

References

1. J. H. Albert and S. Chib, *Bayesian analysis of binary and polychotomous response data*, Journal of the American Statsitcal Association **88** (1993), 669–679.
2. R.B. Avery, P.S. Calem, and G.B. Canner, *Consumer credit scoring: Do situational circumstances matter?*, Journal of Banking & Finance **28** (2004), 835–856.
3. S. P. Brooks and A. Gelman, *Alternative methods for monitoring convergence of iterative simulations*, Journal of Computational and Graphical Statistics **7** (1998), 434–455.
4. P. Burns and C. Ody, *Validation of consumer credit risk models*, Conference Summary, Federal Reserve Bank of Philadelphia & Wharton School's Financial Institutions Center, 2004.
5. M. J. Daniels and J. W. Hogan, *Missing data in longitudinal studies: strategies for bayesian modeling and sensitivity analysis*, Chapman and Hall, 2008.
6. F. W. M. de Andrade and A. L. Sicsú, *A credit risk model for consumer loans portfolios*, Latin American Business Review **8** (2007), 75–91.
7. F. W. M. de Andrade and L. C. Thomas, *Structural models in consumer credit*, Working Paper (2004).
8. Editorial, *Retail credit risk management and measurement: An introduction to the special issue*, Journal of Educational and Behavioral Statistics **28** (2004), 721–725.
9. H. Goldstein, G. Bonnet, and T. Rocher, *Multilevel structure equation models for the analysis of comparative data on educational performance*, Journal of Educational and Behavioral Statistics **32** (2007), 252–286.
10. H. Goldstein and W. Browne, *multilevel factor analysis models for continuous and discrete data*, working paper (2005).
11. Harvey Goldstein, *Multilevel Statistical Models*, John Wiley & Sons Inc, New York, 2002.
12. D. Gross and N.S. Souleles, *An empirical analysis of personal bankruptcy and delinquency*, Review of Financial Studies **15** (2002), 319–347.
13. X. Guo and P. Carline, *Seperate and joint modelling of longitudinal and even time data using standard computer packages*, The American Statistician **58** (2004), 1–9.
14. G. M. Gupton, C. C. Finger, and M. Bhatia, *Creditmetrics-technical document*, J.P. Morgan, 1997.
15. D. J. Hand, *Modelling consumer creidk risk*, Journal of Management Mathematics **12** (2001), 139–155.
16. S. Kealhofer, *Managing default risk in portfolio of derivatives*, Deriivative Credit Risk: Advances in Measurement and Management, Renaissance Risk Publications, 1995.
17. OCC, *Credit card lending*, The Office of the Comptroller of the Currency, Washington, 1998.
18. Basel Committee on Banking Supervision, *A new capital adequacy framework*, 1999.
19. Credit Suisse Financial Products, *Creditrisk+: A credit risk management framework*.
20. D. B. Rubin, *Inference and missing data*, Biometrika **63** (1976), 581–592.
21. E. Segawa, *A growth model for multilevel ordinal data*, Journal of Educational and Behavioral Statistics **30** (2005), 369–396.
22. J. D. Singer, *Using sas proc mixed to fit multilevel models, hierarchical models, and individual growth models*, Journal of Educational and Behavioral Statistics **24** (1998), 323–355.

23. D. Spiegalhalter, N. G. Best, B. P. Carline, and V. Linde, *Bayesian measures of model complexity and fit withdiscussion*, Journal of Royal Statistical Society **64** (2002), 583–640.
24. M. A. Tanner and W. H. Wong, *The calculation of posterior distribution by data augmentation withdiscussion*, Journal of the American Statistical Association **82** (1987), 528–550.
25. K. Wang, K. Yau, A. Lee, and G. McLachlan, *Multilevel survival modelling of recurrent urinary tract infections*, Computer Methods and Programs in Biomedicine **87** (2007), 225–229.
26. Y. Zhao Y. Zhao and I. Song, *A dynamic model for repayment behavior of new customers in the credit card market*, working paper (2006).