

Impact of Digital Inclusive Financial Use on Entrepreneurship

-- An empirical study based on panel data from 31 provinces in China

Manlin Jiang*

School of International Trade and Economics, University of International Business and Economics,
Beijing, 100020, China

* Corresponding Author Email: jml020827@163.com

Abstract. Using the digital inclusive finance index at the provincial level from the Peking University Digital Inclusive Finance Usage Index report, this paper selects panel data from 31 provinces, municipalities directly under the Central Government, and autonomous regions of China (excluding Hong Kong, Macao, and Taiwan) from 2011-2016, and constructs multivariate linear models, multivariate nonlinear models, and fixed effects models to investigate the mechanism of the impact of inclusive finance on entrepreneurship. Regression analysis was conducted. The results show that there is a significant positive relationship between digital financial inclusion and entrepreneurship level. The findings of this paper provide policy implications for how to incentivize entrepreneurship more effectively.

Keywords: Digital Inclusive Finance; Entrepreneurship Levels; Panel Data.

1. Introduction

With the wide penetration of the concept of mass entrepreneurship and innovation, new market players are flourishing and becoming a new force to promote the transformation of old and new dynamics [1~3]. To promote sustainable and balanced development of the financial industry, boost the optimization and upgrading of the economic development mode, and enhance social equity and social harmony, the country is vigorously developing inclusive finance and highlighting the construction of an inclusive financial supply system, and the issue of the relationship between finance and entrepreneurship is drawing more and more attention from the academic community [4~6]. This paper will focus on the relationship between the digital inclusive finance index and regional entrepreneurship level based on the report of the Peking University Digital Inclusive Finance Index and conduct an empirical analysis.

Considering both macro and micro variables, this study improves the overall effect of the Digital Inclusive Finance Index on entrepreneurship in theory and proposes countermeasures on how to improve entrepreneurship level in practice, which helps to further develop the incentive effect of financial infrastructure on active emerging market players and entrepreneurship [7~10].

2. Data collection and organization

This paper selects panel data for 31 provinces, municipalities directly under the Central Government, and autonomous regions in China from 2011-2016 (excluding Hong Kong, Macao and Taiwan), with data obtained from the Digital Inclusive Finance Index published by the Digital Finance Research Center of Peking University, macro variables provided by the National Bureau of Statistics, Internet reports published by the China Internet Information Center (CNNIC), the China Financial Statistics Yearbook, the official website of the People's Bank of China, and the information of newly registered enterprises obtained from the website of AIC. The report first uses 186 data for cross-sectional regressions, followed by panel regressions using panel data from 2011-2016.

3. Multiple linear regression

The model construction followed the order of explanatory variables from less to more, and finally, nine linear models were constructed as shown in table 1.

Table 1. Multiple linear regression model (regression-by-regression analysis)

Regressor entre	(1)	(2)	(3)	(4)
<i>lnindex</i>	0.43870*** (4.77)	0.33987*** (3.81)	0.57668*** (6.17)	0.57731*** (6.13)
<i>lined</i>		2.06484*** (4.83)	3.16742*** (7.12)	3.18074*** (6.57)
<i>In web</i>			-1.47079*** (-5.49)	-1.47186*** (-5.47)
<i>lnopen</i>				-0.00458 (-0.07)
<i>public</i>				
<i>Antec</i>				
<i>lnGDP_per</i>				
<i>NTFS</i>				
<i>On loan</i>				
Constant	4.49695*** (22.90)	2.74865*** (6.76)	0.69673 (1.31)	0.67888 (1.15)
Observations	186	186	186	186
R-squared	0.110	0.211	0.323	0.323
F test	3.77e-06	3.96e-10	0	0
R̄2	0.105	0.202	0.312	0.308
F	22.74	24.42	28.94	21.59

