Research article

Measuring provincial digital finance development efficiency based on stochastic frontier model

Guang Liu, Hong Yi* and Haonan Liang

School of Economics and Statistics, Guangzhou University, Guangzhou, China

* Correspondence: Email: 2009700003@e.gzhu.edu.cn.

Abstract: Effective development of digital finance is vital to closing the regional economic disparities. This study aims at investigating the efficiency of digital finance development in China and its implications for closing regional economic disparities. Using the stochastic frontier model, we estimate the development efficiency of digital finance in 31 provinces in China from 2011 to 2020, and reveal their characteristics of temporal evolution and spatial distribution. The results show that the efficiency of digital finance development in each province shows a tendency to increase quickly first and then slowly decline. The provinces with a higher level of digital finance development always have higher development efficiency at the beginning of the sample period, which then declines rapidly after reaching the maximum, and even less than the national average value at the end of the period, with significant regional disparities observed. The provinces with a higher level of digital finance development always have higher development efficiency at the beginning of the sample period, which then declines rapidly after reaching the maximum, and even less than the national average value at the end of the period. The imbalance of development efficiency among different provinces is increasing, and the potential for development efficiency in the central and western regions is relatively greater. These findings have important implications for promoting high-quality economic development and common prosperity in China. In the future, we should continually prevent the development efficiency of digital finance to decline rapidly in all provinces (especially in the eastern region), and strive constantly to bridge the gap of development efficiency among different province, so as to provide a better surrounding for promoting high-quality economic development and common prosperity.

Keywords: finance development; digital finance efficiency; stochastic frontier model; regional economic disparities; spatial-temporal characteristics

1. Introduction

The original concept of digital finance referred to the provision of financial services through various digital technologies, such as credit cards, chip cards, and automated teller machines (ATMs) (Banks, 2001). As early as 2004, some financial institutions and industries in Europe had already recognized the potential of the Internet to empower finance and expand the industry, and began investing in Internet-related financial businesses (Barbesino, 2005). As the Internet and mobile applications gradually became ubiquitous in people's lives, the definition of digital finance expanded accordingly (Manyika et al., 2016). Digital finance surpasses the geographical limitations of traditional finance, and offers more accurate services at lower costs, with prominent features such as remote access, real-time transactions, and inclusiveness. In China, digital finance emerged in 2013 and has grown rapidly since 2016, with its market size ranking among the highest in the world. Industries such as third-party payments, online lending, digital insurance, and digital currency have all reached the forefront of the world.

In addition to benefiting individual users, digital finance can also benefit the macroeconomic development and the daily operations of the financial sector. Firstly, digital finance can improve the payment and transfer efficiency of individual users, thereby promoting consumer willingness and increasing total consumption, ultimately promoting economic development. Additionally, due to the traceability of digital transactions, government spending and tax losses can be reduced, and the circulation of counterfeit and low-quality banknotes can decrease (Ozili, 2018). Secondly, from an international perspective, digital finance has also helped many developing countries overcome difficult times, created a large number of jobs (Rizzo, 2014), and changed the payment methods of some underdeveloped countries in commerce and trade (CGAP). Finally, for financial institutions, digital finance can promote the expansion of their balance sheets and reduce service costs (Scott et al., 2017). In summary, digital finance can drive the existing financial system to reshape to a certain extent, improve resource allocation efficiency, and enhance capabilities of risk control. This is of great practical significance for breaking through the constraints of traditional finance, enhancing financial support for the development of the real economy, and expanding the inclusiveness and popularity of financial services.

Although the development of digital finance brings many benefits, the serious challenges it faces cannot be ignored. For example, the impact of digital finance of income inequality has led to the emergence of the Kuznets effect, where income inequality shows a "bell-shaped" curve with the development of digital finance (Yao and Ma, 2022). Many regions in China have not yet crossed the turning point of the bell-shaped curve of income inequality, which is a problem that requires attention in the future.

Digital finance has quietly infiltrated every aspect of economy and social life, playing a significant role. However, our understanding of the evolutionary characteristics and regional differences in digital finance development efficiency is still limited, and it is necessary to study it in depth. Therefore, we use the Peking University Digital Inclusive Finance Index and the stochastic frontier model (SFM) to measure the efficiency of digital finance development and grasp the dynamic and efficiency of digital finance development from a macro perspective. This allows us to better identify the operating mechanism of the digital finance market and better utilize its key role. This is helpful in theory to enhance the scientific understanding of the current level and the effects of digital

finance development in China, enrich the theoretical achievements of digital finance development, and provide a theoretical basis for addressing the problem of imbalanced development of digital finance. In practice, it helps to objectively examine the development status of the industry and promote the coordinated and balanced development of digital finance in various regions.

2. Literature review

Existing studies have mainly focused on the level of development of digital finance and its relationship with economic and social development, regional imbalances, and technological innovation in small and medium-sized enterprises. For example, Zhang and Xing (2021) used county-level data from 2014 to 2018 to explore the dynamic distribution and regional differences of digital inclusive finance development in rural areas. They found that the overall level of digital inclusive finance in rural areas and in the eastern, central, and western regions tended to rise, and the absolute differences decreased significantly, but there was a weak trend of dispersion in the later period.

The digital finance development efficiency discussed in this paper essentially belongs to technical efficiency, which reflects the input-output status of the industry. Technical efficiency refers to the situation where it is technically impossible to increase any output (and/or decrease any input) without reducing another output (and/or increasing any other input) in a given feasible input/output vector (Koopmans, 1951). Debreu (1951) and Farrell (1957) further clarified the meaning of technical efficiency, which refers to the maximum proportional reduction of all inputs being consistent with the equivalent output, reflecting the degree to which the output approaches the production frontier under given input conditions.

The accurate measurement of technical efficiency depends heavily on appropriate methods. Aigner and Chu (1968) provided a deterministic method for measuring technical efficiency, achieving the leap from theory to practice. They specified the Cobb-Douglas production function (C-D function) as the measurement model, used quadratic programming as the estimation method, assumed that all deviations from the production frontier were attributable to inefficiency, and restricted the "residual" to be positive. The main shortcomings of this method are that the selected model is parametric and deterministic, and requires a prior function form that does not allow for measurement errors and other statistical noise.

In the late 1970s, the data envelopment analysis (DEA) method emerged, reigniting people's interest in the theory and practice of measuring technical efficiency. The DEA was first proposed by Charnes et al. (1979), who defined the optimal frontier as the envelope of all observed production possibilities and solved it using mathematical programming methods. Improvements were made by Fare and Lovell (1978), Banker and Cooper (1984), and Fare et al. (1985). The DEA is popular in evaluating technical efficiency mainly because it is non-parametric and can handle multiple outputs. Its shortcomings are that it also requires a deterministic model and does not allow for measurement errors.

