LUND UNIVERSITY

SCHOOL OF ECONOMICS AND MANAGEMENT

The impact of Digital Inclusive Finance on Green Total Factor Productivity under the conditions of Regional Development,

a Chinese case based on a spatial perspective

by

Xiao Zhang

May 2021

Master's Programme in Economics

Supervisor: Andreas Ek

Abstract

As a country with a high proportion of population and land area in the world, China is in a stage of changing its economic development direction, from the previous extensive economic growth to the intensive economic growth. At the same time, due to the limited total amount of resources and limited environmental bearing capacity on the earth, certain constraints are found in terms of amount and structure, which makes the contradictions existing in resources, environment and development of China become prominent. In this case: how to improve the efficiency of resource use, increase social productivity, and reduce environmental pollution has become a major problem that needs to be solved urgently.

To answer this question, the Chinese government proposes to accelerate the improvement of technology and enhance the efficiency of technology utilization, that is, to develop Green Total Factor Productivity（GTFP）. Therefore, on the basis of studying the development of GTFP, this article introduces the factors of Digital Inclusive finance services（DIF） and Regional Development Index（RDI）, focusing on the analysis: DIF's promotion of GTFP development, and RDI's regulatory role in it.

This essay takes 30 provinces in China from 2011 to 2017 as the research object. Based on manual data collection, a total of 210 panel data have been obtained. First, the spatial autocorrelation of DIF and RDI is analyzed through Global Morans'I and local Morans'I. Second, the spatial dubin model is used to empirically test the influence of DIF and RDI on the development of GTFP. Finally, the spatial effect decomposition is carried out on how the first two influence GTFP with spatial spillover effect. The empirical results are the same as the research hypotheses: 1. DIF and RDI have a positive spatial autocorrelation. 2. DIF will accelerate the development of GTFP. 3. RDI will reduce the effect of DIF on GTFP.

Key words: Green Total Factor Productivity, Digital Inclusive Finance, Regional Development Index, Spatial Autocorrelation, Spatial Dubin Model, Spatial Heterogeneity.

Acknowledgement

These days, I finally finished this essay. My heart is full of gratitude, and I am extremely grateful to the teachers, classmates and academic predecessors who helped me in writing my thesis!

I encountered many difficulties and problems in the writing of the thesis, but my supervisor, Andreas Ek, gave me a lot of help. He patiently helped me improve my thesis and provided a lot of suggestions for revision. I learned a lot from it, and finally completed my thesis on schedule. I am very grateful to him!

Due to my limited academic ability, the thesis may still have many deficiencies. I am extremely grateful if you can not hesitate to criticize and correct me, so that I can learn from it.

Table of Contents

1 Introduction 6

2 Theoretical framework 12

2.1 The definitions of GTFP、DIF and RDI and how to measure them 12

2.1.1 Green total factor productivity 12

2.1.2 Digital Inclusive Finance 16

2.1.3 Regional Development Index 18

2.2 The hypotheses 22

2.2.1 Hypothesis 1: DIF and RDI have a positive spatial autocorrelation. 22

2.2.2 Hypothesis 2: DIF will accelerate the development of GTFP 25

2.2.3 Hypothesis 3: RDI will reduce the effect of DIF on GTFP 27

3 Empirical analysis design 30

3.1 The methods we will use 30

3.2 The data we will use 32

3.3 Model setting 33

3.3.1 Spatial weight matrix 33

3.3.2 Spatial econometric model 34

3.3.3 Direct Effects、Indirect Effects and Total Effects 37

4 Empirical results and analysis 39

4.1 Spatial correlation analysis 39

4.2 Spatial econometric model analysis 42

4.2.1 Model checking 42

4.2.2 Model's results and analysis 42

4.2.3 Spatial effect decomposition result 46

4.3 Expanded analysis by region 50

4.4 Robustness test to solve endogenous problems 52

5 Conclusion 54

6 Reference List 57

7 Appendix 60

List of Tables and Figures

Table 2-1 Evaluation system of GTFP 12
Table 2-2 Peking University Digital Financial Inclusive Index 17
Table 2-3 Regional Development Index Evaluation System 19
Figure 3-1 Local Morans' I scatterplot in general 31
Table 3-2 Summary of variables in model 33
Table 4-1 Global Morans'I results 39
Figure 4-2 Global Morans'I of DIF during 2011-2017 40
Figure 4-3 Global Morans'I of RDI during 2011-2017 40
Figure 4-4 Local Morans'I scatterplot of DIF 41
Figure 4-5 Local Morans'I scatterplot of RDI 41
Table 4-6 The results of Spatial Dubin Model 42
Table 4-7 The direct effects results 48
Table 4-8 The Indirect effects results 49
Table 4-9 The total effects results 50
Table 4-10 The results of spatial Dubin Model by Region 51
Table 4-11 Decomposition of spatial effects by region 52
Table 4-12 Analysis of Endogenous Problems 53
Table 7-1 Variable descriptive statistics 60

1 Introduction

With the rapid development of society, countries all over the world pay more and more attention to green ecology and environmental friendliness. In particular, as developing countries are getting rid of extreme poverty, while emphasizing the increase in economic aggregate, they also focus on the level of development, instead of pursuing GDP expansion alone. At the same time, due to the limited total amount of resources and limited environmental bearing capacity on the earth, certain constraints are found in terms of amount and structure, which makes the contradictions existing in resources, environment and development of countries around the world become prominent. Therefore, "green development" has become a new requirement for development. In this case: how to improve the efficiency of resource use, increase social productivity, and reduce environmental pollution has become a major problem that needs to be solved urgently.

As a country with a high proportion of population and land area in the world, China is in a stage of changing its economic development direction, from the previous extensive economic growth to the intensive economic growth. In this stage, the problems and difficulties it faces are more severe than other countries. Because it accounts for a large proportion of the global economy and population, if it can solve this problem well, it will effectively promote the improvement of the global economy and the improvement of the world environment. In addition, its experience in dealing with this problem will also provide references for countries around the world. Therefore, this article will take China as a case to analyze the problem of improving resource utilization efficiency and reducing environmental pollution.

Green total factor productivity is proposed by mankind to answer the problem above, and it is currently receiving widespread attention in China. Scholars get the concept of Green total factor productivity from the total factor productivity, taking into account the needs of green development, and incorporating energy consumption and environmental pollution into the economic analysis framework, which is an improvement on the past economic development theories. But what are green total factor productivity and total factor productivity? And how can it solve the question above?

First, let us introduce total factor productivity. Total factor productivity is different from single factor productivity. Ordinary single factor productivity is usually the ratio of output to a certain factor input. For example, labor productivity measures the ratio of output to labor input. But in the actual production process, the demand for elements is not singular, but all elements such as labor, capital, and resources. The economic growth is determined by the factors of these inputs. Solow (1957) explained that the root of economic growth is the contribution of capital accumulation, labor growth and technological factors. It means that part of economic growth comes from the growth of input factors, and the other part is the growth of technological factors. The growth of this technological factor is the total factor productivity we mentioned. It can also be called the rate of technological progress. It measures the level of technology and the efficiency of technology utilization in the process of economic development. Since it measures the utilization efficiency of technology, it is difficult for us to obtain it directly from the total output and total input, and an indirect method is often needed for calculation. We can use the following formula for a preliminary understanding.

$$GY = GA + a * GL + b * GK$$

Where GY is the growth rate of the entire economy (usually a measure of GDP growth rate), GA is the total factor productivity, GL is the labor growth rate, and GK is the capital growth rate. a is the labor share and b is the capital share.

The green total factor productivity is an expansion of total factor productivity during the research process of scholars. Green total factor productivity still measures the level of technology and the efficiency of technology utilization in the process of economic growth. But the difference between GTFP and TFP is that resource and environmental factors are incorporated into the analysis model of economic growth. That is to say that when measuring economic growth, scholars no longer believe that the input of economic activities is only capital, labor, and technology; the output is only GDP. In the initial research, Ranmanathan (2005) usually regarded environmental pollution as an input factor, which other researchers do well (Hailu & Veeman 2001). They believe that only by paying a certain environmental pollution cost, can economic production be carried out smoothly. But later scholars realized that the pollutants discharged in the production process should be regarded as an output. Although this output is what we do not expect. Simply put, the fundamental difference between green total factor productivity and total factor productivity is that in the framework of analyzing economic growth, whether input contains resource factors and whether output contains environmental pollution.

Therefore, we define green total factor productivity as total factor productivity with resource consumption as an input and pollutant emissions as an undesired output. The advantage of green total factor productivity compared with total factor productivity is that it takes into account the consumption of resources and environmental pollution by production in the process of economic growth. And this is very important. Because the current Chinese government is no longer only focused on pursuing the expansion of economic scale, but on the level of economic development. It means the government wants to reduce energy consumption and environmental pollution. Green total factor productivity can help it focus on improving the level of economic development.

For the calculation of green total factor productivity, we can refer to the calculation of total factor productivity and set up the formula:

$$GY = GA + a * GL + b * GK + c * GR$$

GY is the growth rate of the entire economy (GDP growth rate and pollutant emission growth rate), GR is the resource growth rate and c is the resource share. The rest are the same as before. Of course, in the following actual research, the calculation of the green total factor growth rate is not so simple. When conducting empirical analysis, we will introduce some new methods to calculate green total factor productivity. But in terms of its basic definition, this formula can help us get a preliminary understanding of its composition principle.

The increase in green total factor productivity means that the economy will grow while the factor input remains unchanged, and the emissions of pollutants will decrease. That is why we believe GTFP can effectively solve the problem mentioned at the beginning.

Below we propose two concepts that have received extensive attention from the Chinese government: digital financial inclusion and regional development index. We believe that they will have an impact on development of green total factor productivity. Their effect on GTFP will be analyzed in Chapter 2.

The first one we want to introduce is digital inclusive finance. With the development of the economy, due to different factor endowment structures and industrial production structures between regions, the development of regions has shown an imbalance, not only in China, but also in other countries. Developed regions are more attractive to capital and labor in terms of resource accumulation and labor, so they are becoming more and more developed, while underdeveloped regions continue to lag. Therefore, it is very important to solve the problem of regional development imbalance. Therefore, many countries put a focus of domestic economic development on balancing the regional differences, as promoting harmonious development between regions through support and assistance to backward regions. For instance, in China, many provincial governments have tried to build an inclusive financial system to solve the problem of regional development. Inclusive finance is defined as a financial system that can effectively and comprehensively provide services to all classes and groups of society. It can improve the availability of financial services through financial infrastructure constructing. It can provide convenient financial services at lower costs for social life, especially to low-income people in underdeveloped areas. In the era of rapid Internet development, traditional inclusive finance has also been further expanded. The widespread application of digital technology has allowed the rapid development of inclusive finance, making it become digital inclusive finance. Especially after the outbreak of the new coronavirus in 2020, digital technologies based on the Internet have received unprecedented attention. More and more people rely on online economic activities to reduce the risk of public health crises. The role of digital financial inclusion (an inclusive financial system based on digital technology) has been demonstrated as never before.

The second one to be introduced is the Regional Development Index (RDI). The regional development index is a comprehensive index that evaluates the level of regional development. It not only considers the level of economic development—urban population or industrial development structure, but also incorporates other factors that affect residents' lives into the analysis framework, such as the level of environmental protection in the region, the income gap of residents within the region, the equality of human rights, and the level of external openness. Therefore, we believe that the regional development index can measure and judge the overall status quo of a region's development in more dimensions.

In the next chapter, we will introduce in detail the specific definitions of DIF and RDI and how we will measure them in this article.

The main content we want to study is how DIF and RDI will affect GTFP, and what kind of economic relationship exists between them. We understand that there will be some endogenous problems between RDI and GTFP. GTFP can promote regional economic development, reduce the emission of pollutants in production, and then improve the level of regional development. The improvement of the level of development in a region may attract talents from surrounding areas, because developed regions will have more job opportunities and a better living environment. The accumulation of talents will improve the region's innovation capabilities, and then advanced technologies will be developed and applied, and finally GTFP will be improved. Of course, this is only part of the reason for the endogeneity between them, not all of them. But because in the following research, we mainly study how DIF affects GTFP, and how RDI changes the effect of DIF on GTFP. In other words, RDI plays a regulatory role in the process of DIF affecting GTFP. Therefore, RDI is not the main explanatory variable in our research, so we first simply assume that RDI is an exogenous variable for analysis. However, since endogeneity is an unavoidable problem in empirical analysis, after conducting empirical analysis, we will conduct a robustness test of endogeneity. Through the introduction of instrumental variables, two-stage least squares estimation is used to discuss the endogeneity of the model.

After understanding the relationship between the two factors and GTFP, we will put forward some policy recommendations to the government to promote their efficient and rapid development.

