Agglomeration and Employment Density: Test Based on Panel Data of Prefecture-Level Cities of China

Gang WU*, Xiuchuan XU

College of Economics and Management, Southwest University, Chongqing 400715, China

Abstract Related factors for measuring urban agglomeration effect were studied firstly. Then, panel data of 283 prefecture level cities of China were collected to analyze the effect of agglomeration on employment density. Besides, fixed effect model was applied to analyze static panel data, and two-step generalized method of moments (GMM) estimator was employed to analyze dynamic panel data. Results reveal that per capita regional GDP, public medical care level, and population mobility have significant effect on employment density. Therefore, there exists effect of agglomeration economy in prefecture level cities of China in the current stage.

Key words Agglomeration effect, Employment density, Dynamic panel data

1 Introduction

In 1978 - 2013, permanent population of urban areas in China rose from 170 million to 730 million, and urbanization rate rose from 17.9% to 53.7%, with annual increase up to 1.02%. The number of cities rose from 193 to 658, and the number of organic towns increased from 2173 to 20113. Rapid acceleration of urbanization absorbs numerous rural labor to transfer employment. It increases allocation efficiency of urban and rural production factors, drives constant and rapid development of national economy, brings about profound changes in social structure, and promotes overall improvement in living conditions of urban and rural residents. Different from many developing countries, in the rapid urbanization process, China has successfully avoided many problems in megacities, such as slums, urban poverty, high crime rate, and social turmoil. However, restrictive policies for urbanization have price. Henderson stated that many cities of China have little population[1]. According to data of the sixth population census in 2010, in 283 prefecture level cities, only 10 cities had 5 - 10 million residents, 21 cities had 3 - 5 million residents, and 103 cities had 1 - 3 million residents. According to Northam Curve, the urbanization rate of China still remains at rapid development of 30% - 70%. By 2030, there will be one billion people living in cities. How to properly guide population flow, orderly promote citizenship of agricultural transfer population, drive coordinated development of big, medium-sized and small cities and small towns, raise sustainable development level of cities, and promote economic transformation and upgrade and social harmony and progress are challenges faced by China. Unprecedented urbanization is a great opportunity. It is feasible to maintain economic growth and promote urbanization through agglomeration economy.

Received: June 29, 2016 Accepted: August 20, 2016

Supported by Key Project of National Social Science Foundation (12&ZD100); Fundamental Research Funds for the Central Universities of Southwest University (SWU1309005).

* Corresponding author. E-mail: wugang630@sohu.com

2 Literature review

2.1 Theory of urban economics Northam believed that the development process of urbanization has stage rules, and he generalized the urbanization process into an S curve[2]. According to Small Town Theory of Fei Xiaotong, developing small towns is of great significance for China's urbanization; numerous rural surplus labor can be digested locally through developing small towns[3]. Scholars holding views of small town theory stated that we should focus on national conditions when studying development rules of urbanization. There are several hundred million rural surplus labors, so it is impossible for the existing 660 cities to provide so many jobs. In some big cities, many problems become increasingly prominent, including air pollution, water shortage, traffic jam, garbage siege, and shrinkage of green land. However, Fan Gang pointed out big cities have scale benefit much higher than small towns[4]. Niu Wenyuan *et al.* compared new urbanization and traditional urbanization and came up with strategic idea of new urbanization development[5].

2.2 Theory of agglomeration economy According to theory of urban economics, agglomeration economy is a reason for existence of cities[6]. Glaeser studied agglomeration economy through employment growth[7], Rosenthal & Strange measured functions of agglomeration economy using number of new enterprises and number of workers in new enterprises[8]. Bai Chongen *et al.* studied factors influencing industrial agglomeration on the basis of statistical data of industrial enterprises in provinces[9]. Through spatial econometric analysis, Jiang Lei and Ji Minhe found China's innovation activities take on regional state, have obvious spatial agglomeration phenomenon, and are mainly distributed in coastal regions with high urbanization level[10]. In the opinion of Liu Xiuyan and Shaojun, the relation between agglomeration and per capita GDP is subject to regional economic development level[10]. The study of Liao Xinlin indicates that in the context of urban and rural dual economic structure and government excessively pursuing economic growth, financial expenditure exerts positive effect on wide-

ning of regional income gap, while urbanization and economic growth exert negative effect on widening of regional income gap[12]. Nie Hualin and Hanyan found that there is significant positive correlation between China's population urbanization level and economic growth[13].

