Human Capital Dynamics across Provinces in China: A Spatial Markov Chain Approach

Author: Wenxuan Yang[1][2]

Abstract

Human capital is an essential driver of the economy. The regional distribution of human capital affects the development process of the economy. This paper applies the Markov chain and its spatial form extensions to the human capital index of China from 1985 to 2019 and calculates the transition probabilities between different human capital levels. This paper shows that China's regional human capital is unevenly distributed, with significant differences across provinces, and this difference does not tend to decrease over time. Furthermore, by comparing the results of non-spatial and spatial Markov chains, this study confirms the spatial effects from neighboring regions on human capital changes. In other words, rich neighboring provinces increase the probability that a province becomes better off, and conversely, a province surrounded by poor neighboring provinces has a higher probability to worsen. This article emphasizes the importance of incorporating spatial effects into regional human capital in China.[3]

Keywords: Human Capital, Markov Chain, Spatial Effect, China

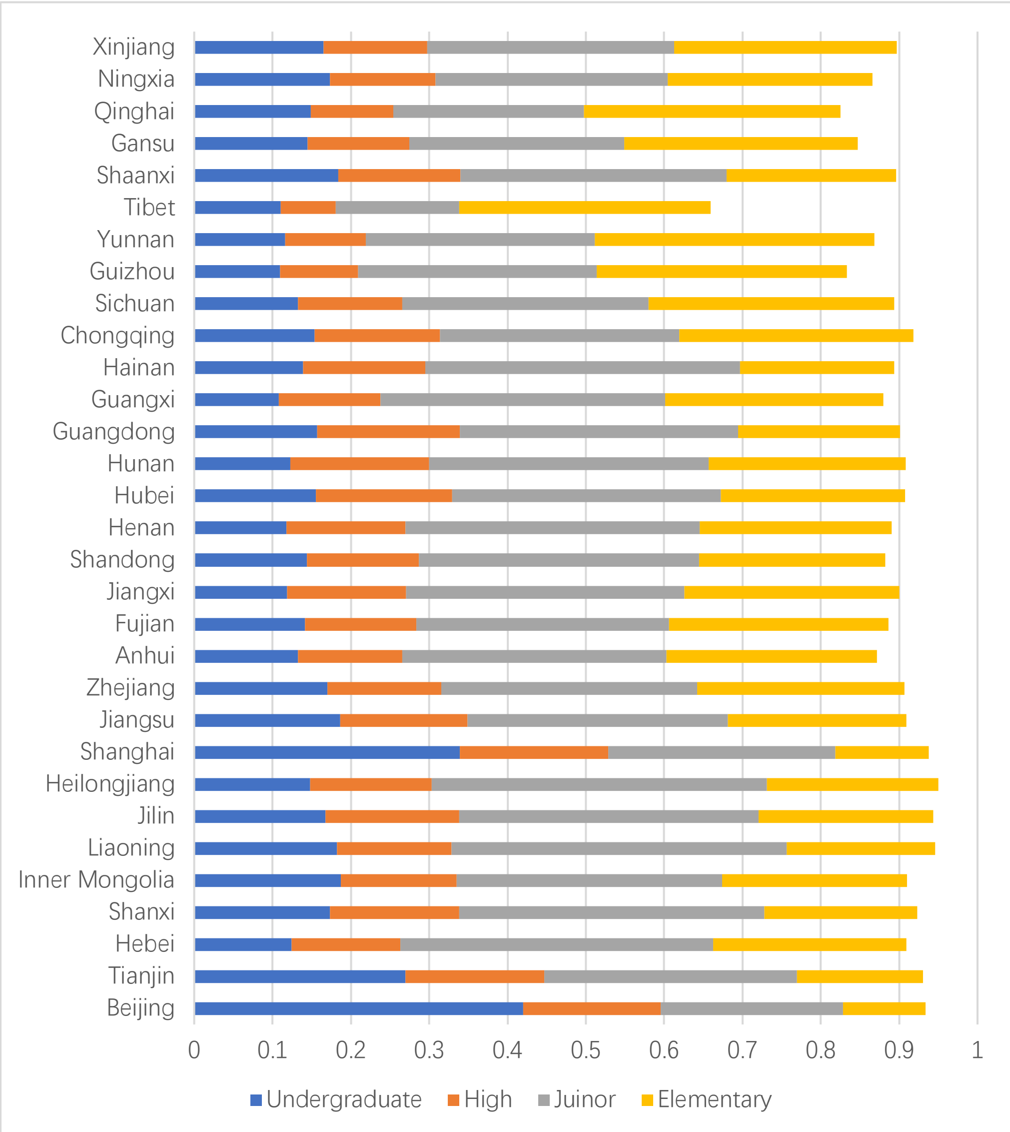
1 Introduction

In recent years, China has placed greater emphasis on talent. In March 2016, the Chinese Communist Party Central Committee issued the "Opinions on Deepening the Reform of the Institutional Mechanism for Talent Development", proposing a more active, open, and effective talent introduction policy. Subsequently, local governments in China's second-tier cities introduced "New Talent Policies," which aimed at attracting college students and other highly educated talents to settle, including cash subsidies, employment settlements, and low-interest housing loans, in what the media called a "war for talent." By the end of April 2021, China had 3,191 items of local talent policies (Xia 2021).

[1] Graduate School of International Development, Nagoya University, Nagoya, Japan
[2] E-mail: yang.wenxuan@h.mbox.nagoya-u.ac.jp

On the contrary, the requirements for settling in big cities such as Beijing and Shanghai are getting higher and higher, and generally, only talents with master's degrees or above can meet the requirements for settling in these cities. Although these first-tier cities have introduced some talent incentive policies, they are unattractive compared to the living and housing subsidies and fast-track settlement policies in second-tier cities. China's first-tier cities are rich in university resources and do not lack talent, so they do not need many welfare policies to attract talent to stay and work locally.

Figure 1 Regional Educational Attainment Percentage in 2020



Source: National Bureau of Statistics of China

The distribution of human capital and labor force in China is uneven. This distribution is influenced to some extent by the talent policy in recent years, and more importantly, it is related to the degree of economic development of a region. Figure 1 shows the proportion of the population with various levels of education in different regions. For human capital, we are more concerned with the proportion of the population with higher education. For most regions, the percentage of undergraduate is between 10% and 20%, but in the more developed provinces or municipalities, the number is much higher, such as Beijing (41.98%), Shanghai (33.87%), Tianjin (26.94%).

