

## Application of Data Mining in Fiscal Expenditure on the New Urbanization and Income Gap based on Network Big Data

Yan Liu

College of Urban, Rural Planning and Architectural Engineering, Shangluo University, Shangluo 726000, Shaanxi, China  
\*liuyan\_ly2004@126.com

### Abstract

*Network big data is an important driving force for the upgrading of information technology, and network data has brought great opportunities to the economic survey. The openness and freedom of the network also produced the possibility of private information and data to be destroyed or violated, and the security of the Internet is becoming more and more important. In this paper, the author analyzes the impact of fiscal expenditure on the new urbanization and income gap by using data mining technology. The result shows that the public expenditure increase 1% will lead urbanization rate increase 0.0462%; also, the effect of public finance expenditure on urbanization has a certain time lag, and the effect will be more obvious in the long term. At the same time, public expenditure will help to narrow the income gap in the long term, public expenditure at lag 2 increase 1% will lead income gap decrease 0.758%. So that, increasing public expenditure is an important way to promote the equalization of basic public services between urban and rural areas.*

**Keywords:** *Network security, big data, data mining, fiscal expenditure, new urbanization, empirical analysis*

### 1. Introduction

In recent years, the development of urbanization in China has made remarkable achievements, has become an important pillar of economic and social development. In 2014, China's urbanization rate reached 54.77%, compared with 2000, the urbanization rate increased by 18%, with an average annual growth of 1.3%. With the rapid development of urbanization, China's economic growth and resident income has increased considerably, 2014 China's gross domestic product (GDP) reached 63.6463 trillion yuan, the growth rate of 7.3%, urban residents per capita disposable income of 2.9 million yuan, rural residents 1.05 million[1-2]. It can be seen that China's urbanization process with more than ten years of process, to achieve a number of developed countries in the last hundred years to achieve the goal. However, in the process of urbanization, there are a variety of problems [3]. First, fiscal expenditure for the influence of urbanization, because the reform of the tax system, resulting in the imbalance of the central and local fiscal powers and financial authority, lack of local financial sector income, and local urbanization need substantial financial support. In this case, local governments often use the means of land finance, increase fiscal revenue, in order to promote the local infrastructure [4]. Since 2008, the number and scale of local government financing platform has rapidly increased, which leads to the increase of local financial risk. Secondly, in the process of urbanization in our country, no good treatment of urban and rural dual structure, the process of urbanization though driven by the development of most of the town, but did not effectively solve the problem of the income gap between urban and rural areas of our country. The income gap between urban and rural areas by 2.7 times increase to 2.9 times in 2014, can be seen, the development of urbanization did not

reduce the income gap between urban and rural areas related to the this phenomenon the production, but also with the historical reasons in our country leads to the dual economic structure.

Starting in 1980s, with the reform and opening up, China has entered a period of economic transformation; at the same time, China's urbanization has reached an unprecedented height[5-6]. From 1978 to 2010, China's urban population increased from the city to the city, the proportion of the total population of the city rose from 50% to 18%. While China is accelerating the process of urbanization, but the urbanization problem has emerged. The main problem is that on the one hand, a large number of rural populations is calculated as urban residents, but they did not get the rights equal to urban residents; on the other hand, urbanization is extremely uneven development, different cities and between urban and rural gap increases day by day[7]. Aiming at these problems, the government proposed the new towns of the concept, and stressed the importance of future development of urbanization to quality oriented, rather than before, so pay attention to speed, in which a key theme is to achieve urban and rural planning and sustainable development. The development of new urbanization is an important pillar of China's modernization process, and it is also the main way to promote social and economic development[8]. We should clearly recognize that fiscal expenditure as an important source of funding to promote local urbanization, is to provide urban infrastructure, to achieve the equalization of basic public services an important way. In the construction of new urbanization process, how to avoid defects of previous, strengthen the financial expenditure of the functions of the allocation of resources, optimize the structure of fiscal expenditure, according to the needs of China's national economy, fiscal expenditure more bias related to people's livelihood education, social security, public health and other fields. Improve the efficiency of the use of financial funds, through financial subsidies, tax incentives and other means, to promote private investment[9-11]. In this paper, we construct the empirical model of fiscal expenditure and urbanization, income gap between urban and rural residents, and analyze the important support for the development of China's urbanization.

