

Quantitative Research on the Internet Consumption Financial Model Based on the Difference Analysis of Risk Score Tracks

Shih-Chieh Lin¹, Ying-Li Lin², Tzu-Ting Chao³

¹Department of Business Administration, Asia University

²Department of Finance, Asia University

³Department of Business Administration, Asia University

Email: a0911399568@gmail.com

Abstracts

The main research purpose of this article is to sort out consumer information without specific personal data and use the target company's neural network risk control decision model which is based on the risk scoring trajectory. We find the main interpretation variables, and then verify the direction and extent of the target's influence to verify empirical results and help the target company make a correct decision analysis on the credit applicants' loans. The final empirical results show that the analysis results of the minimum square method and the analysis results of the truncated regression model can be found: the gender, population identity characteristics, consumer performance capabilities, consumer risk of trust, consumer preferences, consumer behavior characteristics, behavioral characteristics, number of financial institutions the consumer have applied for loans, credit litigation, and loan period were the nine main explanatory variables affecting the composite score. Regarding the influence of each variable on the comprehensive score: the impact of four index variables such as identity characteristics, contract performance ability, consumption preference, and behavioral characteristics on the comprehensive score is positive; it means that with higher identity feature score, contract performance ability, consumer preference score, and behavioral characteristic score, the comprehensive score increases.

Keywords: Internet consumption financial model, Risk score trajectory, Vector autoregression.

1. Introduction

In today's digital era, new financial technologies and big data are accelerating the development of financial transactions. Also with the rise of e-commerce transactions, the financial industry has gradually realized that due to disruptive innovations driven by electronic and third-party payments, technology-based financial-like banking services can be seamlessly integrated into any scenario [8]. With the rise of the Internet, information consumption habits have changed. The

standard process of consuming news through traditional mass media (e.g. newspapers, radio and television) is now replaced or supplemented by news consumption through online sources [9].

Consumer finance refers to a set of flexible financial service models for consumer loans provided by consumer finance companies to consumers from all walks of life. Its advantages include small credit limits, fast approval times, no need for collateral, and high flexibility in service methods. It is flexible and has a short loan period. In addition, a good risk control system also plays an important role in handling consumers' credit risks. Therefore, consumer finance must reduce risks to the greatest extent, and therefore must also rely on consumers' personal credit information in the risk system. The so-called personal credit reporting refers to a personal credit reporting agency established in accordance with the law, which can collect consumers' personal credit information, and can provide consumers' personal credit data according to consumers' requirements to inquire and evaluate the strength and breadth of services; professional Consumer personal credit reporting agencies also provide credit inquiry and other services to each other and other consumer financial service companies, and the development of network technology cannot deny that it also promotes the service process of the consumer finance industry. According to economic theory, developing a consumer credit scoring model can help solve information asymmetry problems such as adverse selection in the lending market, thereby improving the efficiency of both supply and demand sides in the lending market, affecting the monetary policy transmission mechanism, and has profound policy significance. The main purpose of this study is to apply the target company's risk control decision-making model based on the quantitative analysis of risk score trajectory differences, and gradually detect which indicators will mainly affect the main explanatory variables of consumers' comprehensive scores from the actual purchase data, and conduct quantitative analysis. Then verify and interpret the results of the regression empirical analysis and help the target company make correct decisions on whether to approve consumer loan applications for shopping. This study has five chapters. The first chapter is an introduction, explaining the research background, purpose, significance and main findings, and the structure of the full text. The second is a literature discussion and review. It first explains information asymmetry, and then describes the consumer finance operation model. Secondly, the historical research literature is compiled for discussion, and finally the relationship between consumer finance and credit scores is explained and concluded. Section 3 is the research method, explaining the theoretical model, conceptual framework, empirical methods, variable selection and basic statistical analysis, as well as conclusion; Section 4 is the empirical results, the sources of research data and empirical analysis results; Section 5 is the conclusion and policy suggestions, summarizing the main conclusions and implications and proposing policy suggestions.