The above literature about agglomeration economy is mainly based on the data before the 1990s and from the perspective of Marshall externality and Jacobs externality, but does not touch on the effect of agglomeration on employment density (the number of employed in one square kilometer). In view of this, we will firstly discuss factors influencing urban agglomeration effect. Then, on the basis of panel data of 2008 – 2012, we will elaborate the effects of all factors on urban employment density through fixed effect estimation and two-step GMM estimation.

3 Model setting and indicator selection

3.1 Employment density model According to Henderson's urban agglomeration and urban economic growth model[14], we use the employment density to express the employment opportunity, namely, the number of employed in one square kilometer within a city. Suppose the employment density of the city i is y_i, the vector related to factors such as city size and agglomeration effect is x_i, then the function of employment density of the city i can be expressed by following equation:

$$y_i = f(x_i, \theta_i, a_i) \tag{1}$$

where $f(\cdot)$ is probability transformation function, θ_i denotes influence parameters of variable x_i, a_i is fixed effect of a specific city and we suppose it does not change with time.

Because related factors of agglomeration economic effect are positive continuous variables, units are different and values are generally high, for the transformation function of equation (1), using Cobb-Douglas function will be more significant than linear function. Other functions, such as CES function and transcendental logarithmic function, are expansion of Cobb-Douglas function. Using these functions, we can obtain more information, but nonlinear estimation or linear estimation may have the problem of complex estimation or multiple linearity, therefore, for simplicity, we treat the function $f(\cdot)$ using Cobb-Douglas function, then the equation (1) can be changed to:

$$y_i = \prod_j x_{ji}^{\theta_i} e^{a_i} \tag{2}$$

Take the logs of both sides of equation (2), considering random influencing factors, add the constant term. Apart from different cities in the cross section, it is necessary to consider time traced panel data, then we obtain the panel data econometric model:

$$\ln y_{it} = \theta_0 + a_i + \sum_j \theta_i \ln x_{j,it} + \varepsilon_{it} \tag{3}$$

The equation (3) only considers the effect of current explanatory variables on explained variables. In fact, the past of explanatory variables may also exert influence on explained variables. Especially in studying the urban agglomeration effect and analyzing the employment density, it is necessary to consider the effect of current year explanatory variables, and also consider accumulated dynamic effects of various factors on the employment density in certain period. Besides, due to statistical data were collected by year, and it is the accounting system, theoretically, the data over the years can be accumulated through comparable price, but it will lose some sample size and it will be impossible to measure dynamic change effect of factors. Therefore, to survey dynamic effects of agglomeration factors, take full use of the information volume provided by panel data in cross section dimension and time dimension, and utilize sample data in a high efficient manner, we make expansion of infinite distributed-lag for equation (3), and obtain:

$$\ln y_{it} = \theta_0 + a_i + \sum_j \theta_i \ln x_{j,it} + \rho \sum_j \theta_i \ln x_{j,it-1} + \ldots + \rho^p \sum_j \theta_i \ln x_{j,it-p} + \varepsilon_{it} \tag{4}$$

where $\rho \in (0,1)$ is time decay factor. When the lag period of explantory variables $p \to \infty$, we can get infinite distributed-lag model. It is easy to prove that the equation (4) is equivalent to one-period lag explanatory variable model.

$$\ln y_{it} = \theta_0 + a_i + \rho \ln y_{it-1} + \sum_j \theta_i \ln x_{j,it} + \varepsilon_{it} \tag{5}$$

Equation (5) is first-order auto-regression dynamic panel data model.