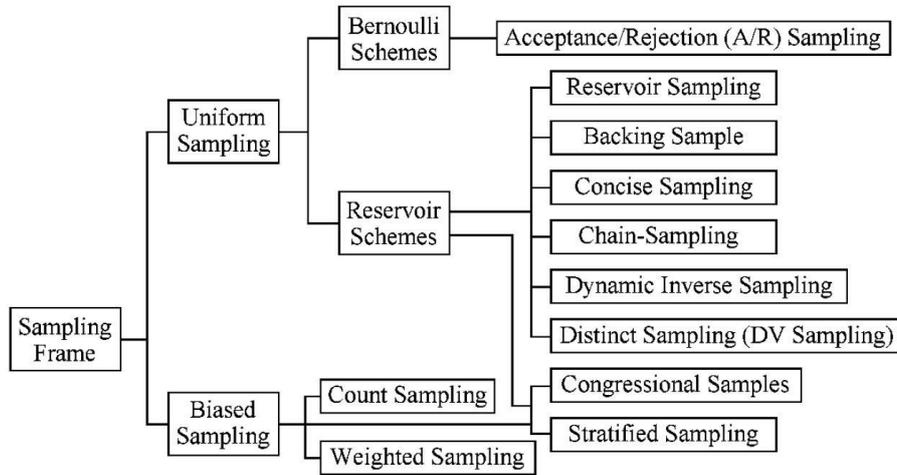
## **2. Data Mining and Support Vector Machine**

### **2.1. Data Mining**

Data mining refers to a large amount of information; there is no law of information, to dig out the value of the people of the knowledge or the law. Since the 80s of the last century until now, after 30 years of development, the development of data mining in various fields have made many achievements, especially play a huge role in the financial industry, medical treatment and electronic commerce industry, therefore also contributed to the calculation of the development of computer science and technology. With the increasing influence of data mining in various fields, a lot of powerful data mining and analysis tools are gradually emerging. Through the use of data mining and analysis tools, combined with the actual demand model, can help users to extract the effective information, so as to analyze the trend of the data, so that the value of the data generated. At present, the research of data mining as long as the research and improvement of the algorithm and the specific application of data mining, such as clustering, classification, association rules, recommend prediction algorithm. The algorithm directly determines the effectiveness and efficiency of data mining and analysis, so the research of data mining is mainly reflected in the algorithm.

After more than 20 years of research and improvement, the traditional data mining algorithm has become more and more mature, more efficient and more accurate, and can solve many practical problems. However people are aware of, a lot of data mining algorithm in the face of the small scale of data will get better effect, when the face of the

large scale data, obviously cannot meet people's needs, not only in execution efficiency, or in the result accuracy are unable to meet. Therefore, many universities and research organizations at home and abroad, and even the research institutes and so on are spending a lot of money and material resources to study the data mining algorithms on large scale data, how to achieve high performance issues.



**Figure 1. Classification of Representative Sampling Methods on Data Ming**

## 2.2. Support Vector Machine

Support vector machine is a new machine learning method based on statistical learning theory, and it is a statistical machine learning theory developed in the case of small samples. Based on the sample machine learning, in fact, it is a link between the input and output through the training samples. In order to find an optimal function, a requirement follows the minimization of the expected risk:

$$R(\pi) = \int L(y, f(x, \pi)) dF(x, y)$$

Therefore, the expected risk of  $R(\pi)$  to replace the experience of risk:

$$R_{emp}(\pi) = \frac{1}{l} \sum_{i=1}^l L(y_i, f(x_i, \pi))$$

Vapnik after in-depth study put forward the principle of structural risk minimization (SRM), that is:

$$R(\pi) \leq R_{emp}(\pi) + \sqrt{\frac{h \ln\left(\frac{2l}{h}\right) - \ln\left(\frac{\eta}{4}\right)}{l}}$$