Regressor entre	(5)	(6)	(7)	(8)	(9)
<i>lnindex</i>	0.63460*** (8.05)	0.30594*** (3.60)	0.30938*** (3.67)	0.28363*** (3.39)	0.34367*** (4.16)
<i>lnedu</i>	-0.25732 (-0.46)	-1.66671*** (-3.09)	-2.06134*** (-3.61)	-2.30798*** (-3.97)	-2.50043*** (-4.43)
<i>lnweb</i>	-0.41170 (-1.62)	-0.45503** (-2.01)	-0.57343** (-2.46)	-0.64203*** (-2.77)	-0.63451*** (-2.83)
<i>lnopen</i>	0.11326** (2.04)	-0.19273*** (-2.89)	-0.21523*** (-3.21)	-0.23321*** (-3.50)	-0.24255*** (-3.76)
<i>public</i>	-1.45534*** (-8.92)	-1.24149*** (-8.34)	-1.40770*** (-8.29)	-1.47203*** (-8.50)	-1.60571*** (-9.36)
<i>Antec</i>		0.55991*** (6.88)	0.46320*** (4.91)	0.46871*** (5.04)	0.45097*** (5.00)
<i>lnGDP_per</i>			0.44853** (1.98)	0.55138** (2.42)	0.81113*** (3.50)
<i>NTFS</i>				0.06225** (2.29)	0.08455*** (3.13)
<i>On loan</i>					-0.25057*** (-3.63)
Constant	8.62688*** (8.47)	11.71172*** (11.56)	10.09865*** (7.81)	10.10022*** (7.90)	9.97250*** (8.06)
Observations	186	186	186	185	185
R-squared	0.530	0.629	0.637	0.647	0.672
F test	0	0	0	0	0
R̄2	0.517	0.616	0.622	0.631	0.655
F	40.66	50.47	44.53	40.40	39.85

t-statistics in parentheses*** p<0.01, ** p<0.05, * p<0.1

Considering the regression of the number of new firms (*entre*) on the logarithm of the digital financial inclusion index (*index*), without any control variables, the equation of this regression takes the form of:

$$\widehat{entre} = 4.97(22.9) + 0.44(4.77) \times \lnindex \quad (1)$$

From the above regression results, we conclude the following.

(1) In all multiple regression results, the digital inclusion index (*index*) is significant at the 1% significance level, so it can be roughly concluded that the digital inclusion index shows a positive relationship with the regional entrepreneurship level.

(2) The regression in column (2) adds the control variable of years of education per capita (*edu*) to control for the level of education, and the difference in regional education levels affects the number of new businesses and thus the level of regional entrepreneurship. The inclusion of this variable increases \bar{R}^2 from 0.105 to 0.202, and the effect of the digital inclusion index on entrepreneurship levels decreases by about 25%.

(3) After adding public expenditure per capita (*public*), \bar{R}^2 increases from 0.308 to 0.517, and public expenditure per capita has a positive effect on innovation, which is significant at the 1% level of significance. At this point, the effect of the level of education (*lnledand* the logarithm of Internet penetration (*lnweweb* entrepreneurship turns from positive to negative.

(4) After adding the level of STI (*Intec*), \bar{R}^2 increases from 0.517 to 0.616, indicating that the level of STI has strong predictive power on entrepreneurship.

(5) After adding the per capita micro-institutional loan balance (*on loan*), all coefficients in the regression in column (9) are significant at the 1% level of significance.

(6) Combined with the above findings, the models (1) (2) (3) (4) have a poor fit and the optimum is excluded. After adding in *public* it is found that *lnedu* and *web* are not significant for both, so (5) is excluded. The regression in column (9) has the best predictive power and the coefficients pass the significance test, so this regression is chosen as the base model to start the discussion later in the paper.

Before conducting the empirical analysis, to avoid the problem of incomplete multicollinearity in the model and the bias of the standard error, which leads to erroneous regression results, we used Stata's VIF check method, which is to check the magnitude of the variance inflation factor, and the test results show that the highest value of VIF for a single variable among the target and control variables is 6.44, and the mean value of VIF for all variables is 3.32.

According to the rule of thumb followed by the academic community, when the VIF value of a single variable is higher than 10, the model is generally considered to have a more serious multicollinearity problem, and it can be assumed that the model constructed in this paper does not have a multicollinearity problem.

Using the test command of Stata software, the F-statistic of the "overall" regression can be calculated to test the joint hypothesis that all the slope coefficients are zero, and the original and alternative hypotheses are $H_0: \beta_1 = 0, \dots, \beta_9 = 0$; $H_1: \beta_j$ not all 0, ($j=1, 2, \dots, 9$). The result of joint significance is $\text{Prob}>F=0.0000$, so the original hypothesis is rejected.