As the deviation from the production frontier cannot be entirely attributed to inefficiency, and may also be related to measurement errors and other statistical noise (such as adverse weather). Aigner et al. (1977) and Meeusen and van den Broeck (1977) proposed the stochastic frontier model (SFM), which was further refined and presented in academia by Greene (2008) and Kumbhakar and Lovel (2003). The SFM is a measurement method that allows for both inefficiency and measurement errors and has outstanding noise handling capabilities. It decomposes the random disturbance term into two independent parts: a random error term and an inefficiency term, and estimates the non-negative inefficiency term to analyze and study the factors influencing inefficiency. Cornwell and Schmidt

(1996) further validated that the error term in previous models could be further decomposed into inefficiency and noise, and also demonstrated that the SFM can be well applied to panel data. After incorporating fixed effects or random effects correction methods, the estimation of firm efficiency can be more accurate. Jondrow et al. (1982) specifically estimated inefficiency and suggested using the measurement of total error to estimate the expected value of the inefficiency component.

The development of the stochastic frontier model has extended to the consideration of sample structure. Yélou et al. (2010) demonstrated, based on an empirical study of the dairy industry in Quebec, Canada, that the threshold effect of the industry should be considered when using the stochastic frontier model. This means that heterogeneity generated by different groups can lead to distorted results. Lee and Lee (2014) proposed a truncated tail SFM, which optimized the complexity of calculations, and concluded through Monte Carlo simulations that "this model approximates the distribution of inefficiency precisely, as the data-generating process not only follows the uniform distribution but also the truncated half-normal distribution if the inefficiency threshold is small" (p. 1). Tsionas (2019) proposed a stochastic frontier model that simultaneously considers technological heterogeneity and universality, making it possible to take into account the production efficiency of different countries. Chen (2014) proposed dividing the industry into several groups based on technological level and measuring the efficiency of enterprises in each group through empirical studies of the computer industry in Taiwan, China. Mastromarco et al. (2012) used the PCCE (Pooled Common Correlated Effects) estimation method to analyze foreign direct investment (FDI) and other indicators of openness in a stochastic frontier model, overcoming the potential bias problem of fixed effects in panel stochastic frontier models.

The SFM not only has the advantage of error handling but also has better actual measurement efficiency than the DEA. Specifically, the choice of production function plays a decisive role in the performance of DEA and SFM, indicating that SFM has better estimation results when the relevant production function is clearly selected, while DEA performs better when the function form is severely non-standard (Gong and Sickles, 1992). Since there is a rank correlation between SFM estimation and the maximum likelihood combination error, this means that the stochastic frontier adjustment of the combined error will not change its ranking, and therefore will not change the ranking results of the deterministic model based on maximum likelihood estimation.

Some domestic studies have used the SFM to measure the efficiency of regional R&D (Research and Experimental Development) innovation, financial development of "Belt and Road" countries, and Chinese OFDI (Outward Foreign Direct Investment) (Bai et al., 2009; Zhang and Yang, 2020). Given those advantages of SFM in measuring technical efficiency, we also adopts it as a measurement model.

3. Research Design

3.1. Samples and data

We need to select appropriate data of digital finance firstly. At present, there are two representative data of digital finance in China, which one is the China digital inclusive finance index developed by the Internet Finance Research Center of Peking University, and the other is the urban digital finance index jointly developed by the Guangzhou Institute of International Finance and Guangzhou University (Guo et al., 2020; Liao et al., 2022). The former includes three first-level indicators of coverage scope, service depth and digitized degree. It is compiled by the data of Ant Financial user from 2011 to 2020, revealing the development of digital finance of 31 provinces, 337 prefecture-level

cities and 2,800 counties in China. It has been widely recognized and used in the academic community. The latter compiles the urban digital financial index system from three dimensions of digital financial services, digital financial technology and digital financial operating environment, analyzes the overall and spatial characteristics of urban digital finance, and reveals the development of digital finance of 278 cities above the prefecture-level in China from 2010 to 2020.

Based on our research objectives and data availability, we choose 31 provinces of China as samples. Referring to the work of Fu and Huang (2018) and Zhang and Hu (2022), we use the China Digital Inclusive Finance Index as the digital finance output data (the explanatory variable). The index has only been updated to 2020, thus the sample period is 2011-2020 (see Table 1).

Table 1. Digital inclusive financial index descriptive statistics of each province.

No.	Province	mean	std	min	25%	50%	75%	max	kurtosis
1	Beijing	275.92	108.78	79.41	220.56	281.38	358.89	417.88	−0.82
2	Tianjin	234.88	97.12	60.58	181.49	241.69	308.67	361.46	−0.86
3	Hebei	201.01	95.12	32.42	148.93	206.95	276.62	322.70	−0.94
4	Shanxi	204.74	95.46	33.41	150.08	215.56	277.73	325.73	−0.90
5	Neimenggu	201.76	91.30	28.89	153.08	222.24	268.30	309.39	−0.71
6	Liaoning	214.77	92.00	43.29	166.96	228.91	285.01	326.29	−0.74
7	Jilin	197.29	93.03	24.51	145.18	212.64	270.75	308.26	−0.80
8	Heilongjiang	199.30	90.67	33.58	148.00	215.91	270.24	306.08	−0.86
9	Shanghai	280.96	112.76	80.19	226.49	280.17	367.46	431.93	−0.84
10	Jiangsu	244.23	104.06	62.08	186.78	248.88	324.94	381.61	−0.93
11	Zhejiang	265.68	105.90	77.39	210.44	266.48	347.60	406.88	−0.87
12	Anhui	215.71	102.88	33.07	158.27	220.03	295.77	350.16	−0.90
13	Fujian	244.29	103.62	61.76	187.97	248.94	325.65	380.13	−0.92
14	Jiangxi	209.87	100.86	29.74	153.52	216.06	288.97	340.61	−0.89
15	Shandong	218.17	99.81	38.55	164.95	226.62	293.86	347.81	−0.85
16	Henan	207.49	103.19	28.40	148.22	214.23	288.55	340.81	−0.99
17	Hubei	227.06	105.02	39.82	171.11	233.31	310.93	358.64	−0.90
18	Hunan	205.63	96.68	32.68	152.60	212.04	280.39	332.03	−0.89
19	Guangdong	244.00	101.04	69.48	188.97	244.48	322.98	379.53	−0.96
20	Guangxi	204.76	96.80	33.89	147.63	215.28	282.42	325.17	−0.97
21	Hainan	220.64	98.98	45.56	163.60	230.95	301.20	344.05	−0.95
22	Chongqing	219.03	98.70	41.89	166.07	227.87	295.23	344.76	−0.87
23	Sichuan	212.21	96.14	40.16	158.24	220.45	287.68	334.82	−0.90
24	Guizhou	190.27	97.05	18.47	129.57	201.37	270.55	307.94	−0.99
25	Yunnan	199.64	97.04	24.91	144.44	210.55	278.41	318.48	−0.90
26	Xizang	185.91	99.02	16.22	122.30	195.56	267.14	310.53	−1.08
27	Shanxi	213.95	98.26	40.96	155.96	222.75	288.68	342.04	−0.94
28	Gansu	189.24	93.80	18.84	136.23	201.95	261.06	305.50	−0.85
29	Qinghai	182.35	94.77	18.33	124.99	197.77	257.39	298.23	−1.02
30	Ningxia	197.83	91.75	31.31	143.87	213.53	268.59	310.02	−0.88
31	Xinjiang	194.73	93.46	20.34	148.47	207.11	266.05	308.35	−0.73
	mean	216.23	98.55	40.00	161.45	225.21	293.15	341.22	−0.89