In the linear separable model, classification of SVM in a  $H$  inner space structure for super plane:

$$\langle w, x \rangle + b = 0, \quad w \in H, b \in R$$

For linearly separable sets of samples, the total can be adjusted by  $w$  and  $b$ :

$$\langle w, x \rangle + b \geq 1, y = 1$$

$$\langle w, x \rangle + b \leq -1, y = -1$$

Obviously, this adjustment does not change the discriminant function (5). Define two standard canonical hyper planes as:

$$\langle w, x \rangle + b = 1 \quad \text{and} \quad \langle w, x \rangle + b = -1$$

The sample points on the two standard hyper planes are called support vectors, and the name of the support vector machine is obtained. Define the classification interval (margin) as:

$$\rho = 2r = \frac{2}{\|w\|}$$

As a result, the optimization problem for linearly separable SVM is:

$$\min J(w, b) = \frac{1}{2} \|w\|^2$$

The above problem is the main problem of SVM, which is a convex two optimization problem. The problem can be solved by constructing the Lagrange multiplier method to solve the problem.

$$\begin{aligned} \max W(\alpha) &= \sum_{j=1}^l \alpha_j - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j x_i x_j \\ \sum_{j=1}^l \alpha_j y_j &= 0 \end{aligned}$$

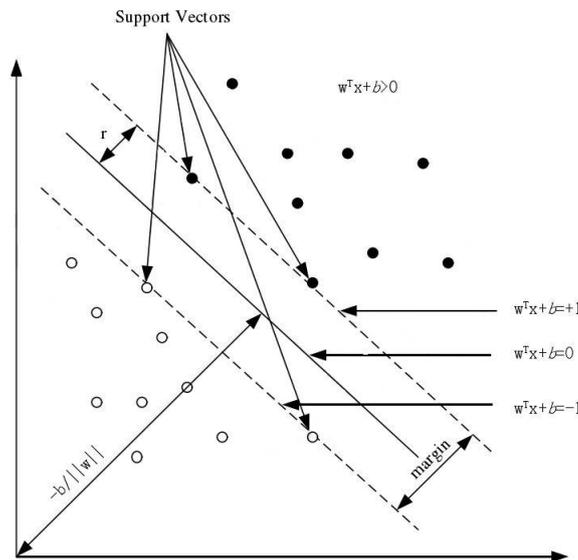


Figure 2.SVM Classification

### 2.3. Network Data

In recent years, with the rapid development of the Internet, networking, cloud computing, network convergence and IT communication technology, the rapid growth of data has become a serious challenge faced by many industries and precious opportunities, so the information society has entered the era of big data. The emergence of big data is not only changing the mode of operation of enterprises and the people's life and work, and

even caused a fundamental change in the pattern of scientific research. Generally speaking, big data is not in a certain period of time by conventional machine hardware and software tools on the perception, acquisition, management, processing and service of the network data set. Big data refers to the people, machine, three dimension in world space network interact and fusion produced and available on the Internet at large the data, referred to as the network data. Network data is an important driving force to upgrade the core information industry of the national economy. The integration of human, machine, material as three dimension of the world has produced large-scale data, how to use the network measurement and perception, big data has become a common problem facing many industries in the national economy, and become the digital the information barriers in these industries and barriers. How different sectors can break through the barriers, the key is to solve the common problems of basic network data. For example, for a unified representation of unstructured data and analysis, the current lack of effective methods and tools. Therefore, through the analysis and Research on the common problems of big data network so, enterprises can grasp the processing capacity of large data network or network can withstand large data processing cost and the cost, and make the whole the industry has entered a new stage of digital and information technology. In this sense, to solve the common problems on the basis of large data networks will be the new focus of application integration of the new generation of information technology, information industry is a new engine of sustained rapid growth, new power is to promote the competition ability of industry users. Big data by the academic circles again review on the methodology of scientific research; scientific research is causing a revolution of thinking and methods of scientific research. Due to the many problems, the theoretical analysis method is so complicated that it is difficult to solve the problem, people began to seek the simulation method this has created a scientific calculation. And the emergence of big data has spawned a new pattern of scientific research, in the face of big data, researchers only directly from the data to find, analyze or dig the needed information, knowledge and wisdom, and even without direct contact with the object to be studied.