Large cities and developed regions can attract talents, while less developed regions will face the problem of brain drain. Nie and Liu (2018) found talent flows come from provinces with many students and universities and tend to flow to regional economic centers or developed provinces, such as Beijing, Shanghai and Guangdong Province. It may affect the dynamics of human capital and exacerbates the imbalanced distribution of human capital in China. Human capital is an essential factor for economic development, which can improve the efficiency of using physical capital. Chu and Cao (2019), Verginer and Riccaboni (2020) have found that talent mobility will make talents gather in developed areas and lead to shortage of talents in backward areas. It will intensify the imbalance of regional development.

Massive talent flow will cause changes in regional human capital, and these changes may be influenced by neighboring regions due to spillover effects. There is an emerging academic consensus that human capital has spillover effects like knowledge and technology (Fang & Luo 2019) (Lu & Zhou 2014). With the help of spatial econometrics, we can capture the spillover effects of human capital. In general, the spillover effects of human capital are most pronounced for neighboring regions. It implies that the role of neighbors has an essential effect on human capital accumulation, which is also called the spatial effect.

Spatial effect should not be ignored when considering regional development and human capital accumulation. However, before Rey (2001), research using Markov approach did not include the role of geography. Based on the definition of Markov Chain, the probability of transferring from one state to another depends only on the state of the previous period. Spatial Markov Chain developed by Rey (2001) integrated local spatial statistics into Markov Chain framework. This is a new way to consider spatial effects in dynamics. In China's case, most papers are using spatial modelling in measuring spillover effects from human capital. However, spatial modelling cannot detect the role of spatial effects in human capital evolution. This paper aims to fulfill the research gap by combining spatial econometrics and traditional Markov Chain methodology together and check whether spatial effects would affect human capital dynamics. The purpose of this paper is to investigate the characteristics of dynamic changes in regional human capital, to calculate the transition probabilities of each level of human capital by applying Markov chains, and to apply Markov chains in the spatially extended form to compare the differences in dynamic changes in spatial and non-spatial forms.

This paper shows that regional human capital imbalances are increasing in China and spatial effects matter in this process. By comparing Markov chains in non-spatial and spatial forms we find that a region is more likely to improve if it is surrounded by neighbors with high human capital level, and conversely, a region has a higher probability to relegate if it is surrounded by neighbors with less human capital.