3.2 Data explanation and descriptive statistics According to existing literature, related factor indicators publicly available for measuring urban agglomeration effect mainly include:

(i) Per capita GDP. Lu Ming, Gao Hong and Zuo Tenghong stated that, in current period of China, economic agglomeration not only brings increase in labor productivity and per capita income, but also realizes higher employment rate, thus more laborers can integrate into the process of economic growth and enjoy benefits of urbanization and economic growth[15]. A city with higher agglomeration effect will have higher per capita GDP and per capita income. The agglomeration effect is accompanied with high output and high income, as well as high employment density. In return, the employment density will promote growth of per capita GDP, so the per capita GDP may be endogenous, it is feasible to solve the endogenous problem using GMM estimator. Another approach for solving the endogenous problem is to establish simultaneous equation model. In this study, the explained variable is employment instead of GDP growth, thus, we use single equation GMM estimator, and do not use simultaneous equation model.

(ii) Public service level. Chen Guoliang and Chen Jianjun pointed out that the big gap between eastern and central and western regions in education and medical care especially in higher education directly affects agglomeration degree of secondary and tertiary industries[16]. A city with higher urban public service level will be more capable of attracting talents. In this study, we use urban education expenses and the number of urban hospital beds to reflect urban public service level.

(iii) Population mobility. In the opinions of Ren Xiaohong and Zhang Zongyi, when the urbanization level is relatively low, improving traffic facility can promote flow of production factors and narrow the urban and rural income gap, but with production factors constantly flowing to urban areas, the functions of improving

traffic facilities will be not significant any more; when production factors flowing to cities reach to a certain critical value, the function of production factor to urban and rural income gap will take on reversely expanding trend[17]. If in China, when urban agglomeration effect is high, there will be more employment opportunities, it indicates urban agglomeration effect is positive and it has not reached the critical point of agglomeration diseconomy. In this study, we attempt to test the effects of agglomeration effect on employment opportunity.

According to the above mentioned factors influencing agglomeration effect, explanation of variables is as listed in Table 1.

Table 1 Explanatory and explained variables of agglomeration economic effect

Variable	Indicator	Meaning	Unit
y	Employment density	Number of employed in one square kilometer in a city	Person / km^2
x_1	Per capita GDP	Per capita gross domestic production of a city	Yuan
x_2	Public education level	Urban education expenses	10^4 yuan
x_3	Public medical care level	Number of urban hospital beds	Pcs
x_4	Population mobility	Urban passenger traffic volume	10^4 people

From *Statistical Yearbook of China* and *China City Statistical Yearbook*, we select related data of prefecture level cities, delete those lacking many variables, and use the mean value to replace cities lacking data of some years, and obtain 1430 samples of balanced panel data of 286 prefecture level cities (since cities are different and location is varied, the fixed effect model of panel data can reflect inherent nature of each city not changing with time). Descriptive statistics of variables are listed in Table 2.

Table 2 Descriptive statistics of variables

Variable	Mean value	Standard deviation	Min.	Max.	Number of samples
y	74.18	143.76	0.73	2322.62	1430
x_1	13885.31	12407.95	1660	140960	1430
x_2	69491.68	132683.40	39	2054600	1430
x_3	10362.15	9416.67	865	92800	1430
x_4	6159.68	7192.36	150	72793	1430

4 Empirical analysis

Firstly, we make econometric analysis of panel data for equation (3). Then, conduct F test to determine if mixed regression or panel data model is used. For the factor a_i specific to cities, the treatment is different. It is able to determine whether the fixed effect model or random effect model is suitable for panel data through Hausman test. F test value is 110.3, the significance is zero, indicating it rejects mixed regression hypothesis, so we should adopt the panel data model. Hausman test value is -275.04, the test fails. The efficiency of fixed effect is lower than that of random effect, but the endogenous steadiness is higher than random effect, thus we apply the fixed effect model.

Secondly, we carry out dynamic panel data estimation for equation (5). The estimation method can select linear differential lag variable as instrumental variable to conduct two-step generalized method of moments (GMM) estimation. The first step is to estimate the coefficient and make statistical judgment, and the second step is to conduct Sargan test and analyze excessive identification restrictive conditions of instrumental variables. Roodman summarized linear differential system GMM estimation and believed that two-step GMM estimation is especially suitable for dynamic estimation of data with small T but big N (*i.e.* short time dimension and big cross section dimension). Since the data we applied in this study have $T = 8$, $N = 286$, $N >> T$, theoretically, the linear differential system two-step GMM is better than ordinary two-step GMM. In addition, for excessive identification restrictive conditions of instrumental variables, the linear differential system two-step GMM is more steady in Hansen test than Sargan test, so we make two-step GMM estimation and linear differential system two-step GMM estimation for equation (5).