2. Literature Review

Information asymmetry means that the parties involved in a transaction have different information that can affect the transaction. The seller may have more information about the item being traded than the buyer, but the reverse may also be true. An example of the former can be found in the sale of used cars, where the seller knows more about the vehicle than the buyer does. An example of the latter is health insurance, where the buyer usually has more information. This

argument can be applied to various markets, such as the capital market, which provides a bridge between capital supply and demand at the same time, but at the same time there is also a common problem of information asymmetry. Agency theory was developed in 1976, describing the information asymmetry relationship between the principal and the agent. If there is a lack of information transparency and information symmetry between the two, moral hazard and adverse selection are prone to occur. These two problems are in the financial market. It will also have a profound impact. Applying this argument to this study, there is information opacity between banks and borrowers, and the information is inconsistent. The bank does not even understand the borrower's solvency, leading to adverse selection and ethical issues. These may be one of the problems that trigger debt insolvency. Information economists believe that in the lending market, lenders and borrowers often have information asymmetry (Asymmetric Information) due to institutional factors. Therefore, the type of borrower may not be observable. When the type of borrower cannot be completely observed, it is very likely to cause adverse selection (Adverse Selection) and moral hazard (Moral Hazard) problems.

From the perspective of information economics, moral hazard stems from the existence of asymmetric information. Specifically, in the operation of commercial banks, one party of economic entities owns information, while the other party does not own the information or has partial information. When one party is more powerful than the other party, having more relevant information, in order to maximize one's own interests, brings the possibility of harm and loss to those with inferior information, thus creating moral hazard. Moral hazard means that after an individual applies for a cash card, due to the asymmetry of information between the bank and the individual, the bank is often unable to effectively monitor the individual. Individuals may engage in behavior that harms the interests of the bank, for example, maliciously overdraft regardless of their own income status and refuse to repay, causing losses to the bank, thus creating a moral hazard problem between the bank and the card holder.

As for the consumer finance market, due to the asymmetric information in the market, consumer credit is also constrained. It may be that the ability of family members to repay debts is reduced, which will cause consumers to be unable to borrow the money they need, or there may be mutual distrust and unequal information between the supply side and the demand side in the market [1][2][3]. Faced with information asymmetry in the consumer finance market, most scholars have different levels of information sharing. Due to adverse selection and moral hazard, for example, the document proposed by [4] reflects the common sharing of information between banks. The higher the level, the higher the amount of consumer loans that each bank in these countries can provide to consumers, and the risks existing in the bank can also be reduced. This phenomenon illustrates that information sharing is an effective channel for reducing moral risks and adverse selection.

To sum up the above, this study will focus on consumer finance, analyze the actual consumer finance model, and apply the risk control decision-making model to collect statistics and analyze the demographic characteristics of consumers, the history of consumer credit, and behavioral records of consumers and consumer historical transaction records, etc. After cleaning and preprocessing the data, the so-called systematic collection and analysis of data are carried out, and then statistical methods are used to verify the model one by one using other measurement

methods such as data conversion and feature selection and sort out the results. Isomorphic behavior patterns and credit characteristics contained in the data are researched and developed to predict risk, and the model is applied to the target company's consumer credit behavior in the future to make relevant risk control decisions, and quantitative analysis of the difference in risk score trajectories of target companies. For example, banks can improve default prediction models to create value by exploring machine learning techniques [7].

3. Research methods

In terms of the use of measurement methods, this study first implements the least squares method to predict, screen and streamline variables, and then uses principal component analysis and Lasso algorithm for verification. The main purpose is to screen out the factors that affect personal credit scores variables, the second is to use truncated regression analysis to actually test the significant impact degree and direction of the main variables that affect the credit score, and the third is to use quantile regression analysis to explore the significant impact of the main variables under different credit score quantiles, extent and direction.

I. Principal Component Analysis

The main function of principal component analysis (PCA) is to reduce the value of variables, while also maintaining the characteristic variables that contribute the most to the variation in the data set as much as possible, and retaining the most important explanatory variables. PCA also uses eigen decomposition of the covariance matrix to sort out the principal components of the data (i.e., eigenvectors) and their weights (i.e., eigenvalues). PCA uses feature quantities to analyze multivariate statistical distributions, and uses the variation in the original data to find out which direction of data values has the greatest impact on the variation. Principal component analysis can reduce the instability of data and avoid collinearity problems. At the same time, when the amount of data is large, it can reduce excessive adaptation in data operations. A small number of variables that explain the most variation is the purpose of using principal component analysis. Therefore, when there is a large amount of variation, principal component analysis can be used to limit the breadth of the data to facilitate subsequent data analysis. Use principal component analysis to display data and variations using the least number of data dimensions.