The results in Table 1 show that the five provinces of Shanghai, Beijing, Zhejiang, Guangdong and Fujian are among the best in digital finance development. Based on Table 1, it can be concluded that the digital inclusive financial index differs significantly across different provinces in China, with the mean ranging from a minimum of 185.91 to a maximum of 280.96 and the standard deviation ranging from a minimum of 91.30 to a maximum of 112.76. Furthermore, the distribution of the indicator across provinces shows a negative skewness, indicating that the tail of the distribution is skewed to the left. These results suggest that there are notable differences in the performance of this indicator across different provinces in China.

Referring to the regional classification standards of the National Bureau of Statistics in 2017, we divide the 11 provinces of Beijing, Tianjin, Liaoning, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, Hebei, Guangdong and Hainan into the eastern region, 8 provinces of Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, Hunan and Shanxi into the central region, 12 provinces of Guangxi, Shanxi, Gansu, Qinghai, Neimenggu, Xinjiang, Xizang, Sichuan, Guizhou, Yunnan, Ningxia and Chongqing into the western region. We further investigate the average level of digital finance development in the whole country and the eastern, central and western region, finding that the average level of the eastern region is consistently high, with a first-mover advantage, while the development of digital finance in the central and western regions is relatively lagging behind, lower than the national average.
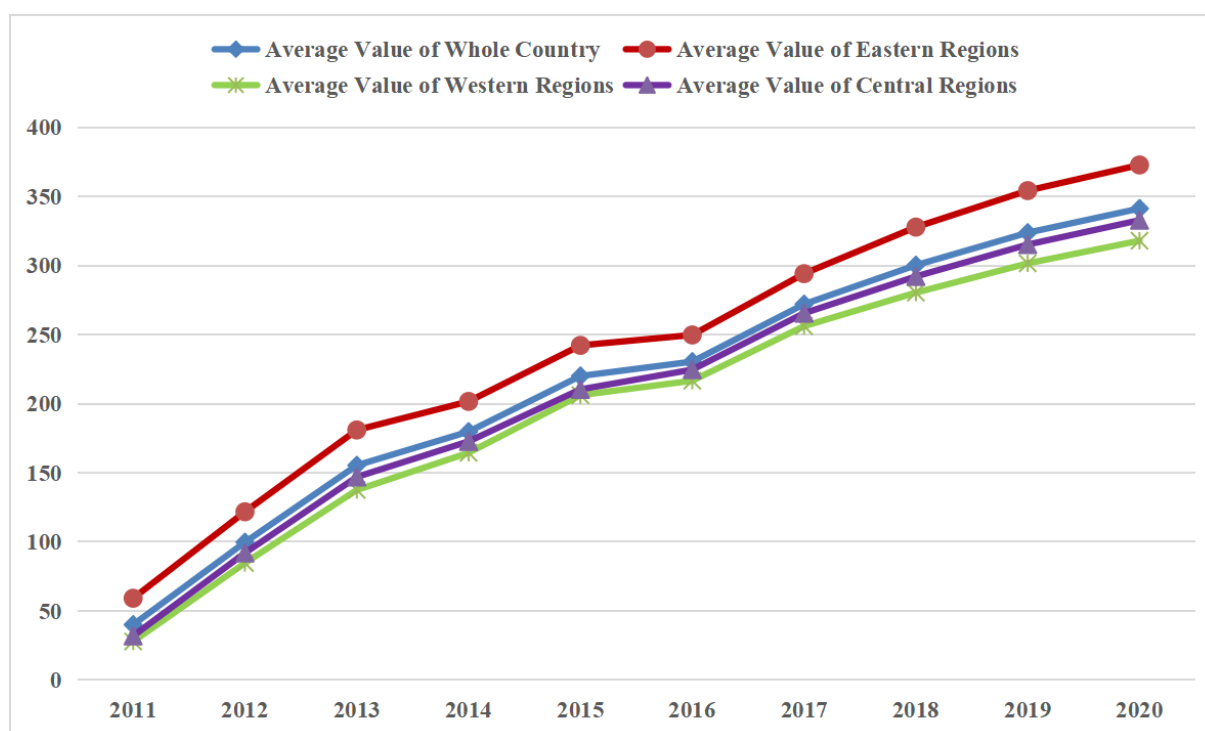

Figure 1. Average level of digital finance development in different regions.

The input data (explanatory variables) of digital finance development in each province is divided into two parts: capital input and labor input. The capital input is further divided into fiscal input and finance institution input. The former is represented by R&D investment in each province, and the latter is represented by the quantity of capital input in banking industry (accounting for about 68.6% of capital input in all finance institution), insurance industry (about 18% of capital input in all finance institution) and securities industry (about 12% of capital input in all finance institution) in each province. The labor

input is represented by the quantity of financial employees in each province. The input data is manly collected from state and each provincial statistical yearbooks and the database of RESSET.

3.2. Model setting

At present, the C-D function and translog production function (translog function) are the main models of the SFM, and both of them have their own advantages and defects (Christensen et al., 1973). The C-D function is simple and easy to understand. Its parameters can be concisely and directly expressed as the output elasticity of capital and labor, so its economic meaning is clear. The translog function relatively comprehensively takes into account the interaction between capital and labor, and their effect on real output, which better solves the problem that the alternative elasticity of the C-D function is 1.

Referring to the work of Battese and Coelli (1995), when just considering two inputs of capital (K) and labor (L), we construct the SFM based on the C–D function as shown in the following Equation (1).

$$\ln C_{it} = \alpha_0 + \alpha_1 \ln K_{it} + \alpha_2 \ln L_{it} + v_{it} - \mu_{it} \tag{1}$$

where the subscript i represents each sample province, and its value range is 1–31. The subscript t represents each year, and its value range is 2011–2020. The variable $\ln C$ represents the logarithm of the digital financial inclusion index of each province, which is the output term of the production function. The variables $\ln K$ and $\ln L$ are the logarithms of capital and labor of each province, and are the input term to the production function. The v is a non-negative stochastic variable related to the inefficiency of technique, and the μ is the observation error and other stochastic terms. The intercept term α_0 and coefficients α_1 and α_2 are the parameters to be estimated.

The translog function can be regarded as a variant of C–D function, and its form is shown in the following Equation (2).

$$\begin{aligned} \ln C_{it} = \beta_0 + \beta_1 \ln K_{it} + \beta_2 \ln L_{it} + \beta_3 (\ln K_{it})^2 + \beta_4 (\ln L_{it})^2 \\ + \beta_5 \ln K_{it} * \beta_6 \ln L_{it} + v_{it} - \mu_{it} \end{aligned} \tag{2}$$

The meaning of each parameter in Equation (2) is the same as that in Equation (1).

