

Title Data Environment and the Effectiveness of Risk Control —The Differentiated Role of FinTech Risk Control Capabilities Across Varying Levels of Digitalization

Zichun Feng

*School of Finance, Southwestern University of Finance and Economics, Chengdu, China
42421002@smail.swufe.edu.cn*

Abstract. Despite the rapid development of digital finance, it is not clear whether the macro digital environment in this era increases the risk control effect of promoting household credit expansion. This paper, in a micro-sample from CHFS 2021, is a fit of provincial-level digitalization with the fintech risk control index, and uses the "Broadband China" policy as a variable to estimate IV-2SLS empirically to test the basic hypothesis that "the higher the level of digitalization is, the stronger the risk Control effect". The paper shows that though there is a positive covariance in the descriptive level, under strict causal assumptions, the interaction term and marginal effect of "risk control $\times$ digitalization" are not significant, and the data cannot robustly support this hypothesis. This strongly indicates that digital infrastructure itself does not increase the credit expansion effect of risk control as a "multiplier": on the contrary, risk control technology is risk defense, and its improvement in a highly digitalized environment results in better risk screening and constraint. This is good micro-level evidence for understanding the underlying risk governance logic in the digital economy.

Keywords: Digital finance, Fintech risk control, Credit card installment payments, Household finance, Linear probability model

1. Introduction

In recent years, the digital economy has fundamentally reshaped the way households receive credit, with consumer installment payments and online consumer loans increasingly popular [1, 2]. Banks are facing extremely high demands in low-level cities, and they increasingly rely on the risk control of an algorithm. Risk control in particular performs better in different regions: regions with high digitization have richer data and smoother repayment chains, and models can be more successful. Low digitalization regions are constrained by data sparsity and the digital divide. This raises a timely practical question: Does risk control work differently in different digitalization environments?

There are some published studies on these issues. First, fintech risk control uses digital footprints to help address information asymmetry. Studies generally point out that alternative data and machine learning models can effectively depict consumer default risk portraits, improve the accuracy of risk

control [3-5], thereby reducing default rate and optimizing credit allocation [6]. The second line argues that uncontrolled credit growth over low levels can cause unnecessary consumer debt and financial pressure [7], while excessive credit growth should be curbed [8, 9]. While the existing studies have found promising results, earlier studies have generally put the digital environment in isolation from risk control capabilities, discussing their indirect impacts on the credit market. Few studies have analyzed how macro digital infrastructure alters the effect of meso risk control capability on micro household credit.

This paper reviewed the moderating impact of the digital environment on fintech risk control from a micro-household perspective in order to test the theoretical assumption that "the greater digitalization, the more important the risk control effect." This paper benchmarked the cross-level matching between the micro sample of 2021 CHFS and the provincial digitalization level and fintech risk control index, and propose a linear probability model (LPM) to test this interaction between the three. For the sake of endogeneity reasons, this paper selects the population coverage rate of "Broadband China" demonstration cities as an exogenous instrumental variable for IV-2SLS causal inference [10-12].

2. Method

2.1. Research design and data sources

This paper aims to assess the difference between the impact of provincial fintech risk control capabilities on household installment payment behaviors in digital environments. This paper matches "household micro-data vs. provincial environmental indicators" for households to the use of credit card installment payments as an observable indicator for the installment payment or online consumer loan usage to the corresponding level of digitalization and the provincial fintech risk level of formalization, and determine the interaction effect between risk control and digital environment.

Household-level data is collected by the China Household Finance Survey (CHFS) project of Southwestern University of Finance and Economics. The baseline empirical focuses on the 2021 cross-sectional sample. The provincial level digitalization level is collected using the 2020 provincial "Digitalization Level" indicator of Peking University Digital Inclusive Finance Index (PKU-DFIIC). As the short-term inertia of provincial digital financial infrastructure and digital development is of increasing significance, this paper takes the 2020 indicator to approximate the digital environment of households in 2021 and match it with the CHFS sample by province. Provincial-level fintech risk control capability is collected with the "Fintech Risk Control Index" of the Institute of Finance, Chinese Academy of Social Sciences, for various fintech risk identification, risk control, and governance abilities in each province.