II. Least Absolute Shrinkage and Selection Operator

The Lasso calculation method can help identify and compare variables before and after streamlining. Through Lasso calculation, the sum of the absolute values of the regression coefficients is less than a certain value, for example, some regression coefficients are converted into the number 0, and the covariance that do not include the corresponding regression coefficients are screened out most efficiently and become feasible. Filter the reduced model for the most consistent signal. Because the Lasso calculation method can automatically optimize the model and automatically perform variable screening. The Lasso calculation method was originally designed and calculated by the least squares method. The Lasso application scenario of the least squares method also briefly highlights the main specificity of Lasso.

It is assumed that a sample includes N types of events, and each event includes p co-variations and an output value. Let y_i be the output value, and $x_i := (x_1, x_2, \dots, x_p)^T$ be the covariation vector of the first case. The objective equation of Lasso is:

$$\min_{\beta_0, \beta} \left\{ \frac{1}{N} \sum_{i=1}^N (y_i - \beta_0 - x_i^T \beta)^2 \right\}, \text{ for all } \sum_{j=1}^p |\beta_j| \leq t \quad (1)$$

Where t is a predetermined free parameter that determines the degree of regularization. Assuming to be X a covariation matrix, $X_{ij} = (x_i)_j$, where x_i^T is the row element of X , the above formula can be rewritten as :

$$\min_{\beta_0, \beta} \left\{ \frac{1}{N} \|y - \beta_0 - X\beta\|_2^2 \right\}, \text{ for all } \|\beta\|_1 \leq t \quad (2)$$

where $\|\beta\|_p = \left(\sum_{i=1}^N |\beta_i|^p \right)^{1/p}$ is the standard ℓ^p norm, 1_N a $N \times 1$ vector of dimension 1. because $\hat{\beta}_0 = \bar{y} - \bar{x}^T \beta$,

$$\text{so } y_i - \hat{\beta}_0 - x_i^T \beta = y_i - (\bar{y} - \bar{x}^T \beta) - x_i^T \beta = (y_i - \bar{y}) - (x_i - \bar{x})^T \beta,$$

This is centered on pairs of variables. At this time, the covariance is $(\sum_{i=1}^N x_{ij}^2 = 1)$, this method can remove the influence of the measurement scale.

III. Censored Regression Models

Consider the following latent variable regression model:

$$y_i^* = x_i' \beta + \sigma \epsilon_i, \quad (3)$$

where σ is the scale parameter, which can be estimated together with the coefficients. When the residuals are normal, the Tobit truncated regression model is set as:

$$y_i = \begin{cases} 0 & \text{if } y_i^* \leq 0 \\ y_i^* & \text{if } y_i^* > 0 \end{cases} \quad (4)$$

When y_i^* negative and y_i the value is 0, it is said to be left censored at 0 .

When there are cutoff points on both the left and right sides, the model can be set as:

$$y_i = \begin{cases} \underline{c}_i & \text{if } y_i^* \leq \underline{c}_i \\ y_i^* & \text{if } \underline{c}_i < y_i^* \leq \bar{c}_i \\ \bar{c}_i & \text{if } \bar{c}_i < y_i^* \end{cases} \quad (5)$$

where $\underline{c}_i, \bar{c}_i$ is the fixed cutoff point reflected. If there is no left truncation, $\underline{c}_i = -\infty$. And $\bar{c}_i = \infty$ when means there is no right truncation point. Generally speaking, a typical Tobit model

is $\underline{c}_i = 0$ and $\bar{c}_i = \infty$ A special case of. This study uses the maximizing logarithmic approximation function to solve the parameters β, σ , such as (6):

$$l(\beta, \sigma) = \sum_{i=1}^N \log f((y_i - x_i' \beta) / \sigma) \cdot 1(\underline{c}_i < y_i < \bar{c}_i) + \sum_{i=1}^N \log \left(F((\underline{c}_i - x_i' \beta) / \sigma) \right) \cdot 1(y_i = \underline{c}_i) + \sum_{i=1}^N \log(1 - F((\bar{c}_i - x_i' \beta) / \sigma)) \cdot 1(y_i = \bar{c}_i) \quad (6)$$

Among them f, F are the ϵ density function and the cumulative probability function.