As we mentioned, countries around the world are increasingly emphasizing the reduction of energy consumption and the emission of pollutants in the process of economic development. Therefore, it is necessary to promote economic growth by developing green total factor productivity. The reason is that it allows us to achieve economic growth by improving the level of technology or improving the efficiency of technology utilization while the capital, labor, and resource inputs remain unchanged. This is a very good way to solve the problems raised at the beginning of the article. Next, we will analyze how DIF and RDI affect the development of GTFP, and how we promote the development of GTFP.

2 Theoretical framework

2.1 The definitions of GTFP、DIF and RDI and how to measure them

2.1.1 Green total factor productivity

As we mentioned above, the definition of green total factor productivity in this article is as follows. We take energy consumption as an input, and pollutant emissions as an undesired output into the analysis framework of the production process. We call the total factor productivity (technical progress rate) obtained based on this analytical framework as green total factor productivity.

Based on this definition, we will use the Malmquist Index to calculate green total factor productivity based on the Undesired-Super-Efficiency Slacks-Based Measure (SBM) Model of "energy consumption" and "environmental pollution". We will first introduce the input and output elements that we use to construct the production process, and then respectively introduce the Malmquist index method and the Undesired-Super-Efficiency Slacks-Based Measure Model we use. In the previous content, we mentioned such a formula:

$$GY = GA + a * GL + b * GK + c * GR$$

This formula helps us define the inputs and outputs in the production process. Therefore, we will first introduce what indicators we use as inputs and outputs in Table 2-1.

Table 2-1 Evaluation system of GTFP

Inputs	Labors	Number of employees in each region measured in units of 10,000 people
	Capital	Physical capital stock measured in units of 100 million RMB
	Energy	Total energy consumption by region measured in units of 10,000 tons of standard coal
Outputs	Expected output	Regional real GDP measured in units of 100 million yuan
	Undesired output	SO2 emissions measured in units of 10,000 tons COD emissions measured in units of 10,000 tons

Due to the need to study total factor productivity under environmental conditions, that is, the green total factor productivity mentioned in this article, Fare (1989) et al. used the directional distance function in the study to satisfy the out of the requirements of the efficiency evaluation process. The directional distance function is a production function that takes environmental factors into consideration and can help us measure the efficiency of production through linear programming. However, due to the problem of variable slack in the input and output process, there will be deviations in the results. Tone (2001) used the slips-based measure model in the research process to solve this problem. Therefore, we also focused on this theory in the research process.

However, because the model used by Tone is a standard model, the value obtained will not exceed 1, so it is difficult for us to distinguish the effective DMU (Decision Making Unit) with efficiency being 1. Therefore, we also refer to the research results of Li & Shi (2014) and use the super-efficiency SBM model to solve the problem that the measured value can only be less than or equal to 1. Finally, an undesired-super-efficiency SBM model proposed by Yishan Yu and Ping Wei (2021, p.45) is used in this article. The following is a specific explanation about the model.

> *"First of all, we assume that the production of a region is to input multiple factors and then obtain multiple expected and undesired outputs. Therefore, we refer to the environmental technology framework to establish a set of production possibilities, and then determine the Each city serves as a DMU to formulate the optimal production technology frontier. The set of environmental technology that measures the efficiency of the green economy, that is, the set of production possibilities, is defined as:*
>
> $$PPS = \left\{ \left(X, \overline{Y^g}, \overline{Y^b}\right) \middle| \bar{X} \geq \sum_{\substack{j=1 \\ j\neq 0}}^{L} \lambda_j x_j \, , \overline{Y^g} \leq \sum_{\substack{j=1 \\ j\neq 0}}^{L} \lambda_j y_j^g \, , \overline{Y^b} \geq \sum_{\substack{j=1 \\ j\neq 0}}^{L} \lambda_j y_j^b \, , L \leq e\lambda \leq \mu, \lambda_j \geq 0 \right\}$$
>
> *In the formula: the input-output mode of each city corresponds to m kinds of inputs, s1 kinds of expected outputs and s2 kinds of undesired outputs. X represents the m-dimensional input vectorY^g, represents the s-dimensional expected output, and Y^b represents the s-dimensional undesired output vector, corresponding to $X = (x_1, x_2, \cdots, x_L) \in R_+^m$ 、$Y^g = \left(y_1^g, y_2^g, \cdots, y_L^g\right) \in R_+^{s_1}$ 、$Y^b = (y_1^b, y_2^b, \cdots, y_L^b) \in R_+^{s_2}$.Among them X、$Y^g$meet the condition of strong disposability, and Y^b meet*

the condition of weak disposability. $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_L)$ *represents the L-dimensional weight vector.* $(L = 1, \mu = 1)$ *means variable returns to scale (VRS) in the production of green technology;* $(L = 1, \mu = 1)$ *means constant returns to scale (CRS) in the production process. Excessive investment is represented by the vector* $S^- \in R_+^{S_m}$*,* $S^b \in R_+^{S_2}$ *indicates excessive undesired output, and* $S^g \in R_+^{S_1}$ *indicates insufficient expected output. According to the super-efficiency SBM model, the undesired output-super-efficiency SBM model can be derived. The model is constructed as:*

$$\rho = \min_{\lambda,\bar{X},\overline{y^g},\overline{y^b}} \frac{\sum_{i=1}^{m} \frac{\overline{x_i}}{x_{io}}}{\frac{1}{S_1 + S_2}\left(\sum_{r=1}^{S_1} \frac{\overline{y_r^g}}{y_{ro}^g} + \sum_{r=1}^{S_2} \frac{\overline{y_k^b}}{y_{ko}^b}\right)}$$

$$s.t. X \geq \sum_{\substack{j=1 \\ j\neq 0}}^{L} \lambda_j x_j$$

$$\overline{Y^g} \leq \sum_{\substack{j=1 \\ j\neq 0}}^{L} \lambda_j y_j^g$$

$$\overline{Y^b} \geq \sum_{\substack{j=1 \\ j\neq 0}}^{L} \lambda_j y_j^b$$

$$\bar{X} \geq x_o, \overline{Y^g} \leq y_o^g, \overline{Y^b} \geq y_o^b$$

$$\overline{Y^g} \geq 0, \overline{Y^b} \geq 0, L \leq e\lambda \leq \mu, \lambda_j \geq 0$$

$$\overline{x_i} = x_{io} + s^- \ (i = 1, \cdots, m)$$

$$\overline{y_r^g} = y_{ro}^g - s^g \ (r = 1, \cdots, s_1)$$

$$\overline{y_k^b} = y_{ko}^b + s^b \ (k = 1, \cdots, s_2)$$

In the formula, $\bar{x}$、$\overline{y_r^g}$、$\overline{y_k^b}$ *represent the projected value (target value) of the input and output of the evaluated unit, and*x_{io}、y_{ro}^g、y_{ko}^b *are the original values."*

After the undesired-super-efficiency model is established, we still use the Malmquist productivity index to measure the growth rate of green total factor productivity. According to the method of Bin Qiu, Shuai Yang, Peijiang Xin (2008, p.23), from the perspective of input, GTFP can be expressed by the Malmquist productivity index.

$$M_i^t = \frac{D_i^t(y^t, x^t)}{D_i^t(y^{t+1}, x^{t+1})}$$

"This index measures the rate of change of GTFP from period t to period t+1 under technical conditions in period t. Among them, $D_i^t(y^t, x^t)$ *is the mixed distance function we introduced, which is the reciprocal of technical efficiency. Based on the same definition, the Malmquist Index under the technical conditions of t+1 period is:*

$$M_i^{t+1} = \frac{D_i^{t+1}(y^t, x^t)}{D_i^{t+1}(y^{t+1}, x^{t+1})}$$

To reduce the interference caused by random selection of periods, we calculated the geometric mean of the two MI indices of period t and period t+1, using it to express the change of GTFP from period t as the base period to period t+1.

$$M_i(x^{t+1}, y^{t+1}, x^t, y^t) = \left\{\left[\frac{D_i^t(y^t, x^t)}{D_i^t(y^{t+1}, x^{t+1})}\right]\left[\frac{D_i^{t+1}(y^t, x^t)}{D_i^{t+1}(y^{t+1}, x^{t+1})}\right]\right\}^{\frac{1}{2}}$$

$$= \frac{D_i^t(y^t, x^t)}{D_i^{t+1}(y^{t+1}, x^{t+1})}\left[\frac{D_i^{t+1}(y^{t+1}, x^{t+1})}{D_i^t(y^{t+1}, x^{t+1})} \times \frac{D_i^{t+1}(y^t, x^t)}{D_i^t(y^t, x^t)}\right]^{\frac{1}{2}}$$

$$= EC(x^{t+1}, y^{t+1}, x^t, y^t)TP(x^{t+1}, y^{t+1}, x^t, y^t)$$

The Malmquist productivity index can be decomposed into changes in relative technical efficiency and changes in technological progress. In the above formula, EC is the relative efficiency change index under the condition that the return to scale is constant and the factors are strong and can be handled. It measures the degree of the observation object's catching up to the production possibility boundary from period t to period t+1. TP is the technological progress index, which measures the extent of the expansion of the possible ij boundary from period t to period t+1. Four mixed distance functions are used in the Malmquist productivity index solution process $D_i^t(y^t, x^t)$、$D_i^t(y^{t+1}, x^{t+1})$、$D_i^{t+1}(y^t, x^t)$、$D_i^{t+1}(y^{t+1}, x^{t+1})$.*"*

Through the combination of the undesired-super-efficiency SBM model and the Malmquist Index, we can calculate the value of growth rate of GTFP in each province in China.

2.1.2 Digital Inclusive Finance

Some financial institutions or companies use the Internet and digital technology to share information and accelerate the flow of resources, thereby effectively reducing transaction costs and access thresholds for financial services, so that some people who were previously excluded from traditional financial services can obtain efficient financial services. The financial services provided by these companies are called digital inclusive financial services.

We believe that the digital inclusive finance in a region is the sum of digital financial inclusion services provided by institutions or enterprises in the region. The higher the level of digital financial inclusion services provided by enterprises in a region, the higher the level of digital inclusive finance in the region. The development of digital inclusive financial services for enterprises will effectively improve the availability of financial services in the region. In other words, the regional DIF used in this article can tell us how many difficulties the people will face in the region to obtain financial services. The higher the value of DIF, the better the development of digital inclusive financial services in the region, and the less difficult it is for people to obtain financial services.

We use the 2011-2018 Peking University Digital Financial Inclusive Index published by Feng Guo, Jingyi Wang, Fang Wang, Tao Kong, Xun Zhang, Zhiyun Chen (2020, p.1406-1409) to indicate the level of digital inclusive finance development in a region. The reason is that this indicator system can fully reflect the development level of digital inclusive finance in various regions of China. And this total index is divided into three dimensions: the coverage of digital financial services, the depth of use, and the degree of digitization.

The indicators covered in this evaluation system are summarized in the Table 2-2.

Table 2-2 Peking University Digital Financial Inclusive Index

First-level dimension	Secondary-level dimension	Specific indicators
Coverage breadth	User coverage	The number of Alipay accounts per 10,000 people
		Percentage of Alipay users tied to cards
		Average number of bank cards bound to each Alipay account
Usage depth	Payment business	Per capita payment
		Amount paid per person
		High frequency (50 times or more per year) active users accounted for 1 time or more active users
	Money Fund Business	Per capita purchases of Yu'e Bao
		Per capita purchase amount of Yu'e Bao
		The number of people who bought Yu'e Bao per 10,000 Alipay users
	Credit loan business	The number of users with Internet consumer loans per million Alipay adult users
		Per capita loans
		Per capita loan amount
		The number of Internet small and micro business loan users
		per million Alipay adult users
		Average number of loans per household for small operators
	Insurance business	number of insured users per million Alipay users
		Insurance per capita
		Insurance amount per capita
	Investment Business	The number of people participating in Internet investment and wealth management per 10,000 Alipay users
		Investment per capita
		Investment amount per capita
	Credit business	Per capita credit calls for natural persons
		The number of users using credit-based services per million Alipay users (such as finance, accommodation, travel, social networking)
Digitalization level	Mobile	Proportion of mobile payment
		Proportion of mobile payment amount
	Affordable	Average loan interest rate for small operators
		Average personal loan interest rate
	Credit	Proportion of Huabei Payments
		Proportion of Huabei's payment amount
		Proportion of Sesame Credit's free deposit (accounting for all cases where deposits are required)
	Facilitation	Percentage of user QR code payments
		The proportion of the amount paid by the user's QR code

There are three main steps in the calculation process of DIF index.

1. Dimensionless processing of the original data.

Since the evaluation system has both positive and negative indicators, and the unit of each indicator is also different, the data needs to be dimensionless processed.

$$x' = \frac{log\, x - log\, x^l}{log\, x^h - log\, x^l} * 100$$

Positive indicators: Fixed the 95% quantile of the actual value of the indicator data for each region in 2011 as the upper limit x^h and the 5% quantile as the lower limit x^l.