It is observed that the translog function can be regarded as an approximate second-order Taylor expansion of the output function $f(\ln K, \ln L)$ at the point $(0,0)$. If $\beta_3 = \beta_4 = \beta_5 = 0$, then the translog function is equal to the C-D function. Based on this, we choose the translog function as the benchmark model for subsequent estimation.

4. Empirical Analysis

4.1. Descriptive statistics

The descriptive statistical results of each variable are shown in Table 2. The values of $\ln C$、$\ln K$ and $\ln L$ are stable and can be estimated directly.

Table 2. Descriptive statistics results.

Variable	Median	Standard deviation	Minimum value	Maximum
$\ln C$	5.212	0.677	2.786	6.068
$\ln K$	5.411	1.533	0.182	8.155
$\ln L$	2.732	0.907	−0.223	4.557
$\ln K * \ln K$	27.62	6.356	7.763	36.820
$\ln L * \ln L$	8.285	4.218	0.000	20.770
$\ln K * \ln L$	16.100	7.870	−0.131	37.160

4.2. Model selection

(1) Individual Effect and First-order Difference Autocorrelation Test

In order to reveal the development state of digital finance in each province, we number the 31 provinces in Table 1 sequentially and name them $PR - 1$ to $PR - 31$. We use the least square dummy variable model (LSDV) to investigate the intercept term of individual heterogeneity, and show the results in Table 3.

The test results show that the p-value of dummy variable of most individuals is 0 and statistically significant, so the original assumption that all individual dummy variables are 0 can be rejected. We believe that there is an individual effect, so the mixed effect model is not fit for the data.

Then we use the first-order difference method to test the autocorrelation of panel data, and find that the first-order difference estimator FD is quite different from the intra-group estimator FE. Since the p-value of the first-order autocorrelation test is 0.0000, we believe that there is no autocorrelation.

(2) Time Effect Test

In order to test whether there is time effect in the model, we define the annual dummy variable, and use $YR - 1$ to $YR - 10$ to represent 2011–2020 sequentially. Taking 2011 as the base period, we show the test results in Table 4.

Table 3. Individual effect test results.

Variable	Coefficient	Variable	Coefficient
$ln\,K$	0.002 (0.05)	$PR-14$	−0.487*** (−11.67)
$ln\,L$	−0.165** (−2.58)	$PR-15$	0.020*** (4.40)
$ln\,K*ln\,K$	0.116*** (87.34)	$PR-16$	−0.239*** (−11.13)
$ln\,L*ln\,L$	0.046 (1.63)	$PR-17$	−0.282*** (−10.43)
$ln\,K*ln\,L$	−0.050*** (−3.11)	$PR-18$	−0.248*** (−10.90)
C	3.282*** (23.20)	$PR-19$	0.110*** (13.71)
$PR-2$	−0.422*** (−12.10)	$PR-20$	−0.537*** (−10.74)
$PR-3$	−0.250*** (−8.76)	$PR-21$	−0.942*** (−10.22)
$PR-4$	−0.463*** (−10.14)	$PR-22$	−0.431*** (−11.66)
$PR-5$	−0.546*** (−11.05)	$PR-23$	−0.206*** (−10.59)
$PR-6$	−0.257*** (−10.47)	$PR-24$	−0.664*** (−11.12)
$PR-7$	−0.547*** (−11.16)	$PR-25$	−0.591*** (−11.50)
$PR-8$	−0.455*** (−9.95)	$PR-26$	−1.274*** (−8.59)
$PR-9$	−0.144*** (−10.02)	$PR-27$	−0.321*** (−11.33)
$PR-10$	0.020 (0.83)	$PR-28$	−0.680*** (−11.51)
$PR-11$	−0.070*** (−10.34)	$PR-29$	−1.049*** (−10.49)
$PR-12$	−0.297*** (−11.28)	$PR-30$	−0.919*** (−10.95)
$PR-13$	−0.342*** (−11.49)	$PR-31$	−0.705*** (−10.51)

Note: *, * * and * * * represents statistically significant at 10%, 5% and 1% level respectively and t-value is in brackets.

Table 4. Time effect test results.

Variable	Coefficient	Variable	Coefficient
$\ln K$	0.023 (1.11)	$YR-4$	−0.444*** (−12.90)
$\ln L$	−0.021 (−0.60)	$YR-5$	−0.559*** (−13.69)
$\ln K * \ln K$	0.148*** (57.24)	$YR-6$	−0.586*** (−13.70)
$\ln L * \ln L$	0.025** (2.06)	$YR-7$	−0.689*** (−14.34)
$\ln K * \ln L$	−0.018** (−2.65)	$YR-8$	−0.753*** (−14.67)
C	1.656*** (28.10)	$YR-9$	−0.807*** (−14.69)
$YR-2$	−0.183*** (−9.65)	$YR-10$	−0.843*** (−14.75)
$YR-3$	−0.369*** (−12.39)		

Note: *, * * and * * * represent statistically significant at 10%, 5% and 1% levels respectively and t-value is in brackets.

The coefficients of annual dummy variables are significantly negative, preliminary indicating there is time effect. Furthermore, we conduct a joint test whether all annual dummy variables are 0, getting $F(9,30) = 66.18$ and Prob > F = 0.000. Thus, we reject the original assumption, accepting that there is time effect in the model.

(3) SFM with Mixed Effect and Fixed Effect Test

Firstly, we conduct the test of mixed effect for Equation (2), and show the results in Table 5. The coefficients of all variables are mostly statistically significant, suggesting that the overall explanatory power of the model is high.

Since the development of digital finance in different provinces is different, there may be some missing variables that do not change over time. Therefore, we further conduct the fixed effect test of robust standard error for Equation (2), and still show the results in Table 5. This time, $\rho = 0.981$, meaning that the variance of the compound disturbance term of the model mainly comes from the individual effect. Comparing the results of the two models, since the p-value of the fixed effect model is 0.0000, we reject the assumption that the individual residual error is 0 and select the fixed effect model finally. It means each province has its own intercept term.

Table 5. Test results of mixed effect and fixed effect model.

Variable	Mixed effect model	Fixed effect model
$\ln K$	0.088**	0.002
	(2.36)	(0.05)
$\ln L$	−0.074	−0.165**
	(−1.32)	(−2.72)
$\ln K * \ln K$	0.106***	0.116***
	(117.29)	(92.00)
$\ln L * \ln L$	0.039	0.046*
	(1.69)	(1.72)
$\ln K * \ln L$	−0.030**	−0.050***
	(−2.25)	(−3.28)
C	2.161***	2.856***
	(67.32)	(26.96)

Note: *, * * and * * * represents statistically significant at 10%, 5% and 1% level respectively and t-value is in brackets.