### **3. New Urbanization**

#### **3.1. Land Urbanization**

New urbanization is the only way to realize the economic and social development in china. Theoretically speaking, the expansion of the scale of the town to produce the effect of gathering and diffusion effect, optimize the allocation of resources, to improve the efficiency of economic development; cities and towns to promote intensive can diluted public infrastructure construction costs, to provide high quality services to the public. The town of specialized division of labor can create more jobs, and to provide consumers with more choices. From a practical point of view, the new urbanization is also an important driving force for China to achieve industrial transformation and upgrading, to solve three rural issues, promote the coordinated development of the region, and promote social progress. Moreover, the problems in the process of urbanization, it is necessary to build a new urbanization, in the development process has been gradually resolved. Urbanization is a complex and dynamic process of development, as well as the concept of new urbanization from the whole process to the final maturity, but also in the continuous improvement of the theory and practice of continuous inspection. In the process of China's urbanization, the use of land resources has been maintained extensive low efficiency of the state. According to estimates, from 2000 to 2011, urban population increased about 50% and towns built area but was more than the original expanded by more than 70%; rural population decrease in the number of obvious, but rural residential land were significantly increased. And in the past 10 years, 90% of the urban construction land demand is through expropriation of agricultural land, new construction area to meet

the; and the town without the development of the stock of land supply, accounted for only 10%. Urban expansion not only resulted in the fragmentation of the surrounding urban land use, the formation of waste of land resources; and on cultivated land protection and food security also brought certain influence. At the same time, it also contributes to the low efficiency development of the urban interior space utilization. Compared with New York, Seoul, Hong Kong and other international cities, the volume of China's first tier cities is generally low; Beijing volume rate of only 2.8, Shanghai is 2.5, far lower than other international cities. In other words, by raising the population density in urban areas, to further promote the development of urbanization, there is still a lot of space.

### **3.2. The Connotation of New Urbanization**

People of the town is the core of the new urbanization, which includes not only the traditional concepts of agricultural population change for the urban population of the "population urbanization", more important is to ensure that agricultural population migration and the original town residents, in education, employment, medical care, housing, social security and other aspects of enjoy the same level of basic public services, truly realize the citizenization of the population migration, enhance the urbanization development process, inclusive. Early reflect the urbanization development pattern in order to achieve the conversion by industrialization. In the town of the early, this process can promote the whole society to improve productivity and economic development, but our country now in the middle stage of urbanization development, more attention should be paid to the quality of the development of urbanization, pay attention to the citizens in the new town development. At present, a huge gap of our country resident population of the town of rate and household population urbanization rate not only hinder the agricultural population migration flows to urban, expand the urban internal differentiation; lead to movements of people's income and consumption level is low, unable to form an effective internal demand; but also to bring the shortage of human capital investment in enterprises and the workers themselves not only to the enterprise is lack of dynamic convection of workers to invest, lack of stability public expectations of workers to enhance their human capital investment. Therefore, the new urbanization must pay attention to the development needs of the people, solve the urbanization support system for economic development, to form a reasonable and effective employment space; to improve the town's absorbing capacity and public service ability to provide, to remove obstacles to labor mobility, to ensure that agricultural population migration enjoy the same treatment as urban residents basic public services, to achieve the integration of agricultural population migration of people and society, fostering new social strata and the consumption of the main, expanding domestic demand, agricultural population migration can really from the urbanization process of benefit, implementation of national new town planning.