Negative indicators: Fixed the 95% quantile of the actual value of the indicator data for each region in 2011 as the upper limit x^l and the 5% quantile as the lower limit x^h.

2.Calculating the weight of the indicator using the analytic hierarchy process

Firstly, a judgment matrix is constructed for each dimension of the evaluation system. This judgment matrix shows the researchers' evaluation of the importance of each indicator. Then, the eigenvectors of the judgment matrix are constructed, and the eigenvalues are normalized to obtain the index weights.

3.Calculating the index

After the processed index data and the weight of each indicator are obtained, the value of DIF is obtained by the method of weighted arithmetic average.

$$d = \sum_{i=1}^{n} w_i d_i$$

Among them, d is the comprehensive index, w_i is the normalized weight of each evaluation index, and d_i is the score of a single indicator.

2.1.3 Regional Development Index

In this article, we define the regional development index as an index that evaluates the comprehensive development level of a region. It can reflect the degree of development of the region. The high index indicates that the comprehensive development level of the region is high and the region is well developed.

Since we have defined the Regional Development Index as a comprehensive evaluation of the level of regional development, we need to measure it through a comprehensive and complete evaluation standard. After referring to relevant literature and works, we decided to adopt an indicator system and calculation results on regional development index constructed by Zhao Na (2020, p.57-58). This evaluation system not only considers conventional economic development factors, but also takes the five development concepts (innovation, coordination, green, openness, and sharing) proposed by the Chinese government as part of the evaluation dimension. That is why we believe it can fully reflect the development level of various regions in China.

Table 2-3 Regional Development Index Evaluation System

Dimension	Indicators
economy	GDP per capita
	General public budget revenues of local finance
	Fixed asset investment in the whole society
	The proportion of the output value of the tertiary industries
	Average salary of employed employees in urban units
	Resident consumption level
	Per Capita Disposable Income of Urban Residents
Innovation	The ratio of domestic patent applications granted
	R&D expenditures of industrial enterprises above designated size
	Technical market turnover
	Number of new product development projects of industrial enterprises above designated size
	Number of private industrial enterprises
	number of legal entities in information transmission, software and information technology services
coordination	Proportion of urban population
	Urban registered unemployment rate
	Per capita income ratio of urban and rural residents (in rural areas as 1)
	Per capita consumption expenditure ratio of urban and rural residents (in rural areas as 1)
	Number of urban residents with minimum living guarantee
	Number of rural residents with minimum living guarantee

	Urban employment
green	Green coverage rate in built-up area
	Harmless treatment rate of domestic garbage
	Wastewater discharge
	Forest coverage rate
	Park green area per capita
	Sulfur dioxide emissions
	Industrial pollution control completed investment
open	number of foreign-invested enterprises
	Total import and export volume at the location of the operating unit
	Foreign exchange income from international tourism
	Registered capital of foreign-invested enterprises
	Total import and export of goods of foreign-invested enterprises
shared	Public finance general service expenditure ratio
	Public toilets per 10,000 people
	number of health technicians per 10,000 people
	Education expenditure per capita
	number of public library institutions
	Urban road area per capita
	General Public Service Expenditure of Local Finance
	number of participants in the basic medical insurance for urban employees at the end of the year
	number of employees participating in pension insurance
	number of buses and trams per 10,000 people

And she adopts the improved entropy method in the calculation process. Entropy method is a mathematical method used to describe the discrete level of variables. In an index evaluation system, the greater the dispersion level of an index, the greater the impact it has on the index. Therefore, we can use this method to determine the weight of each indicator to calculate the final index. And the improved entropy method has the following advantages: 1. It does not rely on other people's subjective thoughts, but considers the weight coefficient of the indicator information, so it can reduce human error. Specifically, it is determined by the degree of dispersion of indicators. The greater the degree of dispersion of the indicator, the smaller the entropy value, and the greater the corresponding weight, the greater the impact of the indicator on the quality score of the Regional Development Index, and vice versa. 2. It uses standardized transformation processing to make the evaluation result unique and no different results will appear. 3. The standardized transformation is helpful to reduce the influence of extreme values on the evaluation results.

And the following are the steps of the calculation the Regional Development Index (Na Zhao,2020, p.57-58):

"1. Construct the original matrix. Suppose we need to measure the quality of new urbanization development in a province. Since the measurement index system contains n dimensions, the initial matrix A of the measurement system can be formed:

$$A = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix},$$

x_{ij} $(i = 1,2,\cdots m; j = 1,2,\cdots,n)$ *Represents the j-th indicator score of province i.*

2. Dimensionless processing. Since the evaluation system has both positive and negative indicators, and the unit of each indicator is also different, the data needs to be dimensionless processed. Dimensionless processing using extreme value methods.

Positive indicators:

$$x'_{ij} = \frac{x_{ij} - min\ (x_{1j}, x_{2j},,,,,, x_{mj})}{max\ (x_{1j}, x_{2j},,, x_{mj}) - min\ (x_{1j}, x_{2j},,,,,, x_{mj})}$$

Negative indicators:

$$x'_{ij} = \frac{max\ (x_{1j}, x_{2j},,,,,, x_{mj}) - x_{ij}}{max\ (x_{1j}, x_{2j},,, x_{mj}) - min\ (x_{1j}, x_{2j},,,,,, x_{mj})}$$

The processed new matrix A' is as follows ($x'_{ij} \in [0,1]$*)*

$$A' = \begin{bmatrix} x'_{11} & \cdots & x'_{1n} \\ \vdots & \ddots & \vdots \\ x'_{m1} & \cdots & x'_{mn} \end{bmatrix}$$

3. Calculate the entropy value of the j-th index. First determine the proportion of the j-th index in the i measurement object p_{ij}*(*$p_{ij} \in [0,1]$*);then calculate the entropy value* e_{ij}*(*$e_{ij} \in [0,1]$*)of the j-th index, where the constant k is the number of samples m related,make* $k = \frac{1}{ln\, m} > 0$*,then:*

$$p_{ij} = \frac{x'_{ij}}{\sum_{i=1}^{m} x'_{ij}}$$

$$e_j = -k \sum_{i=1}^{m} p_{ij}\, ln(p_{ij})$$

4.Calculate the coefficient of variance of the j-th index. The smaller the entropy value, the larger the difference coefficient, and the higher the importance of the index; vice versa.

$$g_j = 1 - e_j$$

5. Calculate the weight w_j *of each indicator.*

$$w_j = \frac{g_j}{\sum_{j=1}^{n} g_j}$$

6. Calculate the comprehensive score of the new urbanization development quality of the i-th province:

$$S_i = \sum_{j=1}^{n} w_j * p_{ij*10000} \quad "$$

After the calculation, we can get the Regional Development Index of regions in China.

2.2 The hypotheses

2.2.1 Hypothesis 1: DIF and RDI have a positive spatial autocorrelation.

According to the "first law of geography" put forward by Tobler (1970): everything is related to other things, but things near are more relevant than things far away. Therefore, we can imagine: In a region, the distribution of the observed value of a variable may be determined by another nearby observation. The interdependence or determination relationship of the observations themselves is what we call spatial autocorrelation. Spatial autocorrelation can be used to describe a variable in Whether the space is clustered or scattered. For example, the positive spatial autocorrelation reflects the distribution of "High-High" and "Low-Low". This means that high-value variables and high-value variables are clustered together, while the low-value variables and low-value variables are clustered together. Negative spatial autocorrelation reflects the distribution of "High-Low" and reflects the dispersion of variables.

However, whether a variable has spatial autocorrelation is still determined its own characteristics.

We believe that in China, the values of DIF and RDI have positive spatial autocorrelation. In other words, regions with high DIF development levels are geographically close, and they cluster together. The regions with high regional development index are geographically close, and they are clustered together.

In economics research, we usually regard an economic factor, an economic activity or an industry as a variable to study its characteristics and its spatial distribution. Economic geography believes that such distribution may be "aggregation" or "dispersion." Let us take the textile industry as an example to explain why there are clusters or dispersions. The agglomeration of the textile industry may increase the demand for cotton. Local cotton suppliers will expand production and reduce costs through increasing returns to scale. At the same time, the agglomeration of industries is conducive to the exchange and development of production technology, thus forming the phenomenon of knowledge spillover. This is conducive to the development of the local textile industry. And this effect will become larger as more companies gather, thereby further accelerating the flow and agglomeration of companies in the region. This force is called agglomeration power (Krugman 1991); but the same industrial agglomeration may increase competition and decrease corporate profits, thereby prompting companies to stay away to reduce competition. This driving force can be called the separation force. In a certain industry in a certain region, the stronger one between the agglomeration power and the separation power determines whether the industry is agglomerated or dispersed.

One of the important reasons why we believe the region with high DIF and the region with High RDI will gather respectively in the space is that they are affected by a stronger gathering force than the separation force. The reason is as follows.

DIF is an inclusive finance service based on digital technology. This kind of economic activity that relies on the Internet or digital technology is often affected by a strong external effect. 1. DIF has high requirements for regional network development. It needs professional network service providers and good digital infrastructure. Secondly, the development of DIF requires a high level of human capital and puts forward certain requirements on the labor market. In addition, DIF relies on digital technology, which is prone to technology spillover and knowledge spillover. Marshall (1920) believes that the concentration of enterprises will facilitate the emergence of professional suppliers in the local area, the sharing of local labor markets by local enterprises, and the knowledge spillover of enterprises. And these are the conditions required for the development of DIF. The emergence of professional suppliers, the sharing of labor markets, and knowledge spillovers can effectively reduce the costs of companies that provide digital inclusive financial services in the region and improve production efficiency for these companies. Companies that provide inclusive financial services choose to gather for low-cost resources and high-level technology. 2. DIF has strong network externalities. The increase in consumers will lead to an increase in the utility of consumers who use the service. DIF can provide more ways for corporate financing, investment and consumer consumption. The more companies and consumers involved in DIF services, the more resources it can call, thereby providing more powerful help to individuals in the economy. Therefore, we believe that the larger the coverage of DIF, the more utility individuals can obtain from it. This is the meaning of network externalities. And such network externalities require the companies that provide the digital inclusive finance services to aggregate to achieve coverage expansion. Both things are the agglomeration power of the companies that provide digital inclusive finance services.

At the same time, in China, the development of DIF mostly relies on financial institutions supported or supervised by the Chinese government. They provide digital inclusive finance services in a certain area through fiscal policies or coercive administrative means, so they will be less affected by competition and have less profit requirement, which means the little dispersion power.

Therefore, the agglomeration power is greater than the dispersion force, and the companies that provide the DIF services will agglomerate in regions. Because we define the level of DIF development in a region as the sum of all companies that provide DIF services in this region. Therefore, the agglomeration of such companies can also be understood as the concentration of regions with high DIF development levels.

RDI measures the development level of a region under a multi-dimensional evaluation system. We can simply think that the more developed a region is, the higher its RDI. We can regard the gathering of RDI in space as the closer the two regions are, the closer their development level is. The reason is that adjacent regions usually have similar factor endowments. Based on similar factor endowments, adjacent areas usually have similar industries and similar development paths, so it is easier to have the same level of development. The same development level means the same RDI value, that is, high values are surrounded by high values; low values are surrounded by low values. And this is what we call agglomeration in space, or positive spatial autocorrelation.

2.2.2 Hypothesis 2: DIF will accelerate the development of GTFP

The development of DIF will have an impact on the behavior of enterprises and consumers in the economy. As DIF is an inclusive financial service, its effect on enterprises will be obvious and strong, so we first analyze what impact it will have on the investment and operating behavior of enterprises.

The development of DIF has effectively increased the coverage of financial services, reduced the degree of financial exclusion, and reduced the difficulty of financing for small businesses and start-ups (Agarwal S., Hauswald R, 2010). Hall B H, Lerner J (2010) believe that due to the long cycle, high risk, and high uncertainty of innovative projects, enterprises are subject to a lot of financial constraints when carrying out technological innovation. With the help of DIF, the company's external financing constraints have been eased. Sufficient financial support allows companies to do more R&D activities to improve technology. Jiayu Wan, Qin Zhou and Yi Xiao (2020) proved that DIF can effectively promote innovative behavior, based on this idea. The reduction of financing constraints and the development and application of new technologies give companies the opportunity to use advanced equipment and environmentally friendly production technologies to reduce pollutant emissions during the production process and to treat emissions in a harmless manner. These behaviors can effectively improve the development of GTFP.