This study sets up three residual estimation models, under the standard normal; $E(\epsilon) = 0, \text{var}(\epsilon) = 1$ under the logistic setting $E(\epsilon) = 0, \text{var}(\epsilon) = \pi^2/3$; under the setting of asymmetric extreme values

$$E(\epsilon) \approx -0.5772 (\text{Euler's constant}), \text{var}(\epsilon) = \pi^2/6.$$

IV. Quantile Regression Models

The quantile regression model is the theoretical model constructed by [5]. According to regression analysis, the traditional correlation between independent variables and dependent variables can be highlighted in their conditional expected values, and the calculated regression model can predict the dependent variable through the independent variables. As for the purpose of quantile regression, it is to explore the correlation between conditional quantiles between independent variables and dependent variables. Research shows that regression models can also predict the conditional quantiles of dependent variables from independent variables. Different from traditional regression models that can only show the trend of the dependent variable moving towards the center, the quantile regression model can further infer the conditional probability range coverage of the dependent variable. Quantile regression is also one of the statistical methods without a parameter.

First assume that the random variable Y and its probability distribution function is

$$F(y) = \text{Prob}(Y \leq y) \quad (7)$$

Among them $0 < \tau < 1$, Y the τ quantile can be defined as the smallest y and satisfies $F(y) \geq \tau$:

$$Q(\tau) = \inf\{y: F(y) \geq \tau\} \quad (8)$$

Given a set Y of n observations. The traditional empirical distribution function is:

$$F_n(y) = \sum_k 1(Y_i \leq y) \quad (9)$$

where $1(z)$ is a pointer variable that is 1 if the value is 1 otherwise .0

$$\text{At this time } Q_n(\tau) = \inf\{y: F_n(y) \geq \tau\} \quad (10)$$

The optimization solution can be expressed as:

$$Q_n(\tau) = \operatorname{argmin}_{\xi} \left\{ \sum_{i: Y_i \geq \xi} \tau |Y_i - \xi| + \sum_{i: Y_i < \xi} (1 - \tau) |Y_i - \xi| \right\}$$

$$= \operatorname{argmin}_{\xi} \{ \sum_i \rho_{\tau}(Y_i - \xi) \} \quad (11)$$

in $\rho_{\tau}(u) = u(\tau - 1(u < 0))$ an explanatory variable with vector X_p in the regression, the setting of conditional quantile regression is:

$$Q(\tau|X_i, \beta(\tau)) = X_i' \beta(\tau) \quad (12)$$

where $\beta(\tau)$ is the regression parameter of the quantile. At this time, the conditional quantile regression estimation formula is:

$$\hat{\beta}_n(\tau) = \operatorname{argmin}_{\beta(\tau)} \{ \sum_i \rho_{\tau}(Y_i - X_i' \beta(\tau)) \} \quad (13)$$

In order to evaluate the estimation efficiency of quantile regression, [6] constructed an analysis using a fitness test. Under the setting of linear quantile regression, $Q(\tau|X_i, \beta(\tau)) = X_i' \beta(\tau)$. This study also assumes that the segmented data and coefficient vectors are $X_i = (\mathbf{1}, X_{i1})'$ sum $\beta(\tau) = (\beta_0(\tau), \beta_1(\tau))'$, so it can be further obtained:

$$Q(\tau|X_i, \beta(\tau)) = \beta_0(\tau) + X_{i1}' \beta_1(\tau) \quad (14)$$

This study redefines :

$$\hat{V}(\tau) = \min_{\beta(\tau)} \sum_i \rho_{\tau}(Y_i - \beta_0(\tau) - X_{i1}' \beta_1(\tau)) \quad (15)$$

If the minimization is unrestricted and has only an intercept term, the objective function can be set as :

$$\hat{V}(\tau) = \min_{\beta_0(\tau)} \sum_i \rho_{\tau}(Y_i - \beta_0(\tau)) \quad (16)$$

At this time, the critical value of [6] fitness is:

$$R^1(\tau) = 1 - \hat{V}(\tau) / \hat{V}(\tau) \quad (17)$$

The statistic value is between $R^1(\tau)$ 0 and 1. At the $\tau - th$ quantile, it means that the model is properly fitted.

In order to test whether there is a significant difference between the slope terms between different quantiles, a tenacity test (Slope Equality Testing). [5] established the null hypothesis that all quantile slopes are equal:

$$H_0: \beta_1(\tau_1) = \beta_1(\tau_2) = \dots = \beta_1(\tau_K) \quad (18)$$

several coefficient restrictions in the above formula. $(p - 1)(K - 1)$ Its Wald test statistic obeys the chi-square distribution $\chi_{(p-1)(K-1)}^2$.