Therefore, in the process of urbanization development, we must adhere to the coordinated development of urbanization and industrialization, informatization and agricultural modernization. Although the new urbanization will lead to the gradual transfer of agricultural population to the town, but it does not mean that the gap between urban and rural areas will be more and bigger. Construction of new urbanization is ultimately to eliminate the static two Yuan of urban and rural areas, urban and rural development, urban and rural residents to share the results of the development of urbanization. In development planning and design, infrastructure investment, the provision of public services, and construction of new towns will focus on the promotion of urban and rural elements of equal exchange and balanced allocation of public resources, the industrial sector driven agricultural sector and in urban driving rural and towns and villages will be included in the integrated development of the track, the formation model of workers and peasants, the relationship between urban and rural areas.

In addition, considering China's huge population base, the agricultural population in a short period of time to absorb the entire town is not realistic.

## 4. Empirical Analysis

### 4.1. ADF Stationarity Test

In order to verify the impact of China's public financial expenditure on the urbanization and income gap between residents, we construct the VAR model of China's fiscal expenditure, urbanization and income gap between residents and residents. Data used in 2001 to 2014 China's public finance expenditure public finance expenditure (PFE), rate of urbanization (UR) and income gap between urban and rural residents(IG), data from China Statistical Yearbook 2014. At the same time, in order to avoid the influence of different variance, we have carried on the logarithm processing to the data, recorded as LnPFE, LnUR and LnIG. Firstly, we analyze the stability of the data. Data stability is the basis for empirical analysis, we use ADF unit root test for analysis, the results are shown in table 1. As can be seen from the table, LnPFE, LnUR and LnIG are not smooth, so we have to deal with the data, recorded as d.LnPFE, d.LnUR and d.LnIG. By testing the difference data, we can see that the variables are stable, so we can carry on the further analysis.

**Table 1. ADF Test Results**

variable	ADF statistics	1%critical value	5%critical value	10%critical value	conclusion
LnPFE	-0.657	-3.750	-3.000	-2.630	nonstable
LnUR	-0.840	-3.750	-3.000	-2.630	nonstable
LnIG	-0.721	-3.750	-3.000	-2.630	nonstable
D. LnPFE	-3.579	-3.750	-3.000	-2.630	stable
D. LnUR	-4.247	-3.750	-3.000	-2.630	stable
D. LnIG	-4.450	-3.750	-3.000	-2.630	stable

### 4.2. VAR Model

First, we for fiscal spending and urbanization rate between the data to test, through the use of Stata.12 statistical software testing, we can draw AIC and SBIC in 2 order to min, respectively, and -11.0524 and -10.6288, so we selection phase 2 for VAR model for the optimal lag behind. On this basis, we also between the fiscal expenditure and income gap between urban and rural areas of inspection, results show that in the factor of income gap, AIC and SBIC in 2 order to min, respectively, and -9.72472 and -9.3011, the optimal lag periods also for 2 order lag. By constructing the VAR model of public finance expenditure and urbanization, as shown in Table 3, we can see from the table that AIC is -9.318718, HQIC is -9.468326, SBIC is -8.914629, the model is more reasonable.

At the same time, through the coefficient table can be seen, lag phase of public finance expenditure expansion of 1% can promote the urbanization rate increased 0.0375%. Public expenditure that is to promote the important strength to support the process of urbanization in China; at the same time, the lag phase of public finance expenditure increase 1%, can promote the urbanization rate increased 0.0876% that the influence of public finance expenditure on urbanization has a certain delay, fiscal expenditure in the long term of town of rate enhancement effect is more obvious.