2.2. Sample processing and variable construction

The households used the "credit card installment usage" as explained (1 for occurrence, 0 for absence). The explanatory variables are the provincial "digitalization level" indicators published by PKU-DFIIC in 2020. The environmental variable is the provincial "fintech risk control index," which is standardized in the regression and classified among high and low risk control groups based on the inter-provincial median. The model has an interaction term between the two to test the effect of the digital environment on the risk control effect. In order to avoid endogeneity issues of the core variables, this paper selects the population coverage rate of "Broadband China" demonstration cities as a key indicator of the provincial digital environment for robustness testing. The models also

control for household income, assets, liabilities and micro-characteristics such as age, education, number of children and head of household status. In data construction, this paper matches the CHFS household survey sample with the provincial macro-indicators mentioned previously to build a cross-level fusion data of "household-province". This paper uniformly removed provincial administrative suffixes and filter misaligned samples in the model such that the matching accuracy is guaranteed.

2.3. Econometric models and estimation methods

Assuming that the dependent variable is binary, this paper used an LPM that is the standard specification. The model is as follows:

$$\text{Install}_{i,p} = \beta_0 + \beta_1 \text{Risk}_p + \beta_2 \text{Digital}_p + \beta_3 (\text{Risk}_p \times \text{Digital}_p) + X_i \quad (1)$$

Here, $\text{Install}_{i,p}$ indicates whether household (i) uses credit card installment payments. Risk_p and Digital_p represent provincial-level risk control capabilities and digitalization levels, respectively. X_i represents household-level control variables (such as household income, assets, age and education level of the head of household, household size, employment status, etc.). The core attention coefficient β_3 reflects the moderating effect of the digital environment on the influence of risk control capabilities on the probability of installment payments. Since the core explanatory variables take values at the provincial level, there may be error correlations in household observations within the same province. Therefore, this paper adopts a robust inference method of clustering by province for standard errors. To convert the interaction term coefficient into interpretable economic meanings, this paper reports the marginal effect of risk control capabilities at different digitalization levels in the results section, and constructs corresponding confidence intervals based on the covariance matrix of the estimators to demonstrate the direction, strength, and significance range of the influence of risk control capabilities on the probability of installment payments in different digitalization intervals.

3. Results

3.1. Descriptive statistics and correlation analysis

Table 1 gives a summary of the main variables. After missing values, the valid household sample size is 11,544. The average value of the explanatory variable "credit card installment usage" is 0.153. The mean values for provincial digitalization level and risk control index are 398.818 and 6.396, respectively, showing significant regional variations in digital infrastructure and risk management. Before econometric inference, this paper first review the correlation between core factors. Provincial aggregated data show positive covariation between the degree of digitalization and household installment payments. Further, by dividing the sample into high and low risk control groups according to the inter-provincial risk control index median found that, under different risk control levels, the fitting slope of degree of digitalization and installment occurrence rate shows a visible group difference. This feature first agrees with the core hypothesis of this paper: "The effect of risk control may have different effects in digital environments. The more digitalization is, the more significant its effect." But whether or not this macro-level correlation is robust at the micro level remains to be tested by econometrics.

Table 1. Descriptive statistical analysis

Variables	N	Mean	Std. Dev.	Min	Max
Credit Card Installment (1=Yes)	11544	0.15	0.36	0.00	1.00
Fintech Risk Control Index	11544	6.39	1.89	2.77	9.84
Digitalization Leve	11544	398.81	21.90	361.52	450.08
Age	11544	45.96	16.18	16.00	104.00
Education Level (1-9)	11544	4.47	1.96	1.00	9.00
Employed (1=Yes)	11544	0.61	0.48	0.00	1.00
Urban (1=City/Town)	11544	0.59	0.49	0.00	1.00
Number of Children	11544	1.46	0.92	0.00	9.00
ln(Assets+1)	11544	10.57	4.20	0.00	17.07
ln(Debt+1)	11544	1.34	3.43	0.00	15.42

Notes: Descriptive statistics are based on the common sample of the baseline regression (N = 11,544). The dependent variable is "Credit Card Installment (1=Yes)". The Fintech Risk Control Index and Digitalization Level are the original values matched at the provincial level.