4. Multiple nonlinear regression

Introduce polynomials of the target variable *lnindex* in order - *lnindex* 2, *lnindex* 3, *lnindex* 4 , *lnindex* 5, *lnindex* 6. The overall regression function of the polynomial regression model presupposes that:

$$\begin{aligned} entre_i = & \beta_0 + \beta_1 \lnindex + \beta_2 \lnindex^2 + \beta_3 \lnindex^3 + \beta_4 \lnindex^4 \\ & + \beta_5 \lnindex^5 + \beta_6 \lnindex^6 + \beta_7 \lnedu + \beta_8 \lnweb + \beta_9 \lnopen + \beta_{10} \lnpublic + \beta_{11} \ln tec + \beta_{12} \ln GDP_{per} + \\ & \beta_{13} \ln TFS + \beta_{14} \ln loan \end{aligned} \quad (2)$$

The regression results are shown in Table 2.

Table 2. Multiple non-linear regression model

Regressor entre	(1)	(2)	(3)	(4)	(5)
lnindex	-0.01618*** (-2.73)	0.21863*** (11.14)	0.22011*** (10.05)	0.23171*** (10.55)	0.24732*** (10.43)
lnindex2	0.09435*** (193.82)	0.11207*** (75.35)	0.11237*** (45.04)	0.13368*** (15.88)	0.20348*** (4.86)
lnindex3		-0.00833*** (-12.27)	-0.00860*** (-4.50)	-0.04226*** (-3.29)	-0.19545** (-2.15)
lnindex4			0.00006 (0.15)	0.01776*** (2.65)	0.14301* (1.93)
lnindex5				-0.00309*** (-2.65)	-0.04807* (-1.81)
lnindex6					0.00598* (1.70)
lnedu	0.19839*** (4.85)	0.14516*** (4.79)	0.14522*** (4.78)	0.13772*** (4.59)	0.13055*** (4.33)
lnweb	0.05355*** (3.42)	0.05667*** (4.94)	0.05645*** (4.87)	0.05727*** (5.02)	0.05844*** (5.14)
lnopen	-0.01515*** (-3.33)	-0.01400*** (-4.20)	-0.01391*** (-4.10)	-0.01420*** (-4.25)	-0.01452*** (-4.37)
public	0.00747 (0.52)	0.00209 (0.20)	0.00197 (0.19)	0.00063 (0.06)	-0.00343 (-0.32)
Intec	0.01845*** (2.83)	0.01575*** (3.29)	0.01564*** (3.22)	0.01687*** (3.52)	0.01551*** (3.21)
lnGDP_per	-0.08328*** (-5.06)	-0.06206*** (-5.09)	-0.06209*** (-5.08)	-0.05712*** (-4.70)	-0.05477*** (-4.50)
lnTFS	0.00195 (1.03)	0.00196 (1.42)	0.00196 (1.41)	0.00196 (1.43)	0.00210 (1.54)
On loan	0.01747*** (3.56)	0.01039*** (2.85)	0.01041*** (2.85)	0.01012*** (2.82)	0.01015*** (2.84)
Constant	2.86875*** (31.22)	2.35204*** (29.60)	2.34910*** (28.66)	2.31695*** (28.44)	2.28681*** (27.57)
Observations	185	185	185	185	185
R-squared	0.998	0.999	0.999	0.999	0.999
F test	0	0	0	0	0
\bar{R}^2	0.998	0.999	0.999	0.999	0.999
F	11491	19440	17720	16928	15893

The above table shows the results of the polynomial regressions, with quadratic, cubic, quadratic, quintuple, and sextuple nonlinear polynomial regression models of the logarithm of the digital financial inclusion index (index) introduced in columns (1) to (5), respectively. The values in parentheses in the above table also use the t-statistic of the coefficient estimates instead of the standard errors, which can also be calculated directly from the estimates and t-values.

Analyzing and comparing the above five models, the following conclusions are drawn: in the introduction of index polynomial regression, \bar{R}^2 is significantly improved, while the estimates of the coefficients of the variables are significant, especially in the model with five quadratic terms in column (4), where the effects of most of the variables are significant at the 1% significance level, and \bar{R}^2 increases to 0.999 compared to 0.655 in the linear regression. The significance of some of the coefficient estimates is found to be affected when the six-order term is introduced, so the regression in column (4) can be used.

$$\widehat{entre}_i = 2.317 + 0.232lnindex + 0.134lnindex^2 - 0.042lnindex^3 + 0.018lnindex^4 - 0.003lnindex^5 + 0.138lnedu + 0.057lnweb - 0.014lnopen + 0.001lnpublic + 0.017ln tec - 0.057lnGDP_{per} + 0.002lnTFS + 0.010lnloan \tag{3}$$

The result of joint significance Prob>F=0.0000 rejects the original hypothesis that the overall regression function is a higher-order polynomial. Using mixed cross-sectional regression (clustering robust standard error), at this time, the time effect and fixed effect of the data were not considered, and it was found that the logarithm of the target variable digital financial inclusion index (index) was negatively related to the number of new enterprises (entre) of the explanatory variable, and this result was contrary to the real economic logic, so the results of this regression were not credible.