According to results of two-step GMM estimation, the coefficient of first-order lag explanatory variable is 0.627, which is relatively high, indicating Sargan test is significant and excessive identification restrictive condition is ineffective, *i.e.* instrumental variables of GMM estimation are not valid. By contrast, the coefficient of first-order lag explanatory variable of linear differential system GMM estimation is 0.225, much lower than GMM estimation. Except the public education level that is not significant, all other explanatory variables are significant at 1% level. Hansen test is not significant, indicating the excessive identification restrictive condition is effective. Therefore, without considering dynamic influence, it is required to adopt the fixed effect result; if considering dynamic influence, it is required to apply results of linear differential system two-step GMM estimation.

For estimation of static panel data model, fixed effect estimation results indicate that per capita income, public education level and public medical care level exert significant effect on urban employment density at 5% or 1% level, while the population mobility exerts insignificant effect on employment density. In comparison, for estimation of dynamic panel data model, results of linear differential two-step GMM estimation results reflect that per capita income, public medical care level, and population mobility exert

significant effect at 1% level, while the effect of public education level is not significant.

The fixed effect model applies (least square dummy variable) LSDV estimation method, so if there is endogenous problem, the estimation may be deviated. In comparison, linear differential two-step GMM estimation improves two-step GMM estimation, and is suitable for treating endogenous variables. Estimation results show that results of LSDV and linear differential two-step GMM are significant in main explanatory variables at 1% level, model results are steady. Thus, it can be considered that there exists urban agglomeration effect in China. Estimation coefficient denotes the elasticity of explanatory variable x to explained variable y.

Table 3 Estimation of the model

	I: Fixed effect				II: Two-step GMM estimation				III: Linear differential two-step GMM estimation			
$\ln y_{it}$	Coefficient	Standard deviation	t value	Probability	Coefficient	Standard deviation	z value	Probability	Coefficient	Standard deviation	z value	Probability
$\ln y_{it-1}$					0.627***	0.174	3.61	0	0.225***	0.080	2.8	0.005
$\ln x_{1,it}$	0.121***	0.023	5.20	0	0.061	0.036	1.67	0.095	0.637***	0.093	6.82	0
$\ln x_{2,it}$	0.033**	0.017	1.99	0.047	0.086***	0.030	2.89	0.004	−0.161	0.097	−1.67	0.096
$\ln x_{3,it}$	0.199***	0.066	3.01	0.003	−0.016	0.109	−0.15	0.883	0.253***	0.085	2.99	0.003
$ln x_{4,it}$	0.0003	0.021	0.01	0.989	−0.024	0.046	−0.52	0.606	0.187***	0.048	3.93	0
Constant term	0.417	0.552	0.75	0.45	0.349	0.803	0.43	0.664	−5.453***	1.062	−5.14	0
R^2	0.375	F_{test} =25.5	Probability	0	Chisquare	98.8	Probability	0	Chisquare	549.9	Probability	0
Hausman test	−261.7	Probability	−		Sargan test	17.9	Probability	0.003	Hansen test	3.42	Probability	0.49
Fixed effect F test	= 110.3		Probability	0								

Note: ** and *** denote significance at 5% and 1% level.

Due to possible endogenous problem, fixed effect model may have deviation, so we mainly focus on results of linear differential system GMM estimation results. Both estimation results indicate that per capita GDP exerts significant effect on employment density. This is because higher per capita GDP brings higher labor productivity and per capita income, and promotes more laborers to get jobs, more laborers can be included in the process of economic growth. This explains why economically developed cities become more and more crowded at present. However, the effect of public education level on employment density is not significant, possibly because explanatory variables we apply in this study are current year education expenditure, but the effect of education may take a long time to manifest. The effect of public medical care level on employment density is significant at 1% level, which is consistent with present conditions of China. Having excellent education and medical care resources will attract numerous high quality talents, thus medical care and education level exerts positive effect on industrial agglomeration, accordingly bringing more jobs to laborers. The population mobility density has significant effect of employment density at 1% level. More frequent flow brings more opportunities for allocation of resources and better play of market competition mechanism.

