

Research on Financial Venture Capital Evaluation based on Principal component regression Analysis

Tang Liyao¹, Li Xiang², Gu Chao³

¹School of Management and Engineering, Capital University of Economics and Business, Beijing, 100070

²Capital University of Economics and Business, Beijing, 100070

³School of Financial Management, Shanghai University of International Business and Economics, Shanghai, 201600

Keywords: Venture capital; motivation index; principal level regression; joint motivation; machine learning

Abstract: Based on the study of the real motivation of Chinese financial venture capital, this paper establishes the joint venture capital motivation, which includes financial motivation, resource motivation and transaction flow motivation, and takes it as an exogenous latent variable. The investment institutions with joint venture capital background on gem are selected to conduct a questionnaire survey, and the motivation of joint venture capital is empirically tested by partial least square method. At the same time, taking the problem of customer default in the current bank credit business as the starting point, the relationship between customer default rate and credit score card score is reasonably mapped. The prediction model is established by logical regression, and the data is analyzed, then the data set is divided, and the cross-time verification set is used as the final verification of the model. Finally, the KS value and AOC curve are used to evaluate the stability of the model. The experimental results show that the scorecard model constructed by the proposed method has good stability.

1. Introduction

Now is the era of artificial intelligence and credit economy, each of the traditional industries have to artificial intelligence transformation. Among them, taking the credit business of the bank as an example, strengthening the supervision of credit business is conducive to the banking industry to integrate into the economic era of artificial intelligence more quickly. Bank Internet credit refers to the economic behavior of using modern network technology to provide money lending to different borrowers. Under the mode of "bank + Internet", Banks score the credit rating of borrowers with the data provided by Internet e-commerce combined with their own credit evaluation standards. Credit investment is a business with both risks and challenges, so we need to consider how to avoid the risks of credit investment reasonably and effectively. Therefore, a stable and accurate bank credit evaluation system is an important guarantee for Banks to avoid lending risks [1]. The credit scoring system of the bank can help the bank to classify the credit rating of the borrower, which helps the bank to predict the repayment ability of the borrower as early as possible, so as to avoid risks and reduce losses.

2. Principal component regression algorithm and its implementation

2.1 Logistic regression algorithm

Logistic regression algorithm is a common algorithm used to solve classification problems. For the input data, when the data is greater than our threshold, output 1; when it is less than the threshold, output 0. The output variables of this model always range from 0 to 1. The hypothesis function of the LOGISTIC regression model is:

$$h_{\theta}(x) = g(\theta^T X) \quad (1)$$

G (z) is a commonly used logical function of S shape, $g(z) = \frac{1}{1+e^{-z}}$. By combining these two formulas, the hypothesis of logistic regression model [2] can be obtained. The function of h (x) is to calculate the probability that the output variable = 1 for a given input variable x based on the selected parameter. For example, for the input x, h (x) = 0.65 is calculated with the given parameter, indicating that the probability of Y being a positive class is 65%, and the corresponding probability of Y being a negative class is 35%.

The cost function of logistic regression is:

$$J(\theta) = \frac{1}{m} \sum_{i=1}^{mt} \text{cost}(h_{\theta}(x^{(i)}))$$

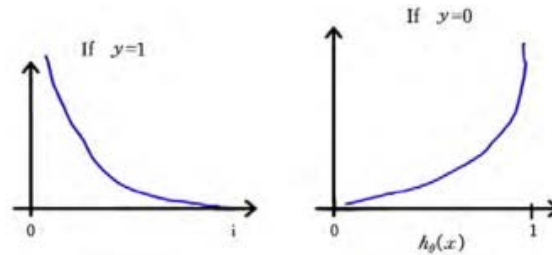


Fig.1 H (x) is related to Cost (h (x), y)

The Cost (h (x), y) constructed is simplified as follows:

$$\text{cost}(h_{\theta}(x), y) = -y * \log(h_{\theta}(x)) - (1 - y) * \log(1 - h_{\theta}(x)) \quad (2)$$

Put it into the cost function of logistic regression to get:

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m [y^{(i)} \log(h_{\theta}(x^{(i)})) + (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)}))] \quad (3)$$

2.2 Mapping of LOGISTIC regression scoring model

Common credit score CARDS are divided into A, B and C CARDS. The main difference of these three CARDS lies in their different time of occurrence, so the corresponding y in the model is also different. A card occurs before the loan and is represented by the maximum number of overdue days in the customer's history; A B card occurs in a loan and is represented by the longest delinquency in a multi-term loan; C card occurs after the loan. The meaning of Y is different for different purposes [3].

Through the preliminary analysis of the data, we draw the following conclusions: UID is the unique identity of each user; Bad_ind is used to judge whether a user is a "good person" or a "bad person", that is, Y in Equation (3); Obs_mth is the time at which each line of data occurs. For the remaining 10 attributes, we can use the characteristics in the model to predict the risk that the customer is a bad guy[4].