V. Variable Selection and Basic Statistical Analysis

This study uses two types of data for empirical analysis. The first is to conduct principal component analysis and regression analysis using 775 personal credit data of the target company collected in 6 months before the COVID-19 epidemic, and 722 effective samples. In this study, the comprehensive score variable is , and there are a total of 24 other independent variables , as shown in Table I, the age variable is the actual age of the customer; this study expresses the qualitative variables quantitatively. For the gender variable, the male value is set to 1 and the female value is 0; for credit variables, the quantity means that if the verification results using the mobile phone are inconsistent, the value is set to 1, and the value is 0 when they are consistent; the mobile phone network access status variable means that the value is when the network access status is abnormal. 1. The normal value is 0; for the mobile phone real-name system variable, the inconsistent value of the verification result is 1 and the consistent value is 0; the anti-fraud risk variable, if the mobile phone number is detected and may have been blacklisted by a third party, the value is 1 , the mobile phone number was not detected to have any risk-related value of 0.The quantitative pointer variables for credit card scenario risks include overall rating Y, identity characteristics, performance capabilities, breach of trust risk, consumption preferences, behavioral characteristics, social influence, and growth potential. Pointer variables in multiple loans are: the industry to which the institution belongs (a total of 7 categories, P2P value is 1, small loan value is 2 , bank value is 3 , consumer finance value is 4 , investment value is 5 , third party value is 6 , leasing value is 7). Quantitative variables such as the number of hit institutions, the number of hit banks, the number of hit consumer finance institutions, the number of hit small loan institutions, the total number of institutional inquiries, the number of institutional inquiries in the past three months, the number of institutional inquiries in the past six months, and business hours. The credit risk is calculated based on the risk score, and if there is no value, it is 0; the credit risk is calculated based on the total number of lawsuits involved, and if there is no value, it is 0.

Table I Variable names and variable codes

variable name	variable name	variable name
age (age)	Risk of breach of trust (Risk of dishonesty)	Hit number of consumer finance (Hit consumer finance)
gender (gender)	consumption preference (Consumer preference)	Number of hit small lending institutions (Number of small loan institutions hit)
credit (Good faith)	Behavioral characteristics (Behavioral characteristics)	Total number of institutional inquiries (The total number of institutional inquiries)
Mobile phone access to the Internet (Mobile phone access)	social influence (Social influence)	Number of institutional inquiries in the past three months (Inquiries of institutions in the past three months)
Mobile phone real-name system	Growth potential (Growth potential)	Number of institutional inquiries in the past six months (Inquiries of institutions in the past six months)

(Phone real-name system) Anti-fraud risk (Anti-fraud risks)	Industry to which the institution belongs	credit risk (Good faith risk)
Identity characteristics Identity characteristics) Performance ability (Performance capacity)	Number of hits (Number of hits) Number of banks hit (Bank hits)	credit involving lawsuit (Good faith involved) business hours (Business hours)

4. Empirical Result

I. Principal Component Analysis Results

This study uses the cumulative contribution ratio (Cumulative Proportion) of the eigenvalues to determine how many principal components need to be taken. When the eigenvalues are greater than 1, the principal components can be selected. As shown in Table II, the principal component analysis table shows that the variation of the principal components after the 10th one tends to be flat, so it is a better choice to choose about the first 9 principal components. Nine variables including: gender, identity characteristics, ability to perform contracts, risk of breach of trust, consumption preferences, and behavioral characteristics, number of hit institutions, credit-related lawsuits, and business hours are the main explanatory variables that affect the comprehensive score.

Table II Principal component analysis table

	Eigenvalues	Eigenvalue variation	Eigenvalue ratio	cumulative eigenvalues	cumulative eigenvalue ratio
1	5.47	3.10	0.22	5.47	0.22
2	2.37	0.66	0.09	7.84	0.31
3	1.71	0.39	0.07	9.55	0.38
4	1.32	0.02	0.05	10.88	0.44
5	1.30	0.17	0.05	12.18	0.49
6	1.13	0.02	0.05	13.31	0.53
7	1.11	0.08	0.04	14.42	0.58
8	1.03	0.02	0.04	15.45	0.62
9	1.01	0.04	0.04	16.46	0.66
10	0.97	0.01	0.04	17.44	0.70
11	0.96	0.06	0.04	18.40	0.74
12	0.90	0.04	0.04	19.30	0.77