3.2. Baseline regression and addressing endogeneity

Table 2. Baseline regression results

Variables	(1)LPM	(2)LPM	(3)LPM	(4)IV-SLS
Constant	0.1492*** (0.0080)	0.1492***(0.0080)	0.0898***(0.0144)	0.0854***(0.0216)
FinTech Risk Control Index	-0.0126***(0.0061)	-0.0194(0.0150)	-0.0112(0.0131)	0.0138(0.0440)
Degree of Digitalization		0.0078(0.0139)	0.0161(0.0178)	-0.0133(0.0528)
Risk Control × Digitalization			0.0022(0.0078)	0.0049(0.0115)
Age			-0.0001(0.0002)	-0.0001(0.0002)
Education Level			-0.0010(0.0023)	-0.0009(0.0025)
Employed			0.0111***(0.0048)	0.0100***(0.0048)
Urban			0.0209***(0.0103)	0.0220*(0.0112)
Number of Children			-0.0019(0.0071)	-0.0019(0.0070)
ln(Assets+1)			-0.0033****(0.0013)	-0.0032***(0.0014)
ln(Debt+1)			0.0716****(0.0030)	0.0715****(0.0030)
Observations (N)	10541	10541	10541	10541
R-squared	0.001	0.001	0.459	0.458
Household Controls	No	No	Yes	Yes
Province-Clustered Standard Errors	Yes	Yes	Yes	Yes

Notes: The dependent variable is the usage of credit card installments. Robust standard errors clustered at the provincial level are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Columns (1) to (3) report the stepwise regression results of the baseline linear probability model (LPM). Column (4) reports the two-stage least squares (IV-2SLS) estimation results utilizing the population coverage rate of "Broadband China" demonstration cities as an exogenous shock.

In order to test the main hypothesis and overcome possible endogeneity issues, this paper reports the estimation results of the linear probability model (LPM) and the two-step least squares method (IV-2SLS) (Table 2). The mean LPM estimates in columns (1) and (3) of Table 2 indicate that after the interaction term and household control factors are full, the coefficient of the core "risk control $\times$ digitalization" interaction term is not significant.

Due to the possibility of missing variables or non-overlapping issues such as reverse causality (e.g., credit demand driving digital infrastructure), further select "Broadband China" demonstration cities as IV to be the indicator for regional digitalization level, and perform 2SLS estimation in column (4). The IV-2SLS estimation results in column (4), where, after correcting for endogeneity bias with instrumental variables, it fails to reveal the coefficient of the "risk control $\times$ digitalization" interaction term (0.0049), which supports the theoretical "the more digitalised the more substantial the risk control effect." From an economic intuition point of view, this suggests that improving the digital environment does not mean enhancing risk control to encourage credit expansion as a "multiplier. A reasonable explanation is that, in a higher level of digital environment, the improvement of risk control capability more transforms into accurate identification and prudent credit constraint towards potential high-risk borrowers, thereby offsetting the pure scale expansion effect. Moreover, the control variables were the same: the number of the logarithm of household assets was very negative, and the number in the log of liabilities was very positive, which fits the speculations of precautionary savings and liquidity limits.

3.3. Marginal effects analysis

This paper analyzes the editing effect, and the results show that although the estimated value of the marginal effect point of the risk control ability shows an extremely weak upward trend with the improvement of the regional digital level, its 95% confidence interval always completely covers the zero line in the actual data distribution interval of the sample provinces. This provides a clear statistical inference: regardless of the level of digitalization in the region, the marginal impact of risk control capability on household phased use is not statistically significant. The result of marginal effect is highly consistent with the conclusion that "the interaction item is not significant" in the benchmark regression and instrumental variable test in Table 2 above, which confirms each other.