3 Financial venture capital motivation model

3.1 Reliability and validity analysis

The PLS structural equation model test is based on reliability and validity analysis. For internal consistency reliability analysis, Cronbach's Alpha) greater than 0.600 is acceptable, while greater than 0.700 indicates good reliability. Average variance extraction rate AVE is used as an indicator for validity analysis, which reflects the total variance of latent variable when interpretation is obtained from explicit variables due to measurement error. Generally speaking, the value of AVE is greater than 0.500, indicating that more than 50% information of observed variable is utilized. The AVE value of the lowest financial motivation is 0.712, which indicates that the latent variable in the model reflects the respective obvious variables well.

Tab.1 PLS path model reliability and validity test index

Joint venture capital data	Composite reliability	Vronbachs alpha	AVE	Hope data
Resource motivation.	0.954	0.939	0.806	0.700
Financial motivation.	0.908	-.865	0.712	
Transaction flow motivation.	0.924	0.837	0.860	
Joint risk trend	0.919	0.823	0.849	

3.2 Principal component analysis

Financial fair target is mainly manifested in the process of regulators for financial institutions, in order to ensure the financial consumers get equal treatment and trading information symmetry of financial service, fair and effective financial system mainly for the perfect laws and regulations system, timely and transparent disclosure of information, complete information, etc., in the real financial consumer goods in the process of trading, because of the complexity of the financial products and trading rules and professional, financial consumers in cognitive level, data collection and processing power of related financial problem and there are large difference between professional financial institutions, Adverse selection and the existence of moral hazard make financial consumers at a disadvantage in the whole transaction process [5]. In this paper, financial equity indicators are taken as exogenous variables, which can be divided into financial consumption information symmetry, financial consumers' independent choice and financial consumption recourse right. Specific indicators are public credit system coverage (percentage of adults), corporate information disclosure degree index, legal rights intensity index and corruption reduction index. Among them, the public credit registry system covering percentage (adult) and enterprise information disclosure index measures the credit information provided by the public and private credit system range, availability and real degree index, information disclosure is the financial transactions of the protection of investors and creditors mechanism, especially in many developing countries, the construction of credit reporting system to reduce the risk of moral hazard and adverse selection risk in financial services is very important.

KS (Kolmogorov-Smirnov) curve: KS is used to assess the model's ability to distinguish risk, i.e. the degree of difference between good and bad samples in the population sample. The greater the cumulative difference between good and bad samples and the higher the KS index, the stronger the risk discrimination ability of the model will be. KS curve drawing step: according to the prediction result of the learner (note, is the probability value of the positive example, non-0/1 variable) to sort the sample in order from the largest to the smallest -- this is the order in which the cutoff points are selected in turn. Select truncated points in order, and calculate TPR and FPR -- Only n truncated points can be selected, respectively at the positions of $1/n$, $2/n$ and $3/n$. The horizontal axis is the percentage of samples (maximum 100%), and the vertical axis is TPR and FPR respectively. KS curves can be obtained.

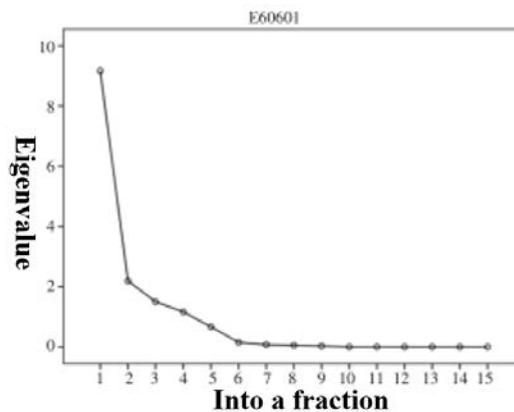


Fig.2 Principal component lithotripsy

The receiver-operating characteristic curve is also known as susceptibility curve, due to various

points on the curve reflects the same sensitivity and the name. They are all responses to the same signal stimulus, but they are the result of several different criteria. The receiver operating characteristic curve is a coordinate chart consisting of false positive probability (the probability of negative example error) on the horizontal axis, recall ratio on the vertical axis, and the curve drawn by different results obtained by different criteria under specific stimulus conditions

According to the above evaluation criteria, the training value of KS is 0.41 and the test value of KS is 0.37 after entering into the data set. In the case of a 4 percentage point difference between the predicted KS and the cross-time validation set, the ROC curve is shown in Figure 3.