13	0.86	0.05	0.03	20.15	0.81
14	0.81	0.08	0.03	20.96	0.84
15	0.73	0.02	0.03	21.69	0.87
16	0.71	0.05	0.03	22.39	0.90
17	0.66	0.10	0.03	23.05	0.92
18	0.56	0.01	0.02	23.61	0.94
19	0.55	0.18	0.02	24.16	0.97
20	0.37	0.09	0.01	24.52	0.98
21	0.28	0.18	0.01	24.80	0.99
22	0.10	0.03	0.00	24.90	1.00
23	0.07	0.05	0.00	24.97	1.00
24	0.02	0.01	0.00	24.99	1.00

II. Lasso Algorithm Results

This study uses the Lasso algorithm to compare the Lambda values of 24 variables and 9 variables to verify the estimation results of the least squares method. Fig 1 is the trend chart of all 24 variable coefficient innovative Lambda; Figure 1 is the trend chart of coefficient innovative Lambda after streamlining the variables. Comparing Fig 2, we can find that after streamlining the variables, the Lambda value obviously compresses the regression coefficient to zero, and the coefficient path converges faster from left to right.

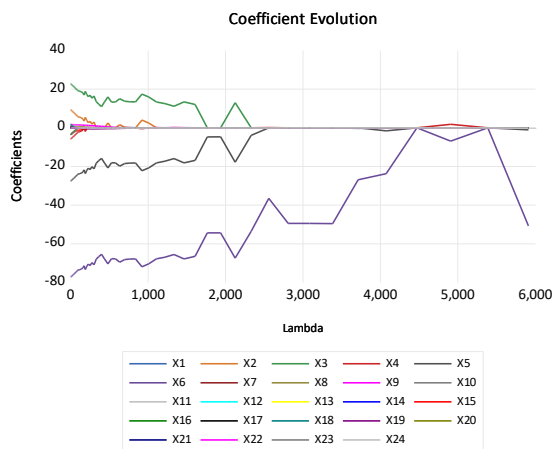


Fig 1 Trend chart of innovative Lambda for all variable coefficients

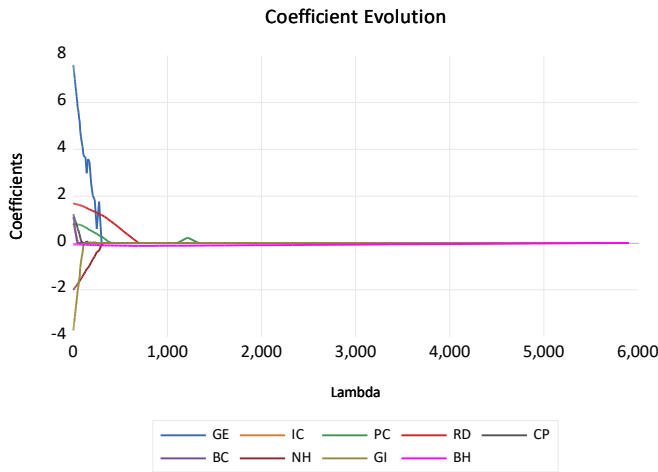


Fig 2 Trend chart of coefficient innovation Lambda after streamlining variables

III. Censored Regression Models Estimation Results

Table III is the estimation result of the truncated normal (Censored Normal) TOBIT model, in which the impact of gender (but not significant), identity characteristics, performance ability, breach of trust risk, consumption preference, and behavioral characteristics on the comprehensive score y is positively correlated; while hit The impact of the number of institutions, credit lawsuits involved, and business hours on the comprehensive score y is negatively correlated. This result is the same as the estimation result of the ordinary least squares method, and the results of each Limit Points are also significant.

Table III Estimation results of Censored Normal

variable	estimated coefficient	standard deviation	z statistic	p value
gender	7.627	4.920	1.550	0.121
Identity characteristics	0.837	0.320	2.616	0.009
Performance ability	0.798	0.113	7.062	0.000
Risk of breach of trust	1.689	0.105	16.006	0.000
consumption preference	1.235	0.262	4.711	0.000
Behavioral characteristics	1.133	0.396	2.862	0.004
Number of hits	-1.998	0.186	-10.759	0.000
credit involving lawsuit	-3.757	1.071	-3.507	0.001
business hours	-0.062	0.009	-6.632	0.000

Table IV is the result of the Censored Logistic model. Among them, the impact of gender, identity characteristics, contract performance ability, consumption preference, and behavioral characteristics on the comprehensive score y is positive; while the impact of the number of hit institutions, credit complaints, and business time on the comprehensive score y is negatively correlated. This result is the same as the estimation results of the above models.