4. Conclusion

Based on the CHFS 2021 micro-sample with provincial macro-data and on the "Broadband China" policy, this paper investigates the moderating impact of the digital environment on fintech risk control. Comprehensive empirical analysis found that, although descriptive statistics show a preliminary positive covariation trend, under rigorous causal inference, the "risk control $\times$ digitalization" interaction term and marginal effect are both not significant, failing to support the traditional hypothesis of "the higher the degree of digitalization, the more significant the risk control effect". This profoundly reveals that digital infrastructure did not act as a "multiplier" to simply amplify the credit expansion effect of risk control. On the contrary, the essence of risk control technology is still risk defense. In highly digitalised regions, better risk control results are translated into accurate screening and proper conditions on high-risk borrowers, and thus offset the temptation to only expand credit. In addition, micro-level properties such as household employment and asset-liability issues take a more significant and dominant role in installment decisions.

Based on this conclusion, this paper proposed that governments and financial institutions that take digital transformation into account should not simply envision digital infrastructure as an agent

of uncontrolled credit growth, but must more fully observe the impact of a change in risk management by technology, and the role of underlying risk control constraints. This results may also be limited: macro- and micro-level cross-level data matching may underestimate the heterogeneity of the information environment between city and county levels, and failed to capture implicit risk preferences of households. Future studies may further consider larger city-level information, large-scale variables like financial institution density and personal credit scores, and other specific digital scenarios such as internet consumer loans, to better understand the relationships between macro-level digital environment and micro risk control.

References

- [1] Cornelli, G., Frost, J., Gambacorta, L., Rau, R., Wardrop, R., & Ziegler, T. (2021). Fintech and big tech credit: What explains the rise of digital lending? *CESifo Forum*, 22(2), 30–34.
- [2] Suri, T., Bharadwaj, P., & Jack, W. (2021). Fintech and household resilience to shocks: Evidence from digital loans in Kenya. *Journal of Development Economics*, 153, 102697.
- [3] Nobanee, H., Ellili, N. O. D., Chakraborty, D., & Shanti, H. Z. (2024). Mapping the fintech revolution: How technology is transforming credit risk management. *Global Knowledge, Memory and Communication*, 75(3-4), 1238–1262.
- [4] Zhu, M., Shia, B.-C., Su, M., & Liu, J. (2024). Consumer default risk portrait: An intelligent management framework of online consumer credit default risk. *Mathematics*, 12(10), 1582.
- [5] Berg, T., Burg, V., Gombović, A., & Puri, M. (2020). On the rise of fintechs: Credit scoring using digital footprints. *The Review of Financial Studies*, 33(7), 2845–2897.
- [6] Goldstein, B., Jagtiani, J., & Lemieux, C. (2025). Fintech innovations in banking: Fintech partnership and default rate on bank loans. *Working Papers 25-21, Federal Reserve Bank of Philadelphia*.
- [7] Di Maggio, M., Kermani, A., & Korgaonkar, S. (2022). Fintech borrowing and consumer stress. *The Review of Financial Studies*, 35(10), 4577–4620.
- [8] Liu, Y., & Li, Z. (2024). The use and disuse of fintech credit: When buy-now-pay-later meets credit information sharing. BIS Working Papers No. 1163, Bank for International Settlements.
- [9] Agarwal, S., Alok, S., Ghosh, P., & Ramadorai, T. (2022). Fintech lending and credit bureau monopolies. *Journal of Financial and Quantitative Analysis*, 57(1), 35–68.
- [10] Kong, D., Li, J., & Jin, Z. (2023). Can digital economy drive income level growth in the context of sustainable development? Fresh evidence from 'Broadband China'. *Sustainability*, 15(17), 13170.
- [11] He, J., Mu, Y., Wang, C., & Mao, Y. (2024). Impact of digital infrastructure construction on financial development: Evidence from the 'Broadband China' strategy. *Heliyon*, 10(15), e35262.
- [12] Wang, X., & He, J. (2023). Digital economy development and the urban-rural income gap: Evidence from Chinese cities. *PLoS ONE*, 18(2), e0279269.