5. Panel regression

In this section, the variable pro_code represents the province number, and different pro_codes represent different provinces. By adding 30 region-specific i.pro_code dummy variables and 5 year-specific i.year dummy variables, we tested the individual fixed effects model, the time fixed effects model, and the individual time double fixed effects model, respectively. Since the regression models may have autocorrelation, cross-sectional correlation, n and heteroskedasticity, to further verify the reliability of the above results, we adjusted each regression equation for the corresponding robustness standard deviation, and the regression results are shown in Table 3.

Table 3. Fixed effects regression results (with the inclusion of dummy variables)

VARIABLES	(1) Individual fixation	(2) Fixed time	(3) Two-way fixed
lnindex	-0.04877 (-1.17)	0.25211 (0.88)	-0.06049 (-0.93)
lnedu	-0.64006* (-1.66)	-2.40530*** (-4.13)	-0.56354* (-1.83)
lnweb	1.07631*** (4.65)	-0.66690*** (-2.93)	0.58988*** (3.09)
lnopen	-0.21460*** (-4.97)	-0.20628*** (-2.80)	-0.04860 (-1.18)
public	-0.29716** (-2.49)	-1.61702*** (-9.18)	0.20892* (1.85)
ln tec	0.20570*** (3.11)	0.42946*** (4.52)	-0.05244 (-0.83)
lnGDP_per	0.55014*** (2.64)	0.79925*** (3.35)	-0.05049 (-0.28)
lnTFS	-0.00746 (-0.56)	0.08293*** (3.03)	-0.01317 (-1.23)
On loan	0.04215 (0.93)	-0.23983*** (-3.39)	0.10472*** (2.88)
Constant	4.84530*** (4.01)	10.03189*** (7.77)	5.19080*** (5.13)
Observations	185	185	185
R-squared	0.987	0.677	0.992
F test	0	0	0
r2_a	0.984	0.650	0.990
F	284.5	25.46	412.1

We regressed the model again by using Stata's xtreg, fe command. Among the results in column (2) are the individual fixed effects, the time fixed effects, and the two-way fixed effects using five dummy variables (i.year) on top of the individual fixed effects using the centralized approach, and the regression results are shown in Table 4.

Table 4. Fixed effects regression results (individual centrality)

VARIABLES	(1)	(2)
lnindex	-0.04877 (-1.17)	-0.06049 (-0.93)
lnedu	-0.64006* (-1.66)	-0.56354* (-1.83)
lnweb	1.07631*** (4.65)	0.58988*** (3.09)
lnopen	-0.21460*** (-4.97)	-0.04860 (-1.18)
lnpublic	-0.29716** (-2.49)	0.20892* (1.85)
Intec	0.20570*** (3.11)	-0.05244 (-0.83)
lnGDP_per	0.55014*** (2.64)	-0.05049 (-0.28)
lnTFS	-0.00746 (-0.56)	-0.01317 (-1.23)
lnloan	0.04215 (0.93)	0.10472*** (2.88)
Constant	5.26911*** (4.51)	5.28524*** (5.40)
Observations	185	185
R-squared	0.853	0.913
Number of prov_code	31	31
F test	0	0
r2_a	0.814	0.885
F	93.63	104.7

$P < 0.01$ for the F-test, with a P-value less than 0.01, it is considered that there is a significant individual fixed effect, indicating that there is a significant individual fixed effect compared to the mixed OLS, and the individual fixed effect model is superior, and the model is finally chosen.