This paper contributes to the literature of China's regional human capital development in three fronts. First, an increasing number of articles use Markov chains to study changes in regional income (Bode & Nunnenkamp 2011; Rey et al. 2016; Kang & Rey 2018). In this paper, Markov chains are used in the human capital index to calculate the dynamic distribution and transition probability of regional human capital in China. Second, a growing of literature finds that human capital has spatial effects(Lu & Zhou 2014; Xu & Li 2020). Based on the methodology of Rey (2001), this paper incorporates spatial factors into the Markov chain framework and verifies that spatial effects affect human capital dynamics. Lastly, most articles on spatial econometrics use spatial autocorrelation analysis to do static explorative analysis. This paper combines local spatial autocorrelation[1] with Markov chain framework in a dynamic way.

The structure of the rest of this paper is as follows. Chapter 2 summarizes the relevant literature and research gap on regional variations of human capital in China. Chapter 3 introduces the research methodology used in this paper. Chapter 4 shows the empirical results, and Chapter 5 summarizes and discusses the results.

2 Related Literature

Schultz (1961), who first introduced the concept of human capital, argued that investment in human capital is as much an explanation of economic growth as investment in physical capital. Moreover, in the case of workers, the human capital investment could explain wage increases.

This revolutionary theory emphasized the importance of people in productive activities and found an alternative source of growth for economic growth.

With the development of economics, the role of human capital for economic growth and social development has become a consensus in the academic community (Lucas 1988). The importance of human capital in reducing regional disparities and promoting balanced economic development has also received increasing attention (Acemoglu 2012). Conversely, the uneven distribution of human capital may also be a reason for the uneven regional development of the economy.

For China, the "strong east and weak west" development pattern has been a problem since the reform and opening up. To alleviate this problem, China has introduced a series of policies to balance regional development since 2000. The most representative of which is the Development of China's Western Regions, which encourages university students to support the construction in the west and gives generous incentives. This policy has promoted the flow of talent and economic development in the western region, but it still has not changed the talent and economic development pattern.

2.1 Dynamics and Spillover Effects of Human Capital in China

Li et al. (2013) measured total and per capita human capital in six Chinese provinces based on the Jorgenson − Fraumeni lifetime income method and found that human capital and per capita income have a similar distribution pattern. The changes in human capital have structural characteristics, i.e., slow growth from 1985-1995 and rapid growth after 1995. In addition, the gap in human capital per capita between developed and underdeveloped provinces is widening.

Zhang and Huang (2020) and Peng (2019) use the Gini and Theil coefficients to measure per capita capital inequality in China and show that the human capital structure in China is characterized by distinct phases, with different trends in each decade. Spatially, spatial agglomeration is evident and has a polarization characteristic of "higher the higher and lower the lower." The study of Li and Chen (2019) also confirms this spatial unevenness and polarization.

Previous studies have shown that the distribution of capital per capita in China is characterized by a "high in the east and low in the west," which is similar to per capita income and that this imbalance is increasing, as evidenced by the fact that regions with high human capital stocks are growing faster and those with low human capital stocks are growing slower.

Most papers are using spatial modeling in measuring spillover effects from human capital. Using exploratory spatial data analysis (ESDA) in a spatial econometric approach, Lu and Zhou (2014) found similar human capital and economic development characteristics in China and developed a spatial Lucas model. This study found significant positive spatial spillover effects of human capital in most Chinese provinces. The study by Fang and Luo (2016) reached similar conclusions and verified that the spillover effect of human capital has a growth effect on the economy using the GMM approach. Previous literature has verified that human capital has spillover effects, but few researches are discussing the role of spillover effects in human capital dynamics.

2.2 Markov Chains and Spatial Applications

This study mainly draws on Rey's (2001) Markov chain and spatially extended form in terms of methodology. Rey (2001) extends the Markov chain framework, integrates Markov chains and spatial correlation analysis, allowing Markov chains to incorporate with regional context. The empirical study of regional income from 1929-1994 concluded that geography impacts the evolution of regional income distribution. After this, many scholars have used Markov chains in the study of regional income changes (Le Gallo 2004) (Hammond 2004).

In this paper, I plan to apply Markov chains in studying regional human capital changes. Human capital, as many economic factors, is mobile and has spillover effects. So theoretically, the dynamics of human capital in a region can also be affected by its neighbors. Peng (2019) found human capital inequality in China fluctuates every ten years, and there is a clear trend towards polarization of the spatial distribution. This implies that the role of space has an impact on human capital changes. In addition, Zhang and Huang (2020) Lu and Zhou (2014) also confirmed human capital has spillover effects to neighboring regions. However, these studies did not find whether spatial effects could impact human capital dynamics or not, due to the limitation of methodology. This study aims to solve this problem by using a Markov chain research framework. It is an attempt to move from static analysis to dynamic analysis in researching human capital issue.