4.3. Estimation of efficiency term

We take the input-output data into the Frontier 4.1 (a software specially used to complete stochastic frontier analysis), obtain the maximum likelihood estimation (MLE) results of C–D function and translog function model respectively, and show the results in Table 6.

Table 6. MLE test results.

Coefficient	C-D function			Translog function		
	Estimated value	p-value	t-value	Estimated value	p-value	t-value
β_0	3.7451	0.1712	21.8751	2.1502	0.0693	31.0245
β_1	1.3339	0.0987	13.5106	0.1208	0.0590	2.0474
β_2	0.4324	0.1623	2.6631	0.2377	0.0770	3.0855
β_3	–	–	–	0.1064	0.0010	102.3278
β_4	–	–	–	−0.0190	0.0314	−0.6056
β_5	–	–	–	−0.0487	0.0227	−2.1466
σ^2	55.1885	15.0607	3.6644	0.0791	0.0227	2.6673
γ	0.9973	0.0008	1242.1860	0.9036	0.0393	22.9672

The results show that the estimated values of all parameters are mostly statistically significant at the level of 10%, suggesting both models have good explanatory power. The γ values of the two models are close to 1, indicating that the differences between the real output and the ideal output are mainly caused by the technical inefficiency term. The test result is consistent with our expectation. In addition, the technical inefficiency factor accounts for more than 90% of the variance of the joint stochastic disturbance term $v - \mu$, indicating that the error variance is mostly from the technical

inefficiency term, but little relates with the stochastic error. Therefore, it is rational and acceptable to choose the stochastic frontier production function with technical inefficiency term.

After obtaining the above results, we choose the log likelihood value to conduct LR test. The value of the one-sided likelihood ratio test statistic is 27.26, which is bigger than the critical value of 11.345 under the condition that the degree of freedom is 3 and the confidence level is 1%. Therefore, the original hypothesis $\beta_3 = \beta_4 = \beta_5 = 0$ is rejected, and the translog function is finally selected. However, the estimated annual digital finance development efficiency of each province may still not be exact because the inefficiency item is not included. Next, we continue to estimate the inefficiency term to obtain a more accurate value of total efficiency.

4.4. Estimation of inefficiency term

We use the OLS model, stochastic frontier semi-normal model and stochastic frontier exponential model to test the Equation (2) respectively, and show the results in Table 7.

Table 7. Test results.

Variable	OLS model	Stochastic frontier semi-normal model	Stochastic frontier exponential model
$\ln K$	0.088*	0.003***	0.000***
	(1.74)	(450.89)	(147.81)
$\ln L$	−0.074	−0.003***	−0.000***
	(−1.23)	(−375.47)	(−21.96)
$\ln K * \ln K$	0.106***	0.102***	0.095***
	(101.10)	(348,463.28)	(321,467.60)
$\ln L * \ln L$	0.039	0.002***	0.000***
	(1.44)	(586.41)	(47.97)
$\ln K * \ln L$	−0.030*	−0.001***	−0.000***
	(−1.73)	(−617.39)	(−134.76)
C	2.161***	2.458***	2.631***
	(27.53)	(237,745.71)	(294,983.47)

Note: *, * * and * * * represent significance at 10%, 5% and 1% levels respectively, and t value is in brackets.

Comparing the test results of the three models, we find that the result of the OLS model is relatively undesirable, so we abandon it. The λ-value obtained by the stochastic frontier semi-normal model is bigger, indicating that the inefficiency term plays a dominant role in the compound disturbance term, and the original hypothesis should be rejected. In addition, we can also reject the original hypothesis according to the p-value and likelihood ratio test. That is to say, there is inefficiency term, and the SFM can be used. Finally, we test the stochastic frontier exponential model, and draw the same conclusion that the original hypothesis should be rejected.

According to the above test results, we use the stochastic frontier semi-normal model to estimate the technical inefficiency term of digital finance development in each province. The results are shown in Table 8.

Table 8. Estimation results of technical inefficiency term based on stochastic frontier semi-normal model.

Year Province	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Beijing	0.0297	0.0000	0.0196	0.0285	0.0487	0.0540	0.0771	0.0977	0.1143	0.1244
Tianjin	0.0688	0.0020	0.0067	0.0148	0.0302	0.0338	0.0530	0.0703	0.0850	0.0944
Hebei	0.2123	0.0192	0.0006	0.0028	0.0145	0.0205	0.0407	0.0532	0.0648	0.0739
Shanxi	0.2033	0.0156	0.0003	0.0041	0.0170	0.0249	0.0416	0.0537	0.0677	0.0764
Neimenggu	0.2470	0.0166	0.0004	0.0053	0.0201	0.0269	0.0405	0.0472	0.0592	0.0674
Liaoning	0.1353	0.0079	0.0020	0.0096	0.0253	0.0276	0.0448	0.0570	0.0678	0.0758
Jilin	0.3019	0.0210	0.0000	0.0035	0.0175	0.0212	0.0387	0.0495	0.0583	0.0668
Heilongjiang	0.2018	0.0204	0.0001	0.0042	0.0185	0.0238	0.0404	0.0494	0.0589	0.0658
Shanghai	0.0290	0.0003	0.0227	0.0306	0.0496	0.0517	0.0803	0.1024	0.1201	0.1315
Jiangsu	0.0629	0.0000	0.0061	0.0142	0.0317	0.0363	0.0586	0.0779	0.0928	0.1034
Zhejiang	0.0332	0.0001	0.0161	0.0239	0.0432	0.0448	0.0709	0.0918	0.1079	0.1183
Anhui	0.2060	0.0124	0.0005	0.0071	0.0183	0.0259	0.0465	0.0632	0.0772	0.0879
Fujian	0.0648	0.0010	0.0081	0.0149	0.0336	0.0372	0.0609	0.0795	0.0936	0.1042
Jiangxi	0.2379	0.0164	0.0004	0.0061	0.0175	0.0240	0.0447	0.0596	0.0717	0.0832
Shandong	0.1635	0.0090	0.0011	0.0068	0.0216	0.0269	0.0460	0.0613	0.0759	0.0872
Henan	0.2527	0.0250	0.0000	0.0036	0.0161	0.0237	0.0444	0.0590	0.0729	0.0827
Hubei	0.1555	0.0090	0.0028	0.0101	0.0248	0.0308	0.0533	0.0713	0.0846	0.0922
Hunan	0.2096	0.0148	0.0004	0.0037	0.0165	0.0212	0.0414	0.0544	0.0672	0.0782
Guangdong	0.0469	0.0000	0.0078	0.0136	0.0306	0.0340	0.0581	0.0773	0.0935	0.1039
Guangxi	0.1991	0.0188	0.0001	0.0037	0.0172	0.0240	0.0423	0.0563	0.0675	0.0756
Hainan	0.1238	0.0086	0.0022	0.0072	0.0267	0.0275	0.0490	0.0669	0.0772	0.0854
Chongqing	0.1433	0.0102	0.0022	0.0088	0.0231	0.0285	0.0492	0.0623	0.0750	0.0856
Sichuan	0.1536	0.0100	0.0010	0.0053	0.0202	0.0246	0.0448	0.0584	0.0702	0.0797
Guizhou	0.4089	0.0357	0.0015	0.0014	0.0118	0.0180	0.0371	0.0500	0.0585	0.0661
Yunnan	0.2962	0.0241	0.0000	0.0032	0.0158	0.0214	0.0394	0.0545	0.0636	0.0716
Xizang	0.4615	0.0488	0.0026	0.0000	0.0102	0.0158	0.0332	0.0494	0.0584	0.0684
Shanxi	0.1486	0.0112	0.0003	0.0066	0.0204	0.0262	0.0443	0.0592	0.0736	0.0838
Gansu	0.4010	0.0352	0.0006	0.0025	0.0144	0.0161	0.0335	0.0449	0.0562	0.0649
Qinghai	0.4132	0.0665	0.0032	0.0012	0.0128	0.0150	0.0329	0.0440	0.0535	0.0617
Ningxia	0.2225	0.0215	0.0005	0.0039	0.0205	0.0196	0.0398	0.0486	0.0591	0.0683
Xinjiang	0.3705	0.0264	0.0001	0.0030	0.0164	0.0177	0.0357	0.0472	0.0590	0.0666
Mean	0.2001	0.0164	0.0035	0.0082	0.0227	0.0272	0.0472	0.0619	0.0744	0.0837