6. Conclusions

(1) Most of the regression results show that the digital inclusion index has a positive effect on regional entrepreneurship levels, consistent with the expected economic logic, and the goodness-of-fit R^2 of the multiple linear regressions is relatively good, with all variables significant at the 1% level of significance. In contrast, for the regression with the inclusion of individual fixed effects, the core explanatory variables are not significant at the 10% level of significance. Therefore, the model with the inclusion of individual fixed effects was not selected. Considering the complexity of the variables in the nonlinear regression model, we used the multiple linear regression model as our optimal model.

(2) There is a significant positive relationship between the digital inclusion index and the level of entrepreneurship, with each 1% increase in the digital inclusion index (lnindex) causing a 0.34367% increase in the number of new businesses, holding all other variables constant.

(3) The coefficient of average years of education indicates that, holding other variables constant, each 1% increase in average years of education (lnedu) causes a 2.50043% decrease in the number of

new enterprises; the coefficient of Internet penetration indicates that, holding other variables constant, each 1% increase in Internet penetration (lnweb) causes a 0.63451% decrease in the number of new enterprises; the coefficient of level of external openness The coefficient of the level of openness to the outside world (lnopen) indicates that a 1% increase in the level of openness to the outside world (lnopen) causes a 0.24255% decrease in the number of new enterprises when other variables are held constant; the coefficient of public expenditure per capita indicates that a 1% increase in public expenditure per capita (lnpublic) causes a 1.6057% decrease in the number of new enterprises when other variables are held constant; the coefficient of the level of science and technology innovation indicates that a 1% increase in the level of science and technology innovation causes a 1.6557% decrease in the number of new enterprises when other variables are held constant. The coefficient of the level of science and technology innovation (Intec) indicates that a 1% increase in the level of science and technology innovation (Intec) causes a 0.45097% increase in the number of new enterprises, while holding other variables constant. The coefficient of the level of development of traditional financial sector (lnTFS) indicates that a 1% increase in the level of development of traditional financial sector (lnTFS) causes a 0.08455% increase in the number of new enterprises, while the coefficient of the balance of loans per capita of microfinance institutions (lnloan) indicates that a 1% increase in the balance of loans per capita of microfinance institutions (lnloan) causes a 0.25057% decrease in the number of new enterprises while holding other variables constant.

(4) R2 indicates that the explanatory variables explain 67.2% of the variation in the explained variable, which has strong explanatory power. The regression standard error SER measures the degree of dispersion of the observations around the regression line.

(5) The intercept term has no real economic meaning.

References

- [1] Liang Beng, Zhang Jianhua. Can digital financial inclusion development stimulate innovation? -- Evidence from Chinese cities and SMEs [J]. *Contemporary Economic Science*, 2019, 41(05): 74-86.
- [2] Jiang Changliu, Jiang Chengtao. Can digital inclusive finance promote high-quality regional economic development? --Based on the experience of 258 cities A validation of the data [J]. *Journal of Hunan University of Science and Technology (Social Science Edition)*, 2020, 23(03): 75-84. DOI:10.13582/j.cnki.1672-7835.2020.03.009.
- [3] Jiao Jinpu, Huang Tingting, Wang Tiandu, Zhang Shaohua, Wang Jian. The development process and empirical study of inclusive finance in China [J]. *Shanghai Finance*, 2015(04): 12-22. DOI:10.13910/j.cnki.shjr.2015.04.003.
- [4] Guo F, Wang JY, Wang F, et al. Measuring the development of digital inclusive finance in China: indexing and spatial characteristics [J]. *Economics (Quarterly)*, 2020 (4): 1401- 1418.
- [5] Yin Zhichao, Song Quanyun, Wu Yu, et al. Financial knowledge, entrepreneurial decision and entrepreneurial motivation [J]. *Management World*, 2015 (1): 87-98
- [6] Spatial Determinants of Entrepreneurship in India [J]. Ejaz Ghani, William R. Kerr, Stephen O'Connell. *regional Studies*. 2014 (6)
- [7] Zhang Na. Research on the influence of financial development on the effect of industrial restructuring of FDI [D]. Ocean University of China.
- [8] [8] Meng Fang. Research on factors influencing informal entrepreneurial behavior in poor areas based on embedding theory [D]. Jishou University.
- [9] Zhang Weiping, Cao Tingqiu. Analysis of systemic risk spillover effects among real estate firms in China - based on tail risk network model [J]. *Financial Research*, 2022(7): 21.
- [10] Ma Yong, Yao Chi. Inflation targeting, policy credibility and macro-control effects [J]. *Financial Research*, 2022(7): 19.