The development of DIF is conducive to improving the efficiency of resource allocation and alleviating the problem of resource mismatch (Xiaoge Zhao, Shihu Zhong, Xiaoxin Guo, 2021). DIF affected the financial industry, from the very beginning. It can promote the financial industry to innovate in financial services, financial products and financial business models, thereby bringing about industry efficiency improvements in the financial industry and effectively reducing its operating costs. The efficiency improvement of the financial industry will form a spillover effect and promote the improvement of the efficiency of resource allocation in other industries. The main reason is that due to its own technological advantages, DIF can effectively obtain information from multiple parties, including enterprises and consumers, and alleviate the problem of information asymmetry in the market. Financial accelerator theory explains why incomplete information can lead to inefficient fund allocation and non-optimal investment in the lending market. DIF has alleviated the information asymmetry, which has brought about the improvement of the efficiency of fund allocation and better investment than before. At the same time, it can also promote the maturity of the technology market, accelerate the development of the factor market, and promote the improvement of the information market, effectively reducing the decision-making costs, transaction costs and payment costs of enterprises. This kind of improvement in resource allocation efficiency and reduction in enterprise costs can enable enterprises to produce more products with fewer resources, thereby realizing the accelerated development of GTFP.

The development of DIF means the improvement of the regional financial level. The extensive improvement of the financial level can promote the upgrading of the regional industrial structure (GOLDSMITHR, 1969). The widespread use of digital technology has accelerated the flow of digital information. It will enable the further integration of information, technology and resources; enable the transfer of new technologies more efficiently and bring about increased productivity in many industries. Finally, it achieved the goal of promoting the optimization and upgrading of the regional industrial structure. The new industrial structure generally has a higher GTFP than the old industrial structure.

In addition to affecting the investment and operating behavior of enterprises, as well as the industrial structure of the region, the development of DIF will in fact have an impact on the economic behavior of residents.

The development of DIF will effectively lower the threshold of financial services and enhance the availability of financial services. It will increase residents' savings, investment and borrowing channels, reduce the cost of obtaining financial services and increase sources of income. In this process, DIF not only eased the liquidity constraints of households through short-term borrowing, but also reduced the payment cost of household consumption through new payment methods. The increase in income sources, the improvement of capital liquidity, and the reduction of consumption costs have contributed to the increase in current household consumption in the region (Xingjian Yi, Li Zhou, 2018). The increase in residents' consumption has brought about an increase in the total demand in the region. In order to meet the demand, enterprises have expanded the scale of production and achieved economies of scale, which has brought about an increase in GTFP. At the same time, the development of DIF has not only alleviated the information asymmetry in the production process of enterprises, but also alleviated the information asymmetry in the consumption process of residents. Because of the increase in consumption choices, consumers will tend to choose high-quality and cheap products. Enterprises with high GTFP have a competitive advantage because the products they produce are more in line with consumer needs. Such a competitive advantage makes them stronger, but the number of low GTFP companies is gradually decreasing. Enterprise competition allows the survival of the fittest, leading to the development of regional GTFP.

Therefore, we believe that the development of DIF will have a significant impact on regional enterprises and households, thereby accelerating the development of GTFP.

2.2.3 Hypothesis 3: RDI will reduce the effect of DIF on GTFP

New urbanization refers to the level of urban development under a multi-dimensional evaluation. We can evaluate from the following different dimensions (in the later empirical analysis, the development level of RDI will also be measured from these aspects): the level of regional economic development, the ability of regional innovation, the coordinated development of urban and rural areas, the level of regional environmental protection, and the level of opening to the outside world. RDI reflects the comprehensive level of urban development in a region, not just the economic development. Therefore, we can think that the higher the RDI, the more developed the region is and the more it meets the requirements of human survival and social development.

We believe that the development of RDI will effectively reduce the effect of DIF on GTFP. There are two main reasons for this statement.

1. Areas with a higher level of RDI development generally have better overall development, and have fewer development disadvantages in the economic system, and there are fewer places where DIF can play a role. For instance, in an exam with a full score of 100 points, student A has scored 98 points, so DIF can only play a role to get 2 points, but for student B with 78 points, DIF has more chance to improve his scores. Not only that, in some areas where DIF plays a role, the high-level development of high RDI in some time will play a substitute role.

Here are a few examples about DIF playing a lesser role and RDI replacing DIF. First, for high RDI areas, companies usually already have a certain scale. At the same time, due to the improvement of the financial system companies, they will face fewer financing constraints, and they can get effective financial support. Therefore, in these areas, DIF will be less helpful to corporate financing. Secondly, high RDI areas usually have a more reasonable industrial structure, can efficiently use various resources, and produce more products under the constraints of certain resources and environmental carrying capacity. Then DIF can promote the upgrading of the industrial structure a little. In addition, the high RDI area can attract more talents and introduce advanced technology. Not only that, but also the government's investments in education will cultivate more high-quality labor, which will lead to an increase in technical level and technical efficiency. And this also overlaps with the role played by DIF. DIF helps companies acquire more talents and advanced technology by solving the problem of information asymmetry.

2. GTFP in high RDI areas is usually higher than low RDI areas. With a larger GTFP,

increasing the same GTFP value will result in a lower growth rate.

Considering the effect of DIF on GTFP, we should not only consider the absolute value of the increase, but also the proportion of its increase. We believe that RDI will reduce the effect of DIF on GTFP, which is based on this statement. The same example are the two students who took the exam above. Student A had 98 points in the last exam, and this time it increased by 2 points, which was only an increase of 2.04%; if student B improved by 2 points, it would increase by 2.5% from 78 points. In comparison, student B's growth rate is higher, and the improvement effect is more significant. The same is true for the relationship between DIF and GTFP.

3 Empirical analysis design

After the theoretical analysis of the three hypotheses, we will adopt the method of empirical analysis and use spatial econometric models to check whether the research hypotheses are consistent with the economic phenomena in real life.

3.1 The methods we will use

Since Hypothesis 1 discusses the spatial distribution of DIF, RDI, it is necessary for us to adopt a model that can measure their spatial autocorrelation. Therefore, we will use the Global Morans' I proposed by Patrick Alfred Pierce Moran (1950) and the Local Moran's I proposed by Anselin (1995) for analysis.

The definition of Global $Morans'I$ is as follows:

$$\text{Global Morans'I} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{S^2 \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}$$

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

$$\bar{x} = \sum_{i=1}^{n} x_i$$

Among them, x_i and $x_j$represent the measured value of DIF or RDI of province i and province j, n=30, $\bar{x}$ is the sample mean, w_{ij} is the spatial weight matrix. In general, $Global\ Morans'I \in [-1,1]$, when $Global\ Morans'I$ is a positive value, it indicates a positive spatial correlation, that is, "high-high adjacent, low-low adjacent"; when $Global\ Morans'I$ is When the value is negative, it means negative spatial correlation, that is, "high-low" agglomeration; if the $Global\ Morans'I$ index is close to 0, it means that there is no spatial correlation and the research objects are randomly distributed.

The definition of Local $Morans'I$ is as follows:

$$Local\ \text{Morans'I} = \frac{Z_i}{S^2}\sum_{j\neq i}^{n} w_{ij} Z_j$$

$$Z_i = y_i - \bar{y}$$

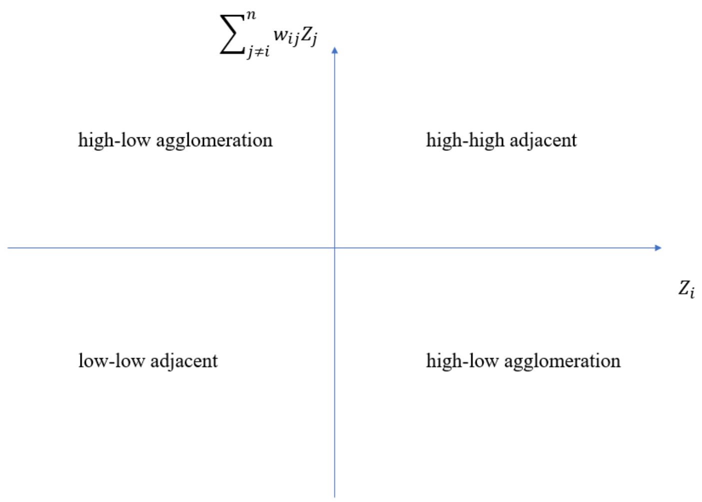
$$Z_j = y_j - \bar{y}$$

$$S^2 = \frac{1}{n}\sum(y_i - \bar{y})^2$$

Among them, y_i and $y_j$represent the measured value of DIF or RDI of province i and province j, n=30, $\bar{y}$ is the sample mean, w_{ij} is the spatial weight matrix.

In general, we can divide the results into four cases. 1. $\sum_{j\neq i}^{n} w_{ij}Z_j$ is greater than 0, and Z_i is greater than 0. It shows that the development level of the area is high, and the development level of the surrounding areas is also high. 2. $\sum_{j\neq i}^{n} w_{ij}Z_j$ is less than 0, and $Z_i$is less than 0. It shows that the development level of the area is low, and the development level of the surrounding areas is also low. 3. $\sum_{j\neq i}^{n} w_{ij}Z_j$ is greater than 0, and Z_i is less than 0. It shows that the development level of the area is low, and the development level of the surrounding areas is also high. 4. $\sum_{j\neq i}^{n} w_{ij}Z_j$ is less than 0, and Z_i is greater than 0. It shows that the development level of the area is high, and the development level of the surrounding areas is also low. The first two case shows the "high-high adjacent, low-low adjacent", and the last two cases show the "high-low" agglomeration. Therefore, we can use the local Moran scatter plot to represent the above four situations.

Figure 3-1 Local Morans' I scatterplot in general



When studying the spatial relationship of economic variables, Paelinck (1979) used spatial econometrics very early. Later, under the research of Anselin (1997) and other scholars, we have a good research method to study in this field. The hypotheses 2 and 3 are all studying the influence of DIF and RDI on GTFP in space, so we will use the spatial Dubin model by Anslin (1988) for analysis. The reasons are as follows: 1. Economic activities will show the characteristics of agglomeration and diffusion, so GTFP will be affected by both the DIF and RDI in the region and the DIF and RDI in the surrounding area. 2. As hypothesis 1 analyzes the spatial distribution characteristics of variables, it provides help for us to conduct research from a spatial perspective. 3. Due to the imbalance of regional development, the research based on spatial perspectives will help the government solve the problem of regional development incoordination.

3.2 The data we will use

Considering the availability of data, we will use panel economic data of 30 Chinese provinces from 2011 to 2017 for research. All data come from the "China Statistical Yearbook", "China Population and Employment Statistical Yearbook", the official website of the National Bureau of Statistics of China, the China Macroeconomic Database, the official statistical yearbooks of various provinces, and the "Peking University Digital Financial Inclusive Index" and published Journal papers, etc.

Considering that there are many factors that affect the development of regional green total factor productivity, not just the level of digital financial inclusion and the level of regional development. Therefore, this article also introduces some control variables that will also have an effect on GTFP. These include the total output value of the region, the population of the region, the freight volume of the region, and the government's expenditure on science and education.The Table 3-2 is a summary of the variables that will be used in the empirical model.

Table 3-2 Summary of variables in model

Type	Variables	Definition
Dependent Variable	GTFP	Green total factor productivity
Independent variables	DIF	The level of development in Digital inclusive finance
	DIF*RDI	The interactive items of digital financial inclusion and the level of regional development
	Cover	Coverage of digital financial inclusion
	Cover*RDI	The interactive items of Cover and the level of regional development
	Usage	The depth of use of digital financial inclusion
	Usage*RDI	The interactive items of Usage and the level of regional development
	Digit	The degree of digitization of digital financial inclusion
	Digit*RDI	The interactive items of Digit and the level of regional development
Control variables	ZCZ	Gro3ss industrial value of the region
	RK	Total population of the area
	HY	Regional freight volume
	KJZC	Fiscal expenditures of regional governments on technological development and education

3.3 Model setting

3.3.1 Spatial weight matrix

In spatial econometrics, we need to use a spatial weight matrix to define the distance between two regions. But the traditional spatial adjacency matrix is not suitable for the research of this article. The reason is as follows: the traditional definition of spatial adjacency matrix is that if two regions are adjacent, then $W_{ij} = 1$, otherwise $W_{ij} = 0$. It is impossible to effectively describe the economic development between the two places, only the geographical relationship between the two places can be measured. Therefore, we mainly consider economic factors and refer to the spatial distance weight matrix by Tiiupass and Friso Schlitte (2006), then construct a spatial weight matrix W that measures economic distance. Among them, $\overline{Y_i}$ represents the per capita GDP of province i from 2011 to 2017.

$$W = \begin{cases} W_{ij} = \dfrac{1}{\left|\overline{Y_i} - \overline{Y_j}\right|}, i \neq j \\ W_{ij} = 0, i = j \end{cases}$$

3.3.2 Spatial econometric model

The setting of the spatial econometric model should meet the actual needs of the research hypothesis and reflect the corresponding economic meaning. In the study of spatial econometrics, scholars mainly choose from the following three models: 1. Spatial Error Model (SEM), which assumes that spatial effects are reflected in error terms; 2. Spatial Autoregressive Model (SAR), which assumes that dependent variables will affect the economy of other regions through spatial interaction; 3. Spatial Dubin Model (SDM) combining SEM model and SAR model. When we choose the specific form of the spatial measurement model, we can refer to the LR test or Wald test results.