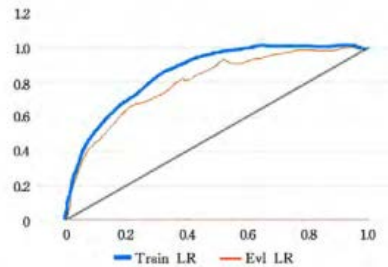


Fig.3 ROC evaluation curve

The result of path coefficient calculation is the most intuitive detection method to verify the PLS structural equation model. The path coefficient reflects the causal relationship and intensity between latent variables as well as between latent variables and observed explicit variables. In this study, Bootstrap resampling was performed on the initial samples to repeatedly test the important parameters in the PATH model of the PLS structure equation, and T statistics were constructed

3.3 Analysis and demonstration

First, the motivations of joint venture are not mutually exclusive or independent. Resource motivations have a positive relationship with financial motivations and transaction flow machines, and they may influence the decision-making of joint venture capital simultaneously. Therefore, venture capital institutions should consider the impact of various factors on the decision before making joint investment decisions.

Second, the empirical conclusion has a profound guiding significance for entrepreneurial enterprises. By understanding the factors that venture capital firms consider when making joint decisions, start-ups can effectively cater to the investment psychology of venture capital firms and thus successfully obtain venture capital

Third, China's venture capital market has been lack of effective exit mechanism, financing channels can not meet the financing needs of venture capital enterprises. In order to solve this situation, China Securities Regulatory Commission (CSRC) stepped up the launch and expansion of "New Third Board". On June 15, 2012, China Securities Regulatory Commission (CSRC) issued "Measures for supervision and Administration of Non-Listed Companies (Draft for Comments)", aiming to change the poor liquidity of venture capital market and strengthen the willingness of venture capital enterprises to participate. As a rule-maker and regulator of venture capital industry, China Securities Regulatory Commission (CSRC) should fully understand the factors that venture capital companies consider when making decisions and give reasonable guidance, so as to bring a new dawn for the long-awaited innovation and high technology.

According to the level of importance, the motivation of joint investment of Chinese venture capital institutions is resource motivation, financial motivation, transaction flow motivation, and resource motivation has a significant impact on financial motivation and transaction flow motivation. The study also holds that various motivations are not mutually exclusive or independent, and venture capital institutions should comprehensively consider the influence of various factors on joint investment decisions before making joint investment decisions, and start-up enterprises should consider the influencing factors of joint decision-making of venture capital institutions. effectively cater to the investment psychology of venture capital institutions, so as to successfully obtain venture capital. Only by understanding the factors considered by venture capital companies in

making decisions, can the CSRC provide reasonable guidance, so as to bring a new dawn to the long-awaited innovation and high technology.

4. Conclusions

In this paper, the empirical research shows that face international competition, is the required amount of capital investment of entrepreneurial innovation enterprises showed a trend of increase year by year, due to the amount needed to raise more capital to build enough moat prevent investment risks, the direct drive more and more risk investment institutions choose joint in order to reduce investment risks individually. Based on the perspective of small and medium-sized venture capital institutions, venture capital with smaller capital scale is more willing to seek joint venture with other venture capital institutions, that is, the greater the financial pressure, the stronger the incentive for joint investment. This is because through co-investment, venture capital institutions can participate in financing projects larger than the amount of capital they manage to share the benefits of co-investment. On the other hand, due to the small amount of their own capital and large operational risks, most small and medium-sized venture capital institutions often adopt the strategy of co-investment, that is, after the start-up has developed to a certain stage, they combine with other large-scale venture capital institutions to resist industrial risks.

For domestic venture capital industry regulator, the regulator also is necessary to deeply understand the risk investment industry joint intention and relevant demands, to develop effective encourage joint venture capital become bigger and stronger policies, guide the venture capital industry into virtuous cycle development, make the risk investment innovation that promotes public entrepreneurship has become a booster.

References

- [1] Lukas Beladi Sihombing, Manlian Ronald Adventus Simanjuntak. Analyzing the Hybrid Financing Deep-Discount Project Bonds and Land Leases for a New Toll Road Financing Model[J]. Journal of Infrastructure Systems, 2020(4).
- [2] Anastasios Sepetis. A Holistic Sustainable Finance Model for the Sustainable Capital Market[J]. Journal of Financial Risk Management, 2020(02).
- [3] Gu Wentao, Wang Ru, Zheng Suhao, Yang Yongwei. Research on the Prediction Model of Financial Market return Direction-- based on the text big data method [J/OL]. Statistical research: 1-12 [2020-11-29]. [Http://kns.cnki.net/kcms/detail/11.1302.C.20201125.0837.003.html](http://kns.cnki.net/kcms/detail/11.1302.C.20201125.0837.003.html).
- [4] Chen Yilin. Research on the interaction between Science and Technology Financial efficiency and Total Factor Productivity-- based on PVAR Model [J]. Modern Business Industry, 2020 (32): 10-11
- [5] Wang Hongqi, Xu Yulian. Research on the Coordination degree Model of Science and Technology Innovation and Science and Technology Finance and its Application [J]. Chinese soft Science, 2012 (06): 129138.