Table IV Censored Logistic model estimation results

variable	estimated coefficient	standard deviation	z statistic	p value
gender	8.225	4.912	1.674	0.094
Identity characteristics	1.019	0.324	3.145	0.002
Performance ability	0.827	0.115	7.177	0.000
Risk of breach of trust	1.765	0.105	16.817	0.000
consumption preference	1.304	0.272	4.802	0.000
Behavioral characteristics	1.024	0.400	2.562	0.010
Number of hits	-2.015	0.184	-10.980	0.000
credit involving lawsuit	-3.556	1.130	-3.147	0.002
business hours	-0.066	0.009	-7.198	0.000

Table V is the estimation result of the Censored Extreme Value model. Among them, gender, identity characteristics, contract performance ability, breach of trust risk, consumption preference, and behavioral characteristics have a positive correlation with the comprehensive score y ; while the number of hit institutions, credit- related lawsuits, and business time have a positive correlation. The impact of the comprehensive score y is negatively correlated. This result is the same as the estimation results of the above models.

Table V Censored Extreme Value model estimation results

variable	estimated coefficient	standard deviation	z statistic	p value
gender	11.645	5.158	2.258	0.024
Identity characteristics	0.533	0.274	1.946	0.052
Performance ability	0.503	0.118	4.252	0.000
Risk of breach of trust	1.786	0.110	16.259	0.000
consumption preference	1.119	0.299	3.741	0.000
Behavioral characteristics	1.772	0.366	4.841	0.000
Number of hits	-2.124	0.187	-11.375	0.000
credit involving lawsuit	-4.190	0.889	-4.711	0.000
business hours	-0.047	0.010	-4.522	0.000

IV. Quantile Regression Model Estimation Results

This study uses quantile regression to examine the impact of variables on the comprehensive score y under different quantiles of the comprehensive score y . First of all, the estimation results of the quantile regression model in Table VI show that all the results of the impact of gender on the comprehensive score y are positive. At the 10% significance level, only the quantiles 0.5 and 0.95 are significant. As the number of digits increases, the effect of gender on scores becomes larger and larger. Regarding the impact of identity characteristics on the comprehensive score y , all the results are positively correlated, but they are only significant at quantiles 0.5 and 0.75. The impact of identity characteristics on the score is getting greater and greater, and the effect of performance ability on the score is getting stronger. The impact of the risk of dishonesty on the score is getting smaller and smaller, the impact of the risk of dishonesty on the score is getting larger and larger, the impact of consumer preferences on the score is getting larger and larger, the impact of behavioral characteristics on the score is generally increasing, and the number of hit institutions has a negative impact on the score. The effect is getting bigger and bigger, the

negative impact of credit-related complaints on scores is more consistent, and it is relatively unaffected by quantile changes. Finally, as for the impact of business time on the comprehensive score y , all results are negative and significant, and the negative impact of the number of hit institutions on the score is getting smaller and smaller.

Table VI Quantile regression model estimation results

Quantile variable	0.05	0.25	0.50	0.75	0.95
Constant term	45401 (0.000)	55675 (0.000)	51650 (0.000)	45698 (0.000)	37218 (0.000)
gender	4.542 (0.537)	7.068 (0.321)	13.979 (0.032)	8.145 (0.281)	23.04 (0.064)
Identity characteristics	0.692 (0.211)	0.536 (0.219)	1.038 (0.011)	1.367 (0.061)	1.249 (0.153)
Performance ability	0.601 (0.000)	1.094 (0.000)	0.807 (0.000)	0.565 (0.010)	0.395 (0.105)
Risk of breach of trust	0.871 (0.000)	1.668 (0.000)	1.897 (0.000)	1.868 (0.000)	1.711 (0.000)
consumption preference	0.548 (0.004)	1.192 (0.000)	1.203 (0.025)	1.312 (0.000)	1.228 (0.005)
Behavioral characteristics	0.520 (0.628)	0.887 (0.001)	1.378 (0.002)	1.039 (0.103)	1.828 (0.000)
Number of hits	-0.992 (0.000)	-1.786 (0.000)	-2.143 (0.000)	-2.026 (0.000)	-2.191 (0.000)
credit involving lawsuit	-1.331 (0.421)	-4.835 (0.400)	-3.442 (0.683)	-3.558 (0.000)	-4.781 (0.000)
business hours	-0.061 (0.000)	-0.075 (0.000)	-0.070 (0.000)	-0.061 (0.000)	-0.050 (0.006)