Therefore, we constructed the Durbin model 1 of the space panel under normal circumstances.

$$\mathrm{GTFP}_{it} = \beta_0 + \beta_1 \cdot W \cdot GTFP_{it} + \beta_2 \cdot \mathrm{DIF}_{it} + \beta_3 \cdot W \cdot DIF_{it} + \beta_4 \cdot (DIF_{it} \times RDIu_{it}) + \beta_5 \cdot W \cdot (DIF_{it} \times RDI_{it}) + \\ + \beta_6 \cdot ZCZ + \beta_7 \cdot \mathrm{RK} + \beta_8 \cdot HY + \beta_9 \cdot KJZC + \mu_i + \nu_i + \varepsilon_{it} --- (1)$$

In this model, we will use the growth rate of GTFP as the dependent variable $GTFP$. This is because our main research is the impact of DIF and RDI on the development of GTFP, so focusing our attention on the growth of GTFP will be more in line with the research requirements. The $W * GTFP$ represents the growth rate of GTFP in the surrounding area. It can help us measure the impact of the GTFP growth rate in the surrounding area on the GTFP growth rate in the region. The DIF represents the overall development level of digital financial inclusion in the region. It can help us measure the impact of the development level of DIF within the region on the GTFP growth rate in the region. The $W * DIF$ represents the development level of digital inclusive finance in the surrounding area. It can help us measure the impact of the development level of DIF in the surrounding area on the GTFP growth rate of the region. The $DIF * RDI$ represents the comprehensive development level of DIF and RDI in the region. It can help us measure how the RDI of the region will change the impact of DIF on the growth rate of GTFP when the DIF remains unchanged. The $W * (DIF * RDI)$ represents the comprehensive development level of DIF and RDI in the surrounding area. It can help us measure how the RDI in the surrounding area will change the impact of DIF in the surrounding area on the growth rate of GTFP when the DIF in the surrounding area remains unchanged. The ZCZ, RK, HY, and KJZC respectively represent the total output value within the region, the number of people in the region, the total freight volume in the region, and the part of government expenditures used to upgrade technology and education. They can help us measure the impact of these factors on the growth rate of GTFP within the region.

Then according to Model 1, the spatial Durbin model 2-4 containing the three sub-indices of DIF, cover, use, and digit, can be described respectively as below.

$$GTFP_{it} = \beta_0 + \beta_1 \cdot W \cdot GTFP_{it} + \beta_2 \cdot COVER_{it} + \beta_3 \cdot W \cdot COVER_{it} + \beta_4 \cdot (COVER_{it} \times RDI_{it}) + \beta_5 \cdot W \cdot (COVER_{it} \times RDI_{it}) + \\ + \beta_6 \cdot ZCZ + \beta_7 \cdot RK + \beta_8 \cdot HY + \beta_9 \cdot KJZC + \mu_i + \nu_i + \varepsilon_{it} --- (2)$$

$$GTFP_{it} = \beta_0 + \beta_1 \cdot W \cdot GTFP_{it} + \beta_2 \cdot USAGE_{it} + \beta_3 \cdot W \cdot USAGE_{it} + \beta_4 \cdot (USAGE_{it} \times RDI_{it}) + \beta_5 \cdot W \cdot (USAGE_{it} \times RDI_{it}) + \\ + \beta_6 \cdot ZCZ + \beta_7 \cdot RK + \beta_8 \cdot HY + \beta_9 \cdot KJZC + \mu_i + \nu_i + \varepsilon_{it} --- (3)$$

$$GTFP_{it} = \beta_0 + \beta_1 \cdot W \cdot GTFP_{it} + \beta_2 \cdot DIGIT_{it} + \beta_3 \cdot W \cdot DIGIT_{it} + \beta_4 \cdot (DIGIT_{it} \times RDI_{it}) + \beta_5 \cdot W \cdot (DIGIT_{it} \times RDI_{it}) + \\ + \beta_6 \cdot ZCZ + \beta_7 \cdot RK + \beta_8 \cdot HY + \beta_9 \cdot KJZC + \mu_i + \nu_i + \varepsilon_{it} --- (4)$$

In these models, we use the growth rate of GTFP as the dependent variable $GTFP$ too. And the $W * GTFP$,ZCZ.RK,HY and KJZC are the same as before. Then we will explain in detail the differences between Model 2-4 and Model 1.

In the model 2, the $COVER$ represents the coverage of digital inclusive financial services in the region. It can help us measure the impact of the coverage of digital inclusive financial services within the region on the GTFP growth rate in the region. The $W * COVER$ represents the coverage of digital financial services in the surrounding area. It can help us measure the impact of the coverage of digital inclusive financial services in the surrounding area on the GTFP growth rate of the region. The $COVER * RDI$ represents the comprehensive development level of COVER and RDI in the region. It can help us measure how the RDI of the region will change the impact of COVER on the growth rate of GTFP when the COVER remains unchanged. The $W * (COVER * RDI)$ represents the development level considering both COVER and RDI in the surrounding area. It can help us measure how the RDI in the surrounding area will change the impact of COVER in the surrounding area on the growth rate of GTFP when the COVER in the surrounding area remains unchanged.

In the model 3, the $USAGE$ represents the depth of use of digital inclusive financial services in the region. It can help us measure the impact of the depth of use of digital inclusive financial services within the region on the GTFP growth rate in the region. The $W * USAGE$ represents the depth of use of digital financial services in the surrounding area. It can help us measure the impact of the depth of use of digital inclusive financial services in the surrounding area on the GTFP growth rate of the region. The $USAGE * RDI$ represents the comprehensive development level of USAGE and RDI in the region. It can help us measure how the RDI of the region will change the impact of USAGE on the growth rate of GTFP when the USAGE remains unchanged. The $W * (USAGE * RDI)$ represents the comprehensive development level of USAGE and RDI in the surrounding area. It can help us measure how the RDI in the surrounding area will change the impact of USAGE in the surrounding area on the growth rate of GTFP when the USAGE in the surrounding area remains unchanged.

In the model 4, the $DIGIT$ represents the degree of digitization of digital inclusive financial services in the region. It can help us measure the impact of the degree of digitization of digital inclusive financial services within the region on the GTFP growth rate in the region. The $W * DIGIT$ represents the degree of digitization of digital financial services in the surrounding area. It can help us measure the impact of the degree of digitization of digital inclusive financial services in the surrounding area on the GTFP growth rate of the region. The $DIGIT * RDI$ represents the comprehensive development level of DIGIT and RDI in the region. It can help us measure how the RDI of the region will change the impact of DIGIT on the growth rate of GTFP when the DIGIT remains unchanged. The $W * (DIGIT * RDI)$ represents the comprehensive development level of DIGIT and RDI in the surrounding area. It can help us measure how the RDI in the surrounding area will change the impact of DIGIT in the surrounding area on the growth rate of GTFP when the DIGIT in the surrounding area remains unchanged.

In these models, $gtfp_{it}$ is the growth rate of GTFP of province i in year t, β_i is the regression coefficient of each explanatory variable, μ_i is the spatial effect, v_i is the time effect, and ε_{it} is the random error term. When μ_i=0, it means time fixed effect. When v_i=0, it means the spatial fixed effects model. When both μ_i and v_i are 0, it represents a mixed panel model.

3.3.3 Direct Effects、Indirect Effects and Total Effects

Since the model contains a spatial lag term, the regression coefficient of the lag term cannot directly represent the true spatial spillover effect of the variable on the explained variable. Therefore, we will decompose the spatial effects of independent variables into direct effects, indirect effects and total effects in the form of partial differentiation. Direct effects represent the influence of changes in independent variables on independent variables in the region. This effect also includes the "feedback effect" in which the spillover effect affects the surrounding area and then reacts to the area. Indirect effects represent the influence of independent variables on dependent variables in other regions. The total effect represents the spatial spillover effect of the independent variable on the population of dependent variables. The specific calculation steps are as follows (Yongjing Wang, Hui Li ,2021, p.159):

"First, convert the general form of the SDM model into a matrix form:

$$(I_n - \rho W)Y = \iota_n \beta_0' + \beta X + \theta W X + \varepsilon$$

$$Let\ \ P(W) = (I_n - \rho W)^{-1} 、 Q_m(W) = P(W) * (I_n \beta_m + \theta_m W)$$

Y is the dependent variable and X is the independent variable, then the formula can be transformed into:

$$Y = \sum\nolimits_{m=1}^{k} Q_m(W) X_m + P(W)\iota_n \beta' + P(W)\varepsilon$$

Then rewrite it into a matrix form:

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \sum\nolimits_{m=1}^{k} \begin{bmatrix} Q_m(W)_{11} & \cdots & Q_m(W)_{1n} \\ \vdots & \ddots & \vdots \\ Q_m(W)_{n1} & \cdots & Q_m(W)_{nn} \end{bmatrix} \begin{bmatrix} X_{1m} \\ \vdots \\ X_{nm} \end{bmatrix} + P(W)(\iota_n \beta' + \varepsilon)$$

m=1,2, ..., K, which means the m-th independent variable. $Q_m(W)_{ij}$ *is the matrix element in* $Q_m(W)$*. Take the partial derivative of the independent variable to get its direct effects, indirect effects and total effects on the dependent variable.*

$$Direct\ effects = \frac{\partial Y_i}{\partial X_{im}} = Q_m(W)_{ii}$$

$$Indirect\ effects = \frac{\partial Y_i}{\partial X_{jm}} = Q_m(W)_{ij}$$

$$Total\ effects = Q_m(W)_{ii} + Q_m(W)_{ij}"$$

4 Empirical results and analysis

Before analyzing the empirical evidence, perform descriptive statistics on the variables. See the appendix for Table 7-1.

4.1 Spatial correlation analysis

As we mentioned, we use $Morans'I$ to analyze the spatial correlation between DIF and RDI to test whether Hypothesis 1 is in line with reality.

The test results are shown in the following table. From 2011 to 2017, the DIF and RDI in all provinces of China were both significantly positive at the statistical level of 1%, indicating that the digital financial inclusion and the level of regional development in all provinces showed significant positive spatial correlation. In addition, from the changes in each variable shown in Figure 4-2 and Figure 4-3, the spatial autocorrelation of DIF and RDI has declined in fluctuations, indicating that the overall development of different regions' DIF and RDI tends to develop in a balanced manner.

Table 4-1 Global Morans'I results

Year	DIF	RDI
2011	0.365***	0.330***
2012	0.358***	0.337***
2013	0.350***	0.331***
2014	0.352***	0.328***
2015	0.363***	0.321***
2016	0.353***	0.308***
2017	0.309***	0.291***

Figure 4-2 Global Morans'I of DIF during 2011-2017

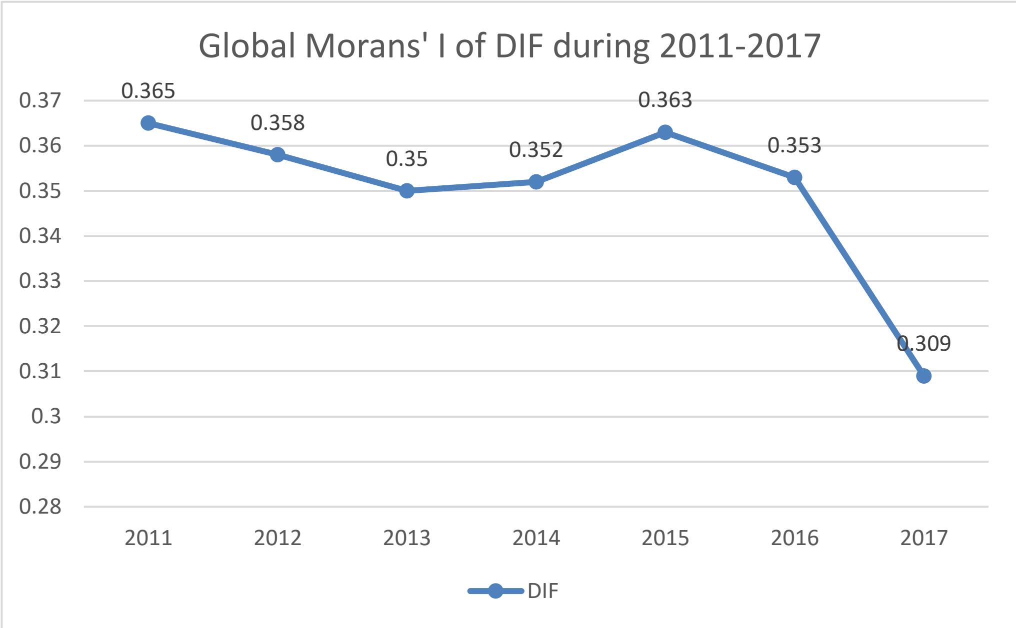
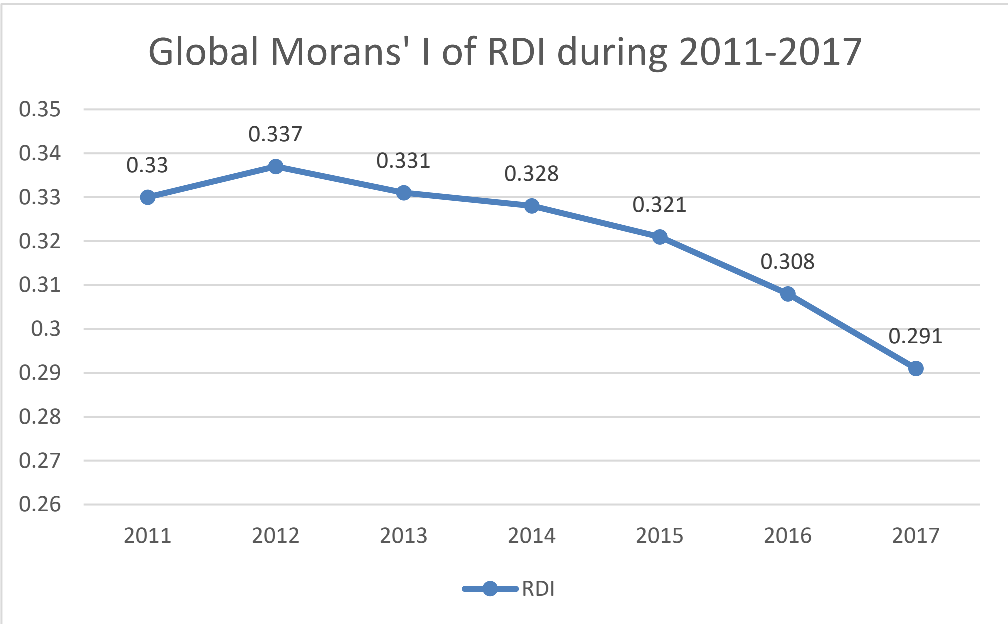