5 Conclusions and recommendations

5.1 Conclusions In this study, we apply fixed effect model to analyze static panel data, and two-step generalized method of moments (GMM) estimator to analyze dynamic panel data. Results reveal that per capita regional GDP, public medical care level, and population mobility have significant effect on employment density, while public education level exerts insignificant effect on employment density. Comparing estimation results of dynamic and static panel data, we can conclude that there exists agglomeration economic effect at prefecture level cities. In general, the agglomeration effect indicators significantly support urban employment density with higher agglomeration level, in other words, the employment opportunity will be greater. This reflects that at prefecture-level cities, the agglomeration economic effect still remains at the stage of scale economy. Moderately raising urban agglomeration level is favorable for bringing into play efficiency of urban agglomeration economic effect and accelerating urbanization process of China.

5.2 Recommendations At present, the CPC Central Committee and State Council printed and distributed *National New Urbanization Plan* (2014 – 2020). In this context, empirical test of agglomeration and employment density has important policy significance. For cities with lower economic agglomeration level, we recommend increasing investment in public infrastructure construction, improving public education and medical care, and promoting population and economic activities to agglomerate in downtown. Besides, we recommend formulating preferential policies in land, tax and employment, creating favorable investment environment, and attracting more enterprises and laborers to move to urban areas, so as to provide fundamental conditions for economic liftoff and urbanization. For mega-cities with high agglomeration level of economic activities, it is recommended to bring into play their radiation functions, strengthen infrastructure connection and public service sharing with surrounding cities, to realize regional coordinated development.

References

[1] HENDERSON JV, QUIGLEY J, LIM E. Urbanization in China: Policy issues and options [J]. Unpublished Manuscript, Brown University, 2009.

(To page 17)

the minimum purchase price is as the expected price, farmers' actual price must be equal to or higher than the expected price, and farmers can obtain the expected income and additional income. Based on the price protection, the expected income is the product of the expected price and yield, so the expected price is the expected market price in years when the minimum purchase price policy was not implemented, which may bring negative effects and affect farmers' enthusiasm for agricultural production. In years when the minimum purchase price policy has been implemented, price protection and realization of the expected income will improve farmers' enthusiasm for planting crops whose prices are protected and their planting area. That is, the implementation of the minimum purchase price policy will promote the sowing area of rice.

Since the implementation of the minimum purchase price policy for rice in Hubei Province in 2004, the planting area of rice in Hubei Province fluctuated but tended to increase on the whole. The regression results of factors influencing the planting area of rice show that the minimum purchase price policy for rice had positive effects on the planting area of rice in Hubei Province, so the agricultural protection policy was an effective measure to increase the planting area of rice. To ensure future food security, the sowing area of rice should be stabilized during the establishment of production input decisions; if farmers' level of education is low, it is needed to guild and protect agricultural production effectively to improve farmers' enthusiasm for agricultural production, insist on optimizing resource environment and policy environment of rice production, and establish more policies to protect farmers' benefit in market. Meanwhile, it is needed to use the price protection policy to popularize and sow high-quality rice varieties and provide effective technical support for farmers, strengthen drought monitoring and early warning capability, improve drought resistant security system, consolidate water conservancy infrastructure construction in paddy field, and enhance rice comprehensive production, drought resistance and disaster reduction capacity. At the present stage, the government should insist on using the minimum purchase price policy to conduct macro-control of market price to realize organic combination of market regulation and macro-control. At the same time, disadvantages of the minimum purchase price policy should be improved.

References

[1] LI W. Peasants' willingness to rice planting and its influential factors: Based on data of Zixing, Hunan [J]. Journal of Hunan Agricultural University(Social Science), 2010,11(5):7 – 13. (in Chinese).

[2] XIAO SX, LIU XH. Analysis on the impact factors for cotton planting area——Based on the investigation at Xinjiang, Henan, Jiangsu and Shandong [J]. Journal of Agrotechnical Economics, 2008(4):79 – 84. (in Chinese).

[3] ZHANG SP. Economic effects of China' s grain price fluctuation [J]. Finance and Economics, 2011(8): 93 – 102. (in Chinese).