Table 9. Test results of translog function.

Coefficient	Estimated value	p-value	t-value
β_0	2.4397	0.1267	19.2537
β_1	0.0398	0.1752	0.2271
β_2	−0.0443	0.2817	−0.1572
β_3	0.1006	0.0015	68.3865
β_4	0.0222	0.1088	0.2038
β_5	−0.0147	0.0589	−0.2493
σ^2	0.0034	0.0202	0.1657
γ	0.9813	0.0829	11.8373

The results show that the technical inefficiency term of digital finance development in each province basically shows a trend of declining at the beginning of the sample period and then of rising. Since 2013, the technical inefficiency term in eastern region has led to rise, and has been higher than that in central and western regions, indicating that the input-output marginal contribution of digital finance in eastern region is declining.

To ensure the robustness of the results, we also use the stochastic frontier exponential model to estimate the technical inefficiency term. Comparing the estimation results of the two models, we find that their correlation coefficient is as high as 91.5%, indicating that they are highly correlated, and the estimation results of both models are rational. Therefore, we bring the estimation results of the inefficiency term in Table 8 into Equation (2), and obtain new coefficient test results with the Frontier4.1 (see Table 9). The results show that after adding the inefficiency term, the γ-value rises, and most coefficients are statistically significant, indicating that the regression results are rational and the model setup is reasonable.

4.5. Analysis of estimation results

After completing the above tests, we finally get the estimation results of digital finance development efficiency of each province and show them in Table 10.

Table 10. Estimation results of digital finance development efficiency (%).

No.	Year Province	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
1	Beijing	97.43	99.61	99.11	98.52	96.77	96.25	94.20	92.67	91.20	90.47
2	Tianjin	91.78	98.47	98.35	97.99	97.49	97.66	96.09	94.52	93.30	92.47
3	Hebei	80.15	97.82	99.42	99.36	98.84	98.38	96.68	95.61	94.67	94.04
4	Shanxi	81.16	98.41	99.49	99.42	98.87	98.26	96.82	95.69	93.85	93.39
5	Neimenggu	77.71	98.34	99.49	99.39	98.67	98.14	97.00	96.47	95.15	94.44
6	Liaoning	86.99	99.10	99.52	99.30	98.26	98.04	96.60	95.57	94.48	93.92
7	Jilin	73.52	97.88	99.48	99.43	98.84	98.61	97.19	96.29	95.41	94.49
8	Heilongjiang	81.29	97.93	99.49	99.41	98.73	98.19	96.64	95.90	94.97	94.39
9	Shanghai	97.26	99.57	98.82	98.26	96.67	96.56	94.21	92.42	90.95	90.10
10	Jiangsu	94.26	99.64	99.62	99.49	98.78	98.54	96.72	95.15	93.96	93.12
11	Zhejiang	96.68	99.52	99.09	98.67	97.02	96.86	94.65	93.14	91.85	91.10
12	Anhui	80.93	98.73	99.51	99.38	98.93	98.45	96.76	95.36	94.22	93.39
13	Hainan	93.41	99.42	99.30	99.05	97.73	97.49	95.48	93.89	92.80	91.95
14	Jiangxi	78.40	98.30	99.48	99.35	98.83	98.35	96.55	95.29	94.40	93.41
15	Shandong	84.75	99.27	99.64	99.55	99.10	98.80	97.38	95.92	94.18	92.99
16	Fujian	77.07	97.41	99.49	99.46	99.02	98.44	96.82	95.69	94.58	93.85
17	Hubei	85.18	99.00	99.49	99.31	98.57	98.15	96.25	94.68	93.58	93.06
18	Hunan	80.58	98.49	99.51	99.45	98.98	98.73	97.16	96.10	94.88	94.19
19	Guangdong	95.34	99.54	99.54	99.44	98.69	98.45	96.65	94.84	92.93	91.93
20	Guangxi	81.61	98.18	99.50	99.45	98.94	98.46	96.89	95.64	94.53	93.84
21	Hainan	88.11	99.04	99.42	99.42	98.78	98.47	96.76	95.03	94.02	93.63
22	Chongqing	86.21	98.87	99.45	99.26	98.41	97.99	96.19	95.00	94.19	93.20
23	Sichuan	85.33	98.91	99.51	99.44	98.82	98.45	96.79	95.71	94.81	94.03
24	Guizhou	66.30	96.72	99.50	99.54	99.25	98.90	97.24	95.89	95.46	94.79
25	Yunan	74.11	97.74	99.52	99.47	98.92	98.51	96.91	95.51	94.81	94.08
26	Xizang	64.35	95.62	99.60	99.59	98.75	99.24	98.85	95.46	95.86	93.74
27	Shanxi	85.72	98.78	99.49	99.37	98.77	98.39	96.86	95.62	94.32	93.52
28	Gansu	66.59	96.37	99.41	99.37	98.87	98.75	97.50	96.28	95.53	94.93
29	Qinghai	64.99	92.60	98.85	99.09	98.95	98.60	96.36	95.49	95.18	94.40
30	Ningxia	79.43	97.50	99.29	99.24	98.47	98.42	96.25	95.29	93.92	93.24
31	Xinjiang	69.01	97.79	99.59	99.55	99.14	99.05	97.74	96.64	95.65	94.93
mean value		82.12	98.21	99.39	99.26	98.54	98.24	96.59	95.25	94.18	93.39

The results in Table 10 reveal some facts as following.