Note : The brackets (.) are p- values

5. Conclusion

This study uses a risk control decision-making model based on quantitative analysis of risk score trajectory differences to actually demonstrate the consumer finance model of real-life target companies. The research methods used in this study include principal component analysis, ordinal dependent variable model, truncated regression model, and quantile regression model analysis to test and predict the number of variables that affect consumers' comprehensive credit scores. Empirical results found that nine variables, including gender, identity characteristics, contract performance ability, breach of trust risk, consumption preference, behavioral characteristics, number of hit institutions, credit litigation, and business time, are the main explanatory variables that affect the comprehensive score. Regarding the impact of each variable on the comprehensive score, the four indicator variables such as identity characteristics, contract performance capabilities, consumption preferences, and behavioral characteristics have a positive impact on the comprehensive score; it means that the higher the identity characteristic score, the increased contract fulfillment ability, increased consumption preference score, and increased behavioral characteristic score can all increase the overall score. In addition, the three indicator variables such as the number of hit institutions, the number of credit lawsuits, and

business hours have a negative impact on the comprehensive score. It means that the increase in the number of hit institutions, the increase in the number of credit lawsuits, and the closer the business time will reduce the overall score. Among the nine variables, the negative factor with the greatest impact is credit litigation. Observed from the estimation results of the quantile regression model: as the quantile increases, the impact of the four indicator variables such as gender, identity characteristics, consumption preferences, and behavioral characteristics on the score becomes more and more significant, but the effect of performance ability on the score becomes larger and larger. Getting smaller and smaller. In addition, as the quantile increases, the number of hit institutions has an increasingly negative impact on the score; credit-related lawsuits have a relatively consistent negative impact on scores and are relatively unaffected by quantile changes. For the impact of business time on the composite score, the results for all quantiles are negative and significant.

Usually when interpreting module data, there are many variables and they are all qualitative explanatory variables. It is difficult to interpret them, and most of them are highly subjective qualitative indicators. Therefore, this study conducts quantitative analysis to find out the most important variables for the target company and provide the most effective interpretation of variables; on the other hand, it also provides analysis of time risks, effectively providing risk control and consumer credit ratings for the target company, so as to reduce the risk of consumer default and the cost of the target company one by one.

WORKS CITED

- Canner, Glenn B, Gabriel, Stuart A, Woolley, J, and Michael. Race, "Default risk and mortgage lending: A study of the FHA and conventional loan markets," *Southern Economic Journal*, vol. 58, no. 1, pp. 249-262, 1991.
- Cox, Donald, Jappelli, and Tullio, "The effect of borrowing constraints on consumer liabilities," *Journal of Money, Credit and Banking*, vol. 25, no. 2, pp. 197-213, 1993.
- Duca, John V, and Rosenthal, Stuart S, "Borrowing constraints and access to owner-occupied housing," *Regional Science and Urban Economics*, vol. 24, no. 3, pp. 301-322, 1994.
- Jappelli, Tullio, Pagano, and Marco, "Information sharing, lending and defaults: cross-country evidence," *Journal of Banking & Finance*, vol. 26, no. 10, pp. 2017-2045, 2002.
- Koenker, Roger, and Bassett Jr, Gilbert, "Regression quantiles. *Econometrica*," *Journal of the Econometric Society*, pp. 33-50, 1978.
- Koenker, Roger, and Machado, Jose AF, "Goodness of fit and related inference processes for quantile regression," *Journal of the American Statistical Association*, vol. 94, no. 448, pp. 1296-1310, 1999.
- Li, S, Li, Y, Ma, F, and Yu, X, "A Credit Risk Assessment Model Based on Deep Belief Networks and Feature Selection," *Journal of Ambient Intelligence and Human Computing*, vol. 12, no. 10, pp. 10207-10218, 2021.
- Lin, L. H, Lin, F. C, Lien, C. K, Yang, T. C, and Y. K Chuang, "Electronic payment behaviors of consumers under digital transformation in finance—A case study of third-party payments," *Journal of Risk and Financial Management*, vol. 16, no. 8, pp. 346-369, 2023.
- Zakaria, B, Ann-Kathrin, B, and Andreas, C, "The effect of traditional media consumption and internet use on environmental attitudes in Europe," *Journal of Evolutionary Economics*, vol. 33, no. 2, pp. 309-340, 2023.