[4] ZHONG J, QIN F. Effect of price support policies on grain production in China [J]. Modern Economic Science, 2012,34(3): 119 – 128. (in Chinese).

[5] GRDEN B. The economies of agricultural policies [M]. 2003.

[6] ANDREW PB. Wheat price policy in Pakistan: A welfare economics approach [M]. 1992.

[7] LI W. Peasants' willingness to rice planting and its influential factors: Based on data of Zixing, Hunan [J]. Journal of Hunan Agricultural University(Social Science), 2010, 11(5): 7 – 13. (in Chinese).

[8] ZHANG JS, WANG YP. Empirical study of factors influencing the yield increase of grain in China [J]. Hubei Agricultural Sciences, 2008, 47(4): 482 – 485. (in Chinese).

[9] WANG W, GUO XY. An econometric analysis of farmers' education level and rural economic growth [J]. Farm Economic Management, 2007(2): 43 – 45. (in Chinese).

[10] LIU XL, ZOU GH, LI Y, et al. Assessment of paddy soil productivity by using the multiple linear mixed-effect regression model [J]. Research of Agricultural Modernization, 2014, 35(5): 603 – 609. (in Chinese).

(From page 11)

[2] NOTTHAM RAY M. Urban geography[M]. New York: John Wiley & Sons, 1979

[3] FEI XT. Small towns and big issues[M]. Nanjing: Jiangsu People's Publishing Ltd., 1984. (in Chinese).

[4] FAN G. Discussion on equilibrium, disequilibrium and sustainability issues in structures[J]. Economic Research Journal, 1991(7): 13 – 20. (in Chinese).

[5] NIU WY, LI QQ. An understanding of China's new urbanization strategy[J]. Impact of Science on Society, 2010 (1): 14 – 20. (in Chinese).

[6] MILLS ES. Studies in the structure of the urban economy[J]. Baltimore, The Johns Hopkins Press,1972.

[7] GLAESER EL, KALLAL HD, *et al*. Growth in cities, National Bureau of Economic Research[J]. 1992.

[8] ROSENTHAL SS, STRANGE WC. Geography, industrial organization, and agglomeration[J]. Review of Economics and Statistics, 2003, 85(2): 377 – 393.

[9] BAI CE, DU YJ, TAO ZG, *et al*. Local protectionism and industrial concentration in China: Overall trend and important factors[J]. Economic Research Journal, 2004, 11(4): 29 – 40. (in Chinese).

[10] JIANG L, JI MH. Urbanization, regional innovation cluster, and spatial knowledge spillover——Empirical study based on spatial econometric model[J]. Soft Science, 2011(12):86 – 90. (in Chinese).

[11] LIU XY, SHAO J, XUE YL. Agglomeration and region economic growth: An empirical review based on panel data of Chinese prefecture cities[J]. Nankai Economic Studies, 2012, 3: 52 – 64. (in Chinese).

[12] LIAO XL. Dynamic analysis of the mechanism of fiscal expenditure and urbanization on urban-rural income gap[J]. Soft Science, 2012(4):33 – 38. (in Chinese).

[13] NIE HL, HAN Y, QIAN L. The dynamic comparison between the level of urbanization and the economic growth based on panel data[J]. Soft Science, 2012(5):27 – 31. (in Chinese).

[14] HENDERSON JV. The urbanization process and economic growth: The so-what question[J]. Journal of Economic Growth, 2003(8):47 – 71.

[15] LU M, GAO H, ZUO TH. Urban size and inclusive of employment [J]. Social Sciences in China, 2012(10): 47 – 66. (in Chinese).

[16] CHEN GL, CHEN JJ. Industrial relationship, geo-spatial database and the secondary and tertiary industries gathered together[J]. Management World, 2012 (4):82 – 100. (in Chinese).

[17] REN XH, ZHANG ZY. Transportation infrastructure, factor mobility and urban-rural income gap[J]. Management Review, 2013, 25(2): 51 – 59. (in Chinese).

[18] ROODMAN D. How to do xtabond2: An introduction to difference and system GMM in Stata[J]. Stata Journal, 2009, 9(1): 86.