Firstly, from the vertical perspective, the development efficiency of digital finance in each province shows a trend of rising rapidly at the beginning of the sample period, then declines steadily and slowly. The development efficiency of digital finance in most provinces reached the maximum during 2012–2014, and declined continuously from 2016. The cause may be that digital finance in each

province is more supported by policies and markets at the initial stage of development, thus the marginal revenue of the industry is higher, and the input-output ratio is relatively larger. After a period of rapid development, the industry tends to be saturated and marginal revenue decreases. The same input can only bring less output, so the efficiency gradually decreases. Additionally, in order to regulate the industry and control the chaos of blind expansion, some government intervention appears, and may limit the rapid expansion. Ultimately, while the development level of digital finance in each province is rising slowly, the efficiency is gradually declining simultaneously.

Secondly, from the horizontal perspective, the development efficiency of digital finance in each province has evolved from divergence to convergence and then to divergence. Those provinces with higher development level of digital finance generally have relatively more development efficiency at the beginning of the sample period. For instance, the five provinces of Shanghai, Beijing, Zhejiang, Guangdong and Fujian have a relatively higher development level of digital finance initially, and the development efficiency of digital finance of them is also more. The development efficiency and development level complement each other, indicating that the development of digital finance may initially have scale effect and Matthew effect. So it is important to obtain first-mover advantage. Thereafter, although the development level of digital finance in those provinces continues to rise slowly, their development efficiency declines rapidly over time. The development efficiency of the above five provinces have ranked from the top five to the bottom five and been lower than their original value by 2020. On the contrary, the provinces of Qinghai, Guizhou, Gansu, Xinjiang and the others have a relatively lower development level of digital finance in the early stage, the development efficiency of digital finance of them has ranked the top of all provinces and been far higher than their initial values by 2020. The result is consistent with the aforementioned vertical analysis result. The cause may still be that as the development level of digital finance improves, its marginal product declines.

Finally, from various regions, the development efficiency of digital finance in the eastern, central and western region also shows a trend of first rising and then declining in the sample period (see Figure 2). The development efficiency of digital finance in the eastern region is the highest and grows fastest at the beginning of the sample period, but it also declines fastest after 2013. The development efficiency of digital finance in the central and western regions starts at a lower level and declines relatively slowly after reaching the maximum. This result is also consistent with the aforementioned vertical analysis result, which again shows that the development potentiality of digital finance in the western region is greater.
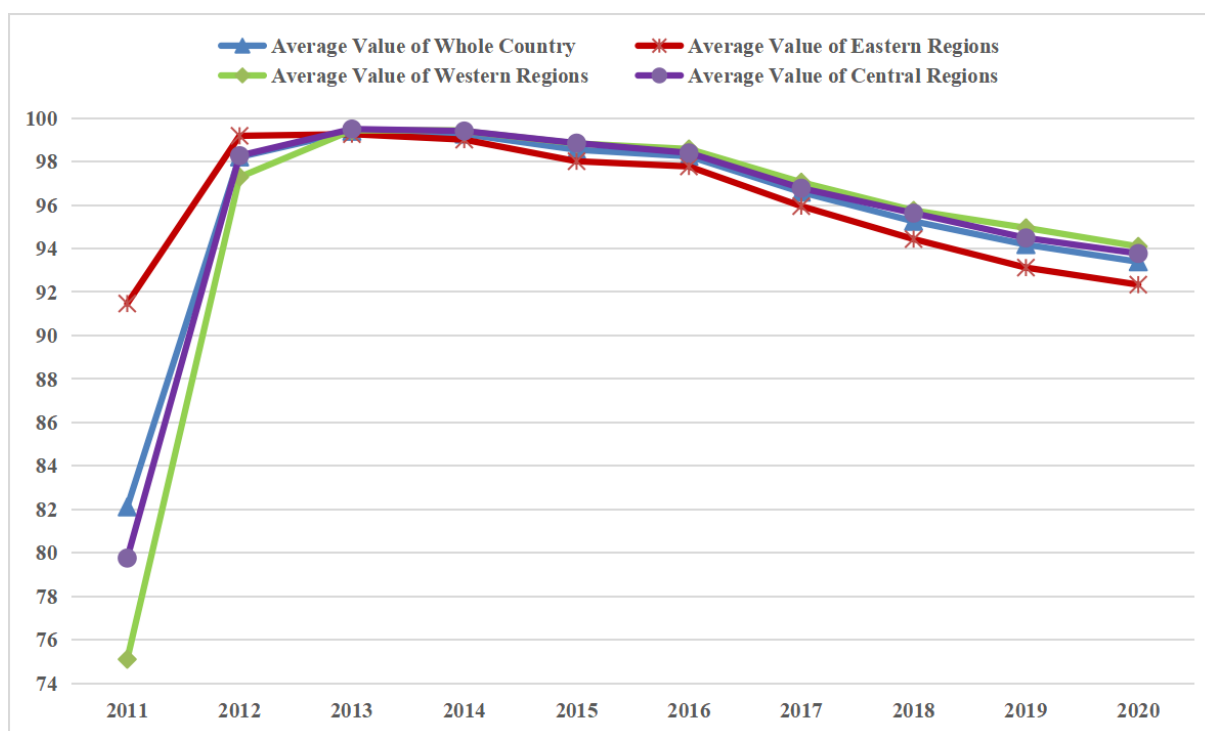
Figure 2. Average value of digital finance development efficiency in different regions.

5. Conclusions

Over the past decade, digital finance has gained significant development in all provincial regions in China, providing a solid foundation and strong impetus for sustainable economic development. Using the digital financial inclusion index for each year from 2011 to 2020 in 31 provincial regions across China as output indicators, we initially measured the digital financial development efficiency of each region with the help of SFM, and revealed its time-series evolution and spatial distribution differences. The results found that the digital financial development level of each provincial region showed a continuous upward trend in general, but the digital financial development efficiency showed a rapid rise and then a slow decline, convergence and then dispersion; the digital financial development efficiency of provinces with higher digital financial development level and eastern regions was higher at the beginning of the period and declined faster after reaching the peak, and its value was already lower than the national average; the imbalance of digital financial development efficiency of each provincial region. There is a tendency for the degree to expand, and the central and western regions have greater potential.

The findings of this paper have the following insights. Firstly, digital finance development plays a significant role, and in order to further promote the development of digital finance in all provinces in China, we should continue to stop the rapid decline of digital finance development efficiency in all regions (especially in the eastern provinces) and continuously improve the uneven digital finance development efficiency among regions. Secondly, the digital finance development efficiency in the eastern region declines faster after reaching the peak, and the central and western regions have more potential for digital finance development. The central and western regions have greater potential for digital financial development efficiency, so resources should be reasonably allocated to guide the reasonable flow of resources and continuously improve the quality of digital financial development in

each region. Thirdly, there are significant differences in the level and efficiency of digital financial development among different regions, so the central government should strengthen the overall planning and promote the coordinated development of each region through functional positioning, policy inclination and cross-regional mutual assistance.