Figure 4-3 Global Morans'I of RDI during 2011-2017



After passing the $Global\ Morans'I$ test, the local spatial correlation analysis is carried out. And 2017 is selected as the time point to draw local Moran scatter plots of DIF and RDI respectively.

Figure 4-4 Local Morans'I scatterplot of DIF

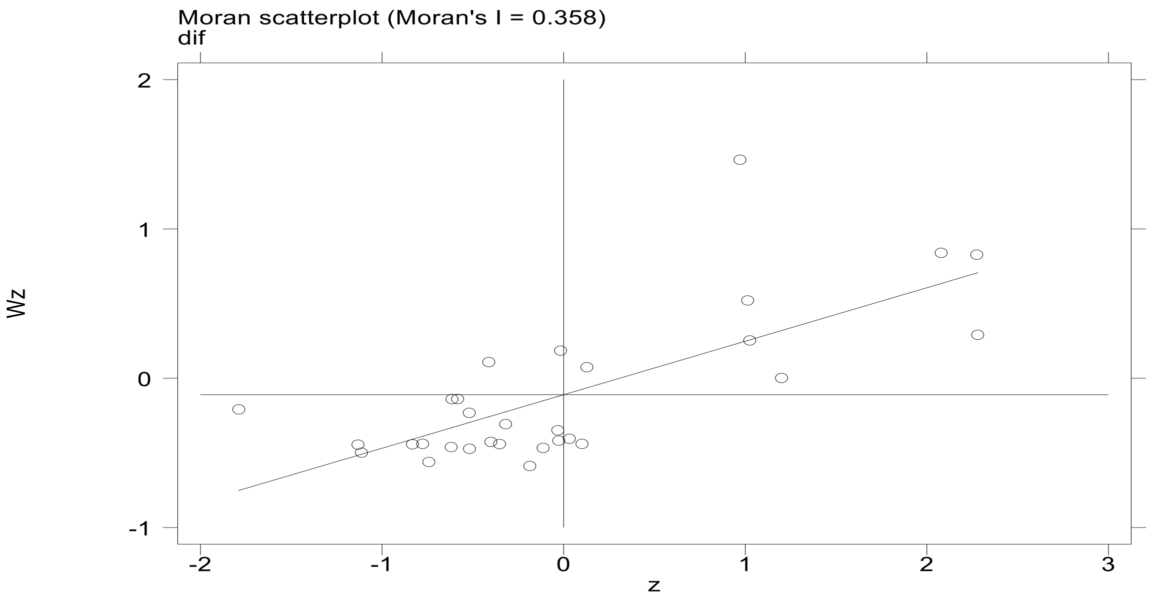
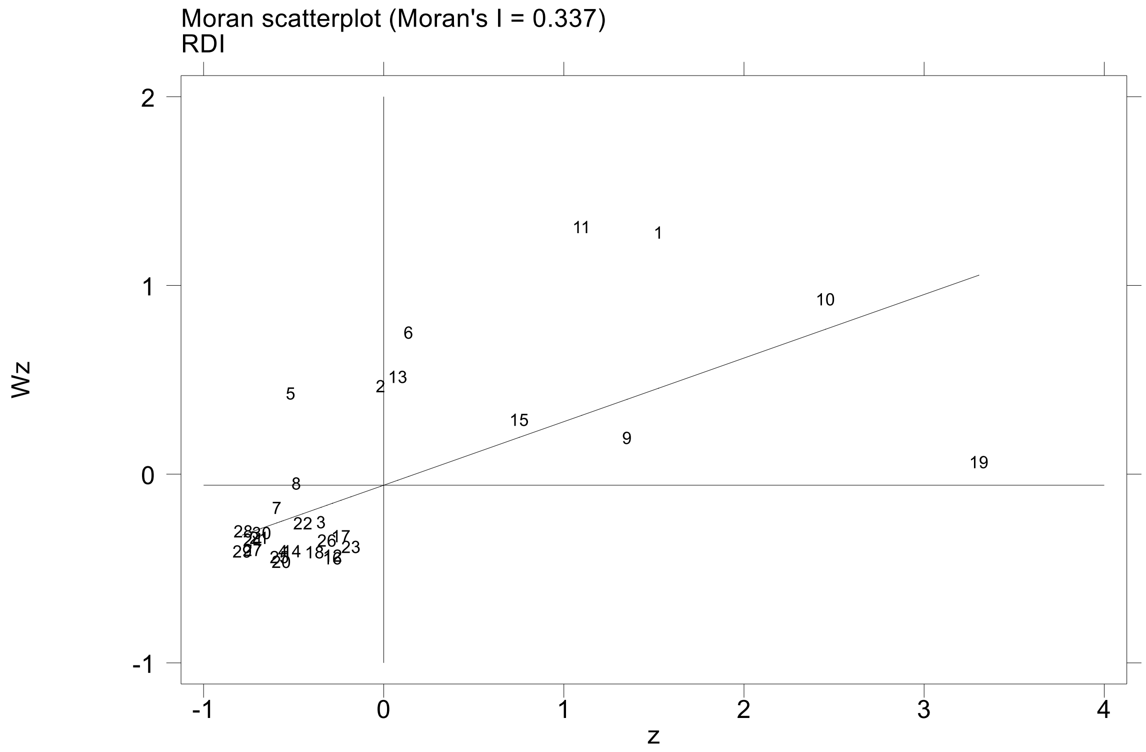


Figure 4-5 Local Morans'I scatterplot of RDI



We can find that the DIF and RDI of most provinces are concentrated in the first and third quadrants, and only a few provinces appear in the second and fourth quadrants, indicating that Chinese DIF and RDI show the obvious positive spatial correlation characteristics of "high-high" agglomeration and "low-low" agglomeration. Obvious spatial agglomeration phenomenon appears. This phenomenon in real life is consistent with that stated in

Hypothesis 1, so Hypothesis 1 is correct.

4.2 Spatial econometric model analysis

4.2.1 Model checking

In the above, we mentioned that there are three models to choose from when performing spatial econometrics, and initially adopted the spatial Doberman model. However, whether this model fits the research is still uncertain. We mainly used the LM test and LR test to get rid of this problem. The results are shown in the Regression Result Table below. The LR test statistics of each model significantly reject the null hypothesis, indicating that the SDM model cannot be degraded to the SAR model and the SEM model. At the same time, the Hausman test values all significantly reject the null hypothesis at the 1% level, so this paper selects the spatial Dubin fixed effects model.

4.2.2 Model's results and analysis

Table 4-6 The results of Spatial Dubin Model

VARIABLES	(Model 1)	(Model 2)	(Model 3)	(Model 4)
WGTFP	-26.70**	-24.05*	-28.66**	-19.27
	(12.96)	(13.03)	(13.23)	(13.23)
DIF	0.0780**			
	(0.0326)			
WDIF	-0.0439***			
	(0.0132)			
DIFRDI	-1.02e-05***			
	(2.00e-06)			
WDIFRDI	0.00150***			
	(0.000350)			
COVER		0.0171		
		(0.0118)		
WCOVER		-0.0468***		
		(0.0139)		
COVERRDI		-9.27e-06***		
		(2.11e-06)		
WCOVERRDI		0.00157***		
		(0.000374)		
USAGE			0.0326*	
			(0.0194)	

VARIABLES	(Model 1)	(Model 2)	(Model 3)	(Model 4)
WUSAGE			-0.0400***	
			(0.0136)	
USAGERDI			-7.40e-06***	
			(1.60e-06)	
WUSAGERDI			0.00145***	
			(0.000349)	
DIGIT				0.00229
				(0.0174)
WDIGIT				-0.0292***
				(0.00871)
DIGITRDI				-4.27e-06***
				(1.43e-06)
WDIGITRDI				0.000960***
				(0.000245)
ZCZ	0.150**	0.152**	0.139**	0.136**
	(0.0597)	(0.0599)	(0.0593)	(0.0583)
RK	0.650**	0.732***	0.863***	0.559**
	(0.276)	(0.283)	(0.260)	(0.260)
HY	-0.0554*	-0.0639**	-0.0567**	-0.0557*
	(0.0293)	(0.0293)	(0.0276)	(0.0284)
KJZC	1.05e-08***	9.20e-09**	5.77e-09	4.45e-09
	(3.96e-09)	(4.00e-09)	(3.73e-09)	(3.86e-09)
Log-likelihood	407.7566	404.1636	407.172	404.4729
LR-spatial lag	25.24***	35.55***	42.05***	27.78***
LR-spatial error	24.55***	36.05***	42.22***	27.42***
Spatial effect	Yes	Yes	Yes	Yes
Time effect	Yes	Yes	Yes	Yes
Observations	210	210	210	210

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The spatial correlation test results show that DIF and RDI are spatially related and heterogeneous. From the regression results of the spatial Dubin model, it can be seen that the spatial lag regression terms of models 1-3 are all significantly negative. Therefore, when analyzing the influence of DIF and RDI on GTFP, we cannot ignore the influence of spatial factors. The following is a detailed analysis of the results of the four models.

Model 1: It is not difficult to find that the log-likelihood of Model 1 is 407.5766, and all variables have passed the T test, and they are significant at the level of 1%, 5% or 10%.

Therefore, we can think that the result of the entire model is good, and these variables have a significant impact on the growth of GTFP. Then we look at the symbols of the variables, which shows how they will impact the growth of GTFP: to promote the development of growth of GTFP, or to hinder the development of growth of GTFP. The coefficient of DIF is positive, indicating that if DIF increases, then the growth of GTFP will also increase, that is, DIF will promote the development of GTFP. This result is consistent with Hypothesis 2. The coefficient of DIFRDI is negative, indicating that the increase of RDI will reduce the growth of GTFP, which means that RDI reduces the promotion effect of DIF on GTFP. This result is consistent with Hypothesis 3. The coefficients of WGTFP and WDIF are negative, indicating that when the GTFP and DIF in other surrounding areas have a higher degree of development, it will hinder the development of GTFP in this area. The main reason may be the siphon phenomenon in the process of economic development. The siphon phenomenon is a physical phenomenon, it refers to the existence of gravitational and potential energy difference between liquid molecules. The liquid will flow from the side with high pressure to the side with low pressure. But in economics, we can understand that cities with a higher level of development are more attractive to capital, labor, and resources, and they will encourage the resources of the surrounding underdeveloped cities to gather there. In Model 1, the higher development level of GTFP and DIF in the surrounding area means that they will be more attractive to capital, labor, and resources. This is not difficult to understand. With higher GTFP, companies can produce more products with less resources; higher DIF development level provides a better business environment for companies, and companies tend to migrate there, then fewer companies stay in the worse region. This result is not conducive to the development of GTFP in the region. And one of the control variables, HY has a negative coefficient, indicating that the larger the freight volume in this area, the more unfavorable the development of GTFP. Areas that rely on cargo transportation usually have primary and secondary industries as their main regional industries. Compared with the tertiary industry, these two industries consume more resources and energy, and at the same time, less added value of products is not conducive to the increase of GDP. And the consumption of resources did not bring more GDP, so it is not hard to understand why GTFP is difficult to achieve growth. The coefficients of the other three control variables are all positive, indicating that they can effectively promote the growth of GTFP. ZCZ will bring economies of scale, RK will bring increased consumption capacity and sufficient labor, and KJZC will effectively improve regional technological innovation and labor quality, which are all beneficial to GTFP.