3 Methodology and Data

3.1 Markov Chain

Traditional ESDA methods can only characterize the spatial features of a given year or compare the changes in spatial features between two years. Markov chains, on the other hand, can provide probabilistic information about dynamic changes. Specifically, Markov chains can discretize a random sequence of continuous states into several types in a specific application and calculate the probability distribution of each type and the general trend of its evolution to approximate the spatio-temporal characteristics of the variables (Tao & Qi 2013).

According to (Rey 2001), in mathematical terms, assuming a total of k species and T times, the distribution of states at time t can be represented by a 1*k vector of

$$P_t = [P_{1,t}, P_{2,t} \dots P_{k,t}] \tag{1}$$

This vector represents the distribution probability of the different species at time t.

The probability of transition in each region can be represented as $m_{t,i,j}$, where t is the time, i is the initial state and j is the end state. $m_{t,i,j}$ represents the probability of that at time t, this region converting from state i to state j at the following period. In the basic Markov model, we consider transition probability is time-invariant, which means $m_{t,i,j} = m_{t+b,i,j}\ \forall b$. Therefore, we can write weight matrix as

$$M = \begin{pmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{n1} & \cdots & m_{nn} \end{pmatrix} \tag{2}$$

which must meet the condition $\sum_{j=1}^{n} m_{ij} = 1$.

Assume one simple scenario where there are only two states, Low and High, for a variable. These two states can switch to the other state or stay the same in the following period. Moreover, there are only two times t_0 and t_1. Then, the transition probability matrix M for these two states can be represented in Table 1.

Table 1 Simplified Markov Transition Matrix

t_1 / t_0	Low	High

Low	m_{LL}	m_{LH}
High	m_{HL}	m_{HH}

Notes: This table is drawn by author himself based on (Rey 2001)

With the help of the transition probability matrix M, we can compute the state distribution for period t + 1 as follows.

$$P_{t+1} = P_t M = P_{t-1} M^2 = \cdots = P_0 M^t \tag{3}$$

where P_0 is the initial state distribution. From this equation, we know the evolution of Markov Chain is totally decided by transition probability matrix.

Markov chains can also give us some other information, such as the time it takes to reach a steady state and it's time to transit from one state to another. The details will be shown in the empirical analysis.

3.2 Spatial Expansion of Markov Chain

There is an assumption in the traditional Markov chain, which is that the effect of spatial factors or spatial homogeneity is not considered. We expand the k*k matrix into the form k*k*k to include effects from neighbors. For simplicity, as in Table 2, we consider here a simple scenario. The variables have only Low and High states, and the neighbors also have only Low and High states. t_0 and t_1 denote the starting and ending time points, respectively.

Table 2 Simplified Spatial Markov Transition Matrix

Spatial Lag	t_0 \ t_1	Low	High
Low	Low	$m_{LL\|L}$	$m_{LH\|L}$
	High	$m_{HL\|L}$	$m_{HH\|L}$

High	Low	$m_{LL\|H}$	$m_{LH\|H}$
	High	$m_{HL\|H}$	$m_{HH\|H}$

Note: This table is drawn by author himself based on Rey (2001)

In Table 2, spatial lag indicates the state of neighbors, which is analogous to the presence of a precondition. For example, $m_{LH|L}$ denotes the transition probability of the region moving from Low to High when the neighbor is Low. By comparing the spatial transition matrix with the traditional transition matrix, we can see if the spatial context has a significant effect on the transition probability. For example, $m_{LH|L} < m_{LH}$ suggests that having poorer neighbors is detrimental to the region's conversion from Low to High. Conversely, if the spatial context does not have a noticeable effect on the transition probability, then

$$m_{ij|1} = m_{ij|2} = \cdots = m_{ij|k} = m_{ij} \ \forall\ i, j. \tag{4}$$

Spatial Markov chains can provide information on whether transition across classes is related to neighbors and the magnitude of the influence from neighbors. But spatial Markov chains cannot measure the specific magnitude of the impact from neighbors.

3.3 Data

The data for this study come from the Human Capital Index Project of China Center Human capital and labor Market Research (CHLR). This project uses and improves upon the internationally widely used Jorgenson-Fraumeni income calculation method (hereafter referred to as the J-F method) to estimate regional human capital in China. Essentially, the J-F lifetime income approach measures the level of human capital as the present value of an individual's lifetime income over his or her expected lifetime. Assuming that an individual's human capital can be traded in the market like physical capital, the price is the present value of the individual's future lifetime earnings over his or her expected lifetime.