Use of AI tools declaration

No artificial intelligence (AI) tools were used in the creation of this article.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

Aigner D, Lovell C, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. *J Econometrics* 6: 21–37. https://doi.org/10.1016/0304-4076(77)90052-5

Aigner DJ, Chu SF (1968) On estimating the industry production function. *Am Econ Rev* 58: 826−839. http://www.jstor.org/stable/1815535

Bai JH, Jiang KS, Li J (2009) Exploiting the model of the stochastic frontier to measure and evaluate the efficiency of the regional R&D innovation in China. *Manage World* 10: 51–61.

Banker R, Cooper WW (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manage Sci* 30: 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078

Banks E (2001) *e-Finance: the electronic revolution in financial services*, In: Banks E, 1 Edn., New York: John Wiley and Sons Press, Inc.

Barbesino P, Camerani R, Gaudino A (2005) Digital finance in Europe: Competitive dynamics and online behaviour. *J Financ Serv Mark* 94: 329–343. https://doi.org/10.1057/palgrave.fsm.4770164

Battese GE, Coelli TJ (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empir Econ* 20: 325–332. https://doi.org/10.1007/BF01205442

Charnes AW, Cooper WW, Rhodes EL (1979) Measuring the inefficiency of decision making units. *European Journal of Operational Research* 3: 339–338. https://doi.org/10.1016/0377-2217(79)90229-7

Chen K-H, Ghosh SN (2014) Threshold effects of technological regimes for the stochastic frontier model. *J Dev Areas*. https://doi.org/10.1353/jda.2014.0040

Christensen LR, Jorgenson DW, Lau LJ (1973) Transcendental logarithmic production frontiers. *Rev Econ Stat* 55: 28–45. https://doi.org/10.2307/1927992

Cornwell C, Schmidt P (1996) Production frontiers and efficiency measurement. The econometrics of panel data: a handbook of the theory with applications. https://doi.org/10.1007/978-94-009-0137-7_33

Debreu G. (1951) The coefficient of resource utilization. *Econometrica* 19: 273–292. https://doi.org/10.2307/1906814

Fare R, Grosskopf S, Lovell CAK (1985) The Measurement of Efficiency of Production[M]. *Springer Sci Bus Media* 6. https://doi.org/10.1007/978-94-015-7721-2

Fare R, Lovell CAK (1978) Measuring the technical efficiency of production. *J Econ Theory* 19: 150–162. https://doi.org/10.1016/0022-0531(78)90060-1

Farrell JM (1957) The measurement of productive efficiency. *J Royal Stat Society* 120: 253–290. https://doi.org/10.2307/2343100

Fu QZ, Huang YP (2018) Digital finance's heterogeneous effects on rural financial demand: Evidence from China household finance survey and inclusive digital finance index. *J Financ Res* 11: 68–84.

Gong BH, Sickles RC (1992) Finite sample evidence on the performance of stochastic frontiers and data envelopment analysis using panel data. *J Econometrics* 51: 259–284. https://doi.org/10.1016/0304-4076(92)90038-S

Greene WH (2008) The econometric approach to efficiency analysis, In: Harold O. Fried, C. A. Knox Lovell and S. S. Schmidt Greene W H, *The measurement of productive efficiency productivity growth*, New York: Oxford University Press, 92–250. https://doi.org/10.1093/acprof:oso/9780195183528.003.0002

Guo F, Wang JY, Wang F, et al. (2020) Measuring China's digital financial inclusion: Index compilation and spatial characteristics. *China Econ Q* 19: 1401–1418.

Jondrow J, Lovell CAK, Materov IS, et al. (1982) On the estimation of technical inefficiency in the stochastic frontier production function model. *J Econometrics* 19: 233–238. https://doi.org/10.1016/0304-4076(82)90004-5

Koopmans T (1951) An analysis of Production as an Efficient Combination of Activities. In: *Activity Analysis of Production and Allocation*, New York: John Wiley and Sons Press, Vol.13 of Cowles Commission for Research in Economics: 33–97. https://cir.nii.ac.jp/crid/1572824499992043008

Kumbhakar SC, Lovell CK (2003) *Stochastic frontier analysis*, In: Kumbhakar SC, Lovell CK, Cambridge University Press.

Lee S, Lee YH (2014) Stochastic frontier models with threshold efficiency. *J Prod Anal,* 4245–4254. https://doi.org/10.1007/s11123-013-0364-9

Liao GK, Li ZH, Wang MX, et al. (2022) Measuring China's urban digital finance. *Quant Financ Econ* 6: 385–404. https://doi.org/10.3934/QFE.2022017

Manyika J, Lund S, singer M, et al. (2016) Digital finance for all: Powering inclusive growth in emerging economies. *McKinsey Global Institute,* 1–5.

Mastromarco C, Serlenga L, Shin Y (2012) Is globalization driving efficiency? A threshold stochastic frontier panel data modeling approach. *Rev Int Econ* 203: 563–579. https://doi.org/10.1111/j.1467-9396.2012.01039.x

Meeusen W, van Den Broeck J (1977) Efficiency estimation from Cobb-Douglas production functions with composed error. *Int Econ Rev* 18: 435–444. https://doi.org/10.2307/2525757

Ondrich J, Ruggiero J (2001) Efficiency measurement in the stochastic frontier model. *Eur J Oper Res* 129: 434–442. https://doi.org/10.1016/S0377-2217(99)00429-4

Ozili PK (2018) Impact of digital finance on financial inclusion and stability. *Borsa Istanbul Rev* 184: 329–340. https://doi.org/10.1016/j.bir.2017.12.003.

Rizzo M (2014) Digital finance: Empowering the poor via new technologies. Washington DC: The Word Bank. Available from: https://www.worldbank.org/en/news/feature/2014/04/10/digital-finance-empowering-poor-new-technologies.

Scott SV, Van Reenen J, Zachariadis M (2017) The long-term effect of digital innovation on bank performance: An empirical study of SWIFT adoption in financial services. *Res Policy* 465: 984–1004. https://doi.org/10.1016/j.respol.2017.03.010

Tsionas EG, Tran KC, Michaelides PG (2019) Bayesian inference in threshold stochastic frontier models. *Empir Econ*, 56399–56422. https://doi.org/10.1007/s00181-017-1364-9

Yao L, Ma X (2022) Has digital finance widened the income gap? *Plos one* 172: e0263915. https://doi.org/10.1371/journal.pone.0263915

Zhang JY, Hu ZM (2022) Spatiotemporal characteristics and influencing factors of China's urban digital inclusive finance development. *J Southwest Minzu University (Humanities and Social Science)* 43: 108–118.

Zhang LY, Xing CH (2021) Distribution dynamics, regional differences and convergence of digital inclusive finance in rural China. *J Quant Technical Econ* 38: 23–42.

Zhang YT, Yang L (2020) National financial development of the "Belt and Road" countries and efficiency of China's foreign direct investment. *J Quant Technical Econ* 37: 109–124.