**Table 2. Model Order Selection**

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	19.9064				.000096	-3.58129	-3.64768	-3.52077
1	49.4479	59.083	4	0.000	6.0e-07	-8.68958	-8.88874	-8.50802
2	69.2619	14.531*	4	0.006	1.1e-07	-11.0524*	-11.5171*	-10.6288*
3	61.9963	25.097	4	0.000	1.3e-07	-10.3993	-10.7312	-10.0967
4	63.1191	11.197	4	0.024	3.6e-07	-9.82382	-10.2885	-9.40021

At the same time, we construct the VAR model of the public finance expenditure and the income gap between urban and rural areas, as shown in Table 4. From the table we can see that AIC is -10.03201, HQIC is -10.18162, SBIC is -9.627924, and the model is more reasonable.

**Table 3. LnPFE and LnUR Regression Analysis**

Sample: 2003 - 2014				No. of obs =12		
Log likelihood = 65.91231				AIC= -9.318718		
FPE= 3.43e-07				HQIC= -9.468326		
Det(Sigma_ml) =5.81e-08				SBIC= -8.914629		
LnUR	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
LnPFE L1.	.0462343	.1200179	0.31	0.754	-.1976965	.2727651
LnPFE L2.	.0876921	.1130439	0.78	0.438	-.1338699	.3092541
LnUR L1.	.1085249	.2663608	0.41	0.684	-.4135328	.6305826
LnUR L2.	.0739437	.2533894	0.29	0.770	-.4226904	.5705778
_cons	1.807754	.6451868	2.80	0.005	.5432111	3.072297

**Table 4. LnPFE and LnIG Regression Analysis**

Sample: 2003 - 2014				No. of obs = 12		
Log likelihood = 70.19208				AIC= -10.03201		
FPE= 1.68e-07				HQIC= -10.18162		
Det(Sigma_ml) = 2.85e-08				SBIC= -9.627924		
LNIG	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
LnPFE L1.	.6820479	.1757006	3.88	0.000	.3376811	1.026415
LnPFE L2.	-.7580285	.179068	-4.08	0.000	-1.082195	-.3802616
LnIG L1.	.5316073	.3394116	1.57	0.117	-.1336272	1.196842
LnIG L2.	-.0476822	.2263313	-0.21	0.833	-.4912835	.3959191
_cons	1.006625	.3133269	3.21	0.001	.3925156	1.620735

At the same time, through coefficient table can be seen, the lag phase of public finance expenditure growth of 1% will lead to the income gap between urban and rural areas expand 682%, the public finance expenditure in the short term will help expand the income gap between urban and rural areas; meanwhile, the lagging phase II of public finance expenditure growth of 1% to cause the income gap between urban and rural areas reduced 0.7312%, indicating that the public finance expenditure long-term help to narrow the income gap between urban and rural areas, increase the public finance expenditure is

an important way to promote the equalization of basic public services in urban and rural areas.

### 4.3. Granger Causality Test

We test the model by Granger causal analysis, and the results are shown in Table 5. The results show that LnPFE is the cause of the Granger LnUR, but LnUR is not the cause of the Granger LnPFE, which shows that public expenditure is the cause of urbanization rate increase. At the same time, we LnPFE LnIG of test, the results show LnPFE is the Granger cause of the LnIG, but Granger reasons for not LnPFE LnIG. This shows the cause of public finance expenditure is the impact of the income gap between urban and rural residents.

**Table 5. Granger Causality Test**

Equation	Excluded	chi2	df	Prob > chi2	conclusion
LnPFE	LnUR	5.3241	2	0.084	LnUR is not Granger reason of LnPFE
LnUR	LnPFE	5.3241	2	0.003	LnPFE is Granger reason of LnUR
LnPFE	LnIG	32.533	2	0.170	LnIG is not Granger reason of LnPFE
Ln IG	LnPFE	26.766	2	0.036	LnPFE is Granger reason of LnIG

## 5. Conclusions

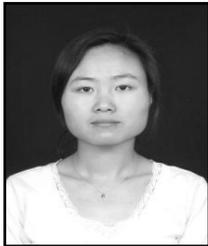
The association between the VAR through the use of econometric model on the development of public expenditure, urbanization and income gap between urban and rural areas were analyzed, the results show: (1) public finance expenditure is an important force in promoting China's urbanization process, the lag of public expenditure a period of expansion of 1% can promote the urbanization rate increased by 0.0375%; (2) effects of fiscal expenditure on urbanization has a certain lag, increase the fiscal expenditure in the long term of the urbanization rate is more obvious; (3) public expenditure will promote the urban-rural income disparity in the short term, but long term public expenditure will help narrow the income gap between urban and rural areas. The two phase of the lag of public expenditure growth of 1% can cause the income gap between urban and rural areas to reduce 0.7312% increase of public expenditure is an important way to promote the equalization of basic public services in urban and rural areas.