Model 2: All variables except COVER passed the T test, and they were significant at the 1%, 5%, or 10% level. The reason why the coefficient of COVER is not significant may be as follows. In the evaluation system of DIF, COVER pays more attention to the scope and number of people covered by DIF, among which consumers have a great influence. But for the development of GTFP, the extent to which companies can profit from DIF is more critical. Therefore, consumers who affect COVER have less impact on the development of GTFP, so this may cause the coefficient of COVER to be insignificant. Then we analyze other variables that have significant effects. The coefficient of COVERRDI is negative, indicating that the increase of RDI will reduce the growth of GTFP, which means that RDI reduces the promotion effect of COVER on GTFP. This result is consistent with Hypothesis 3. The coefficients of WGTFP and WCOVER are negative, and the principle is same as that in Model 1. The coefficient signs of the four control variables are the same as those of model 1, so I will not repeat them.

Model 3: All variables except KJZC among the control variables passed the T test, and they were significant at the 1%, 5%, or 10% level. The reason why the coefficient of KJZC is not significant may be as follows. KJZC's role in promoting the development of GTFP is mainly to improve the skills of people through education, thereby improving the efficiency of technology utilization. But the main explanatory variable USAGE used in this model is the depth of DIF usage. Since DIF is a digital inclusive financial, it requires users to have a certain level of education. Therefore, the roles of KJZC and USAGE may overlap, so KJZC’s coefficient is no longer significant. Then we analyze other variables that have significant effects. The coefficient of USAGE is positive, indicating that if USAGE increases, the growth of GTFP will also increase. Therefore, DIF will promote the development of GTFP, and this result is consistent with Hypothesis 2. The coefficient of USAGERDI is negative, indicating that the increase of RDI will reduce the growth of GTFP, which means that RDI reduces the promotion effect of DIF on GTFP. This result is consistent with Hypothesis 3. The coefficients of WGTFP and WUSAGE are negative, and the principle is the same as it in Model 1. The coefficient symbols of the four control variables are the same as those of Model 1, so I will not repeat them again.

Model 4: All variables except W*GTFP, DIGIT and KJZC passed the T test, and they were significant at the 1%, 5%, or 10% level. The reason why they are not significant may be similar to why COVER in Model 2 and KJZC in Model 3 are not significant. DIGIT is determined by the consumer, but GTFP is determined by the enterprise. Therefore, DIGIT has no significant effect on GTFP. And DIGIT overlaps with KJZC, so their coefficients are not significant. The results of other variables are the same as these in the first three models, which can verify Hypothesis 2 and Hypothesis 3.

In summary, DIF and USAGE have a significant role in promoting the growth of GTFP. The increase of RDI effectively reduces this promotion effect. This phenomenon is consistent with hypotheses, indicating that hypothesis 2 and hypothesis 3 are correct.

4.2.3 Spatial effect decomposition result

The following three tables respectively show the direct effect, indirect effect and total effect results of the spatial effect decomposition of the SDM model. Direct effects represent the influence of changes in independent variables on independent variables in the region. This effect also includes the "feedback effect" in which the spillover effect affects the surrounding area and then reacts to the area. Indirect effects represent the influence of independent variables on dependent variables in other regions. The total effect represents the spatial spillover effect of the independent variable on the population of dependent variables.

Table 4-7 shows the direct effects of independent variables on dependent variables. It is not difficult to find that basically all independent variables have significant direct effects on the dependent variables, indicating that the independent variables in the region will have a significant impact on the dependent variables in the region. According to the results, DIF will significantly improve the development of GTFP, while RDI will significantly reduce the promotion of DIF and its three sub-indices on GTFP.

Table 4-8 shows the indirect effects of independent variables on dependent variables. Since most of the independent variables and all the control variables have not passed the significance test, we believe that DIF and RDI are unlikely to have a strong spatial spillover effect in promoting the development of GTFP. The reasons may be as follows: 1. DIF relies on the Internet and digital technology. Therefore, its development is mainly online, relying on less offline facilities. The construction of offline digital technology facilities can only cover the local area due to the administrative division system with Chinese characteristics, and it is difficult to affect the surrounding areas. Therefore, it is more difficult to form a spatial spillover effect. 2. The development of DIF shows a significant positive spatial correlation, and the degree of development of DIF shows a state of aggregation. Therefore, when the DIF development level of neighboring provinces is relatively close, it is more difficult to receive the influence of DIF development in the surrounding areas.

Table 4-9 shows the total effect of independent variables on dependent variables. The result is the same as the previous analysis: the spatial effect mainly exists in the interaction terms with DIF and RDI. It shows that due to the characteristics of online development, DIF has less spatial effect, but RDI can play a significant role in regulating the development of surrounding areas.

Table 4-7 The direct effects results

VARIABLES	(Model1)	(Model2)	(Model3)	(Model4)
WGTFP	-26.87**	-24.73*	-29.46**	-20.07
DIF	0.0731**			
WDIF	-0.0384***			
DIFRDI	-1.00e-05***			
WDIFRDI	0.00136***			
COVER		0.0144		
WCOVER		-0.0400***		
COVERRDI		-9.02e-06***		
WCOVERRDI		0.00140***		
USAGE			0.0315	
WUSAGE			-0.0344***	
USAGERDI			-7.24e-06***	
WUSAGERDI			0.00131***	
DIGIT				0.00809
WDIGIT				-0.0274***
DIGITRDI				-4.35e-06***
WDIGITRDI				0.000924***
ZCZ	0.150**	0.150**	0.142**	0.141**
RK	0.643**	0.726***	0.840***	0.543**
HY	-0.0510*	-0.0591**	-0.0534**	-0.0514**
KJZC	1.07e-08***	9.27e-09**	5.86e-09	4.64e-09
Observations	210	210	210	210
Number of provinces	30	30	30	30

*** p<0.01, ** p<0.05, * p<0.1

Table 4-8 The Indirect effects results

VARIABLES	(Model 1)	(Model 2)	(Model 3)	(Model 4)
WGTFP	14.39	24.64	34.46	33.36
DIF	0.0861			
WDIF	-0.105**			
DIFRDI	-7.33e-07			
WDIFRDI	0.00280**			
COVER		0.0473		
WCOVER		-0.123**		
COVERRDI		-2.73e-06		
WCOVERRDI		0.00302**		
USAGE			0.00924	
WUSAGE			-0.129***	
USAGERDI			-2.61e-06	
WUSAGERDI			0.00325**	
DIGIT				-0.171***
WDIGIT				-0.0239
DIGITRDI				3.63e-06
WDIGITRDI				0.000413
ZCZ	-0.00914	0.0553	-0.0531	-0.127
RK	0.191	0.123	0.578	0.379
HY	-0.123	-0.122	-0.115	-0.122
KJZC	-3.40e-09	-1.19e-10	1.39e-09	-3.42e-09
Observations	210	210	210	210
Number of provinces	30	30	30	30

*** p<0.01, ** p<0.05, * p<0.1

Table 4-9 The total effects results

VARIABLES	(Model 1)	(Model 2)	(Model 3)	(Model 4)
WGTFP	-12.48	-0.0829	5.002	13.30
DIF	0.159			
WDIF	-0.143***			
DIFRDI	-1.08e-05*			
WDIFRDI	0.00415***			
COVER		0.0617		
WCOVER		-0.163***		
COVERRDI		-1.17e-05*		
WCOVERRDI		0.00442***		
USAGE			0.0407	
WUSAGE			-0.164***	
USAGERDI			-9.85e-06**	
WUSAGERDI			0.00456***	
DIGIT				-0.163**
WDIGIT				-0.0512*
DIGITRDI				-7.14e-07
WDIGITRDI				0.00134
ZCZ	0.141	0.205	0.0884	0.0139
RK	0.834	0.850	1.418	0.922
HY	-0.174**	-0.181**	-0.169**	-0.174**
KJZC	7.28e-09	9.15e-09	7.25e-09	1.22e-09
Observations	210	210	210	210
Number of provinces	30	30	30	30

*** p<0.01, ** p<0.05, * p<0.1

4.3 Expanded analysis by region

Due to the large geographic span of China and the large differences in regional development levels, the following will conduct a further expansive analysis based on the geographic locations of China's provinces. With reference to the “Chinese Provinces: East, West, Central, and Northeast Region Division Method” published by the National Bureau of Statistics of China in 2011, we divided China's provinces into developed regions in the east and underdeveloped regions in the west. Then for the two different types of regions, a classification regression is performed to study the influence of DIF and RDI on the development of GTFP in the two types of regions.

Figure 4-10 shows the results of the spatial Doberman model 1 in different regions. In the two different groups of western provinces and eastern provinces, the independent variables we mainly study are still significant, indicating that they still have a significant impact on the development of GTFP. The reasons why these factors will significantly affect the development of GTFP have been explained above, and they will not be repeated here. The control variables ZCZ, RK, HY and KJZC are basically insignificant. This shows that it is difficult for these variables to have a significant impact on the development of GTFP after distinguishing the geographical locations of provinces. The reason may be that the economic development levels of the provinces in the two groupings are relatively close, and the numerical difference of the variables is small. Therefore, they will not have a significant impact.

Table 4-10 The results of spatial Dubin Model by Region

VARIABLES	The whole China	Western provinces	Eastern provinces
WGTFP	-26.70**	-0.0390**	-15.30***
DIF	0.0780**	0.135	0.125**
WDIF	-0.0439***	0.0580**	0.0750**
DIFRDI	-1.02e-05***	0.226**	1.02e-05***
WDIFRDI	0.00150***	-0.00149*	-0.000938**
ZCZ	0.150**	-3.05e-06	-1.29e-06
RK	0.650**	7.17e-05	0.000170
HY	-0.0554*	-9.87e-08	-3.36e-07
KJZC	1.05e-08***	-1.74e-09	-1.43e-08*
Log-likelihood	407.7566	250.9748	195.0373
Spatial effect	Yes	Yes	Yes
Time effect	Yes	Yes	Yes
Observations	210	119	91

*** p<0.01, ** p<0.05, * p<0.1

We decompose the spatial effect of such regression results. According to Figure 4-11, we find that the decomposition results of the latter two models are basically similar to the decomposition results of the original model. We can get the following conclusion: most variables have significant direct effects on the development of GTFP, and only some variables have indirect effects on the development of GTFP.

Table 4-11 Decomposition of spatial effects by region

Effects	Variables	The whole China	Western provinces	Eastern provinces
Direct effects	WGTFP	-26.87**	-0.0323*	-15.52***
	DIF	0.0731**	0.115	0.123**
	WDIF	-0.0384***	0.0693**	0.0806**
	DIFRDI	-1.00e-05***	0.202**	1.09e-05**
	WDIFRDI	0.00136***	-0.00188**	-0.00102**
	ZCZ	0.150**	-3.42e-06	-1.25e-06
	RK	0.643**	0.000152	0.000179
	HY	-0.0510*	-8.17e-08	-3.70e-07
	KJZC	1.07e-08***	-1.88e-10	-1.43e-08**
Indirect effects	WGTFP	14.39	-0.0621	-26.35
	DIF	0.0861	0.153	0.114
	WDIF	-0.105**	-0.0918	0.220**
	DIFRDI	-7.33e-07	0.209	2.59e-05**
	WDIFRDI	0.00280**	0.00318	-0.00246**
	ZCZ	-0.00914	7.58e-06	-1.07e-06
	RK	0.191	-0.000866*	0.000556
	HY	-0.123	-2.24e-07	-1.76e-06*
	KJZC	-3.40e-09	-1.68e-08	4.35e-09
Total effects	WGTFP	-12.48	-0.0944**	-41.87**
	DIF	0.159	0.268	0.238
	WDIF	-0.143***	-0.0225	0.301**
	DIFRDI	-1.08e-05*	0.410	3.68e-05**
	WDIFRDI	0.00415***	0.00130	-0.00348**
	ZCZ	0.141	4.15e-06	-2.33e-06
	RK	0.834	-0.000714*	0.000735
	HY	-0.174**	-3.05e-07	-2.13e-06*
	KJZC	7.28e-09	-1.70e-08	-9.99e-09
	Observations	210	119	91
	Number of provinces	30	17	13

*** p<0.01, ** p<0.05, * p<0.1

4.4 Robustness test to solve endogenous problems

In the introduction, we have discussed that there may be endogenous problems caused by reverse causality between RDI and GTFP. Therefore, in response to this problem, the endogenous test will help us confirm the robustness of the above-mentioned empirical results. In this article, we choose 'the number of people with high school education and above in the previous year' (EDU) as the instrumental variable of the RDI, and we use the two-stage least squares estimation to correct the endogenous problem of the model. Of course, in the calculation process, we will take the logarithm of EDU

EDU can effectively reflect the level of local labor resources and is positively correlated with the regional development index. In addition, after fixing other control variables, there is no direct correlation between the it and GTFP, which meets the exogenous requirements of instrumental variables. This will make it possible to become an effective instrumental variable for the regional development index. The instrumental variable unrecognizable test shows that the Kleibergen-Paap rk LM statistic is 3.103 (p = 0.0782), rejecting the unrecognizable null hypothesis. Weak instrumental variable Cragg Donald Wald- 1.500, less than the 10% critical value (7.03), unable to reject the hypothesis of "endogenous variables and instrumental variables have a strong correlation". In addition, since the number of endogenous variables and instrumental variables is the same, instrumental variables are just identified, and there is no over-identification

According to Table 4-12, considering the endogenous problems of RDI and GTFP, DIFRDI still has a significant negative impact on green total factor productivity, which is the same as the result above. This shows that the above regression analysis results are robust and consistent

Table 4-12 Analysis of Endogenous Problems

	The first stage	The second stage
	Regional Development Index	Green Total Factor Productivity
EDU	0.617303**	
DIFRDI		-.000011*
WDIFRDI		.0001279
Other control variables	Yes	Yes
Spatial and Time effect	Yes	Yes
observations	210	210

*** p<0.01, ** p<0.05, * p<0.1

5 Conclusion

The writer uses panel data of 30 provinces in China from 2011 to 2017 to do research to get the development level of DIF, RDI and the growth rate of GTFP of each province in China. Then from the perspective of space, using Morans'I and the spatial Dubin model to analyze the spatial distribution characteristics of DIF and RDI, and their influence on the development of GTFP.