J-F income-based approach is the most widely used measure of human capital today. It estimates expected future lifetime earnings based on the currently observed earnings of a cross-section of individuals. In this method, personal income is expected to grow at a certain growth rate, discounted to the present at a fixed discount rate. An individual's expected lifetime earnings are calculated from variables such as years of schooling, gender, age, observed average earnings of the cohort, and induction rate. The total human capital stock of the region is obtained by multiplying the individual human capital of the region by the total number of individuals (Li et al. 2013). The lifetime income approach can accurately reflect the role of long-term investments such as education and health in human capital accumulation.

This paper is using the calculation results from Project 2021, which covers the period from 1985 to 2019. The unit of human capital index is thousand yuan. This paper used real average human capital value for each province, which is adjusted by Consumer Price Index (CPI) and the base year is 1985. This database covers 31 provinces and municipalities in mainland China, and the Human Capital Index data from 1985-2019 are used in this paper.

Table 3 Descriptive Statistics of Regional Human Capital Index

	In 1985				In 2019			
	West	Middle	East	Northeast	West	Middle	East	Northeast
Mean	45.08	45.83	69.9	52.67	406.83	494.5	649.8	422
S.D.	5.68	4.17	18.94	2.31	105.68	92.80	211.00	71.97
Min	36	38	52	50	225	374	357	339
Median	44.5	47	62.5	54	434	501.5	636.5	460
Max	53	49	108	54	563	617	1082	467
Obs.	12	6	10	3	12	6	10	3

Note: China are divided into west, middle, east and northeast, according to the 2011 classification by the National Bureau of Statistics of China

Source: Author’s calculation using data from China Center for Human Capital and Labor Market Research (CHLR)

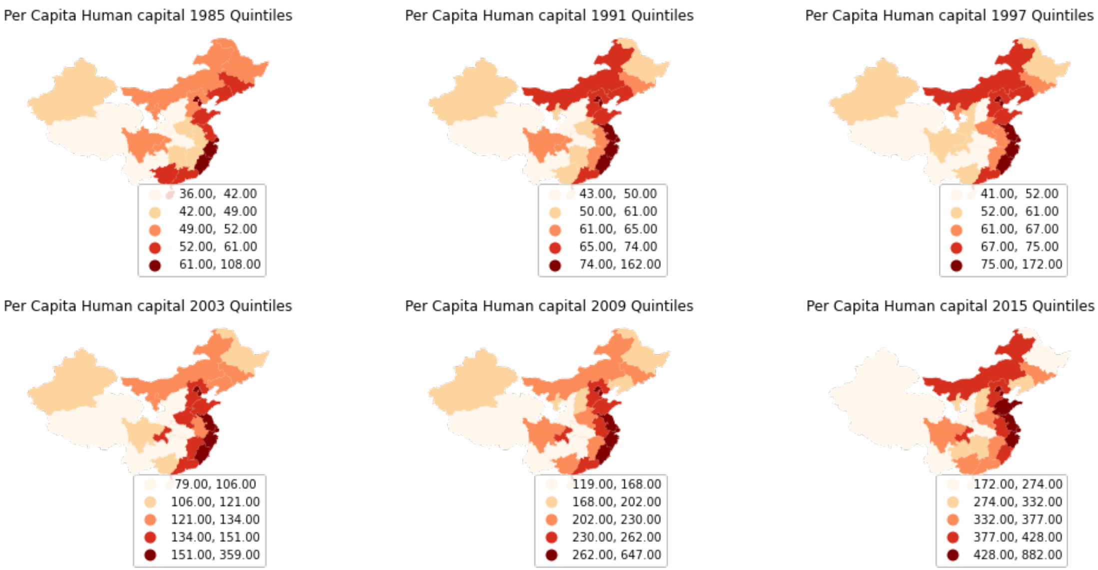
Table 3 compares the descriptive statistics of the human capital index for the provinces between the four geographical regions of China for the starting year (1985) and the final year

(2019). Both the mean and median indicate that in both years, China's regional human capital is characterized by a high east and low west. However, the relative position of the Northeast has changed. In 1985, the human capital index of Northeast was between the East and Middle. But in 2019, the index for the Northeast is lower than that of Middle. This result implies China's regional human capital pattern is changing.

4 Empirical Analysis

4.1 The Dynamics of Human Capital