In the process of urbanization in our country, increase the expenditure of public finance, reasonable adjustment of the structure of fiscal expenditure, financial funds will be more used in the field of people's livelihood. First, we must correctly handle the relationship between the central and local governments; local finance is the main source of funds to support the construction of local urbanization. Through fiscal and tax reform, strengthen the use of fiscal funds to improve the efficiency of government functions, improve the public service system, and constantly reduce the gap between urban and rural areas. At the same time, to build local main taxes, through the reform of the property tax, to provide more sources of local financial revenue. Secondly, to improve the transfer payment system, strengthen the education, health and social security expenditure, and actively promote the equalization of basic public services. Finally, in the process of urbanization, also want to pay attention to the guarantee of land resources and ecological environment, prevent local governments in order to develop the economy and blindly land urbanization transformation, to the neglect of urban sustainable development. In the process of urbanization, to strengthen the financial supervision of local governments, but also pay attention to the protection of the ecological environment, to build a sustainable development of new urbanization development path.

## References

- [1] L. Bertinelli, "Urbanization and growth", *Journal of Urban Economics*, vol. 56, (2004), pp. 80-96.
- [2] F. Johannes, "The Costs of Urbanization in Developing Countries", *Economic Development and Cultural Change*, vol. 30, (1982), pp. 625-648.
- [3] C. James and J. Davis, "Vernon Henderson. Evidence on the political economy of the urbanization process", *Journal of Urban Economics*, vol. 53, (2003), pp. 98-125.
- [4] S. Paul, "Local government fiscal reform in developing countries: Lessons from Kenya", *World Development*, vol. 21, (1993), pp. 901-923.
- [5] C. Xue and F. Guangjing, "Urbanization construction funds scale and land finance reform: a review of the research on land financial compensation mechanism under the background of new urbanization", *China land science*, vol. 11, (2013), pp. 50-53.
- [6] J. Sun, "promote the development of new urbanization of the financial significance and fiscal policy", *Financial research*, vol. 4, (2013), pp. 14-16.
- [7] S. X. Liu, "Five challenges to the financial system of urbanization in China and Its Countermeasures", *Local financial research*, vol. 4, (2012), pp. 71-73.
- [8] J. Cui and Q. Yang, "An empirical study on the effect of Land Finance on Urbanization in the new century: Empirical Evidence from a second tier city", *Journal of Renmin University of China*, vol. 1, (2014), pp. 12-14.
- [9] P. Li, "Fiscal expenditure on agriculture, farmers' income and the income gap between urban and rural residents: An Empirical Study Based on Provincial Panel Data", *Southern economy*, vol. 8, (2012), pp. 102-104.
- [10] X. Wang, "County level financial development, fiscal expenditure and the difference of income gap between urban and rural residents", *Contemporary economic research*, vol. 9, (2014), pp. 22-25.
- [11] V. Henderson, "The Urbanization Process and Economic Growth: The So-What Question", *Journal of Economic Growth*, vol. 8, (2003), pp. 47-71.

## Author



**Yan Liu**, <1981.12>, <ShaanXi luonan, P. R.China>

Current position,grades:geography lecture, college of urban ,rural planning and Architectural Engineering ,Shangluo University  
Scientific interest:Her research interest field include geographical environment and health.

Publications:more than 10 papers published

Experience:She has teaching experience of 9 years, has completed 3 scientific projects.