The research found the following phenomena: 1. Both DIF and RDI have significant spatial positive correlation. It shows that the regions with high DIF development levels and regions with high DIF development levels are geographically close, and they cluster together. The regions with high regional development index and the regions with high regional development index are geographically close, and they are clustered together. The regions with low DIF development levels and low regional development index are in the same way. 2.DIF improves the growth of GTFP, by upgrading the industry structures, promoting the enterprise to produce environmentally friendly products, and changing consumers' behavior. 3. Higher RDI reduces the GTFP's increase caused by DIF. These phenomena are consistent with our hypotheses, indicating that the hypotheses are correct.

Besides that, we also got some additional conclusions. 1. Because of the existence of the economic siphon effect, well-developed areas will significantly hinder the development of GTFP in the surrounding areas. 2. The three sub-indices of DIF, COVER, USAGE and DIGIT, will all promote GTFP, but only USAGE has a significant impact. This is because COVER and DIGIT in the evaluation system affect consumers more than enterprises. However, companies have a greater impact on GTFP than consumers, so it is difficult for COVER and DIGIT to have a significant impact on GTFP.

Based on the above conclusions, we believe that regional governments should focus on the development of DIF in underdeveloped areas and reduce regional development differences. Two reasons for paying attention to DIF in underdeveloped areas are as follows: 1. DIF can promote the development of GTFP in the area; 2. Compared with DIF in developed areas (higher RDI levels), DIF in underdeveloped areas is less hindered by RDI, so it can promote the development of GTFP more strongly. At the same time, we should focus on the depth of DIF usage when developing DIF. Because the deepening of the use of DIF can effectively increase the help enterprises and consumers get from DIF, thereby promoting the development of GTFP. Reducing the difference in regional development can effectively reduce the siphon effect between regions, thereby avoiding the excessive concentration of capital, labor and energy in regional cities. Excessive concentration of resources in large cities will make surrounding small cities lack of development momentum, become more decayed and backward, and it is not conducive to coordinate the regional development. Ways such as fiscal transfer payments and strengthening the construction of digital infrastructure can be adopted to alleviate the regional development differences caused by the uneven development of DIF.

Although this article has made some research results on the influence of DIF and RDI on the development of GTFP, there are still some limitations.

The data used in this article is from January 1, 2011 to December 31, 2017. Since data from a few years ago is used instead of the most recent data, the timeliness of the data is not particularly good. The reason is that there is a certain degree of difficulty in data acquisition and data processing, and it is difficult to complete this work alone due to the academic ability of the author. Therefore, the research results of other scholars and published research conclusions are cited as a substitute for the data we need. At the same time, due to the data calculated by others, there may be some problems in the accuracy of the data.

Although there are many factors that affect the development of GTFP, when setting up the model, we only selected a limited number of variables as the main independent variables and control variables. This is because many economic indicators have collinearity. If too many variables are included, it may not be conducive to the analysis of the model. However, choosing only a few dependent variables to analyze the development of GTFP, it is always difficult to analyze the various factors that affect GTFP perfectly.

Although the expectation of this article is to discuss the spatial impact of DIF and RDI on the development of GTFP, we found that their spatial spillover effects on the development of GTFP are not significant. It shows that in this field, Internet technology can gradually overcome geographic differences, so the spatial impact is no longer significant. Therefore, discussing online factors from a spatial perspective may not be the best choice.

This article has the above-mentioned limitations—there may be some other limitations, which are not described in detail here—but overall, it has innovatively established the relationship between Digital Inclusive finance, Regional Development Index and Green Total Factor Productivity. And we preliminary analyze the mechanism of influence between them. Therefore, we still believe that it has made some contributions to the development of theory.

6 Reference List

1) Hong L, Shi J F. Energy efficiency analysis on Chinese industrial sectors: an improved Super-SBM model with undesirable outputs[J]. Journal of Cleaner Production, 2014, 65(FEB.15):97-107.
2) Tone K. A slacks-based measure of efficiency in data envelopment analysis[J]. European Journal of Operational Research, 2001, 130(3):498-509.
3) Y. H, Chung, et al. Productivity and Undesirable Outputs: A Directional Distance Function Approach[J]. Journal of Environmental Management, 1997.
4) Paas T, Schlitte F. Regional Income Inequality and Convergence Processes in the EU-25[C]// ERSA conference papers. European Regional Science Association, 2006.
5) Paelinck J. Spatial Econometrics: History, State-of-the-art and Challenges Ahead. workshop on spatial econometrics, 2005.
6) Tobler W R. A Computer Movie Simulating Urban Growth in the Detroit Region[J]. Economic Geography, 1970, 46(2).
7) Anselin L. Spatial Econometrics: Methods and Models[M]. Springer Netherlands, 1988.
8) [10] Luc, Anselin, and, et al. Local Geographic Spillovers between University Research and High Technology Innovations[J]. Journal of Urban Economics, 1997.
9) Moran, Ap P. NOTES ON CONTIRDIOUS STOCHASTIC PHENOMENA[J]. Biometrika, 1950, 37(1/2):17-23.
10) Anselin, L. (1995), Local Indicators of Spatial Association—LISA. Geographical Analysis, 27: 93-115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x
11) MARSHALL A.Principles of economics[M].8th edition.London:Macmillan,1920:271-501.
12) Krugman P. Increasing Returns and Economic Geography[J]. Journal of Political Economy, 1991, 99(3):483-499.
13) Hall B H,Lerner J.The financing of R&D and innovation[J]Handbook of the Economics of Innovation,2010,1:609-639.
14) Agarwal S, Hauswald R. Distance and private information in lending[J] Review of Financial Studies,2010(7):2757~2788.

15) GOLDSMITH R. Financial Structure and Development[M]. New Haven:Yale University Press,1969:113.

16) 易行健,周利.数字普惠金融发展是否显著影响了居民消费——来自中国家庭的微观证据[J].金融研究,2018(11):47-67. Xingjian Yi, Li Zhou.Does the development of digital financial inclusion significantly affect household consumption——micro evidence from Chinese households[J].Financial Research,2018(11):47-67.

17) 余奕杉，卫平．中国城市绿色全要素生产率测度研究[J]. 生态经济,2021,37(03):43-52.Yishan Yu, Ping Wei. Research on Green Total Factor Productivity Measurement of Chinese Cities[J]. Ecological Economy, 2021, 37(03): 43-52.

18) 邱斌,杨帅,辛培江.FDI 技术溢出渠道与中国制造业生产率增长研究:基于面板数据的分析[J].世界经济,2008(08):20-31. Bin Qiu, Shuai Yang, Peijiang Xin. Research on FDI Technology Spillover Channels and China's Manufacturing Productivity Growth: An Analysis Based on Panel Data[J]. World Economy, 2008(08): 20-31.

19) 王永静,李慧.数字普惠金融、新型城镇化与城乡收入差距[J].统计与决策,2021,37(06):157-161. Yongjing Wang, Hui Li. Digital financial inclusion, new urbanization and urban-rural income gap[J]. Statistics and Decision, 2021, 37(06): 157-161.

20) 赵娜. 新型城镇化发展质量的测度与评价[J]. 统计与决策, 2020, v.36;No.562(22):59-62. Na Zhao. Measurement and evaluation of the development quality of new urbanization[J]. Statistics and Decision, 2020, v.36; No.562(22):59-62.

21) 郭峰、王靖一、王芳、孔涛、张勋、程志云,《测度中国数字普惠金融发展：指数编制与空间特征》,《经济学季刊》,2020 年第 19 卷第 4 期,第 1401-1418 页。GUO Feng, WANG Jingyi, WANG Fang, KONG Tao, ZHANG Xun, CHENG Zhiyun, "Measuring China's Digital Financial Inclusion: Index Compilation and Spatial Characteristics", China Economic Quarterly, 2020,19(4),1401-1418

22) 赵晓鸽,钟世虎,郭晓欣.数字普惠金融发展、金融错配缓解与企业创新[J].科研管理,2021,42(04):158-169.Xiaoge Zhao, Shihu Zhon, Xiaoxin Guo. Digital financial inclusion development, financial mismatch mitigation and enterprise innovation[J].

Science Research Management,2021,42(04):158-169.

23) 万佳彧,周勤,肖义.数字金融、融资约束与企业创新[J].经济评论,2020(01):71-83. Jiayu Wan,Qin Zhou,Yi Xiao.Digital Finance, Financing Constraints and Enterprise Innovation[J].Economic Review,2020(01):71-83.

7 Appendix

Table 7-1 Variable descriptive statistics

Variable		Mean	STD.DEV.	Min	Max	
GTFP	overall	0.927857	0.049376	0.69281	1.106287	N=210
	between		0.02475	0.88443	0.992135	n=30
	within		0.04293	0.72246	1.102735	T=7
DIF	overall	172.0594	77.74163	18.33	336.6506	N=210
	between		22.06974	139.925	227.0877	n=30
	within		74.63693	25.16165	281.6222	T=7
COVER	overall	151.4972	74.54297	1.96	316.1184	N=210
	between		26.74066	118.4994	223.0039	n=30
	within		69.7289	26.0233	257.5722	T=7
USAGE	overall	168.5765	78.51732	6.76	396.0493	N=210
	between		32.119	120.2174	246.0006	n=30
	within		71.85371	8.815824	318.6251	T=7
DIGIT	overall	246.2881	114.4881	7.58	453.66	N=210
	between		10.1412	226.6012	277.0247	n=30
	within		114.051	9.613407	447.7788	T=7
DIFNU	overall	6530.952	8025.566	119.3283	50321.89	N=210
	between		6709.617	1256.07	27525.72	n=30
	within		4547.875	-13593.1	29327.13	T=7
COVERNU	overall	5923.068	7554.807	12.7596	46880.45	N=210
	between		6350.954	1074.398	25599.49	n=30
	within		4230.652	-12921.4	27204.02	T=7
USAGENU	overall	6594.294	8359.276	44.0076	55758.7	N=210
	between		7013.468	1102.094	27642.82	n=30
	within		4701.158	-12422.8	34710.17	T=7
DIGITNU	overall	8423.266	9471.004	268.9296	52039.04	N=210
	between		7445.301	2135.743	33674.09	n=30
	within		5988.219	-17936.5	26788.22	T=7
ZCZ	overall	9.730327	0.838215	7.420842	11.40428	N=210
	between		0.833594	7.697857	11.12876	n=30
	within		0.166356	9.270338	10.13526	T=7
RK	overall	8.19831	0.735897	6.342421	9.320897	N=210
	between		0.746491	6.368281	9.285559	n=30
	within		0.017105	8.105253	8.253441	T=7
HY	overall	11.56628	0.815465	9.440328	12.98149	N=210
	between		0.817059	9.606798	12.78454	n=30
	within		0.128695	11.08428	11.86382	T=7
KJZC	overall	15.76821	0.656778	13.91888	17.3417	N=210
	between		0.63379	14.16376	16.93098	n=30
	within		0.202974	15.23214	16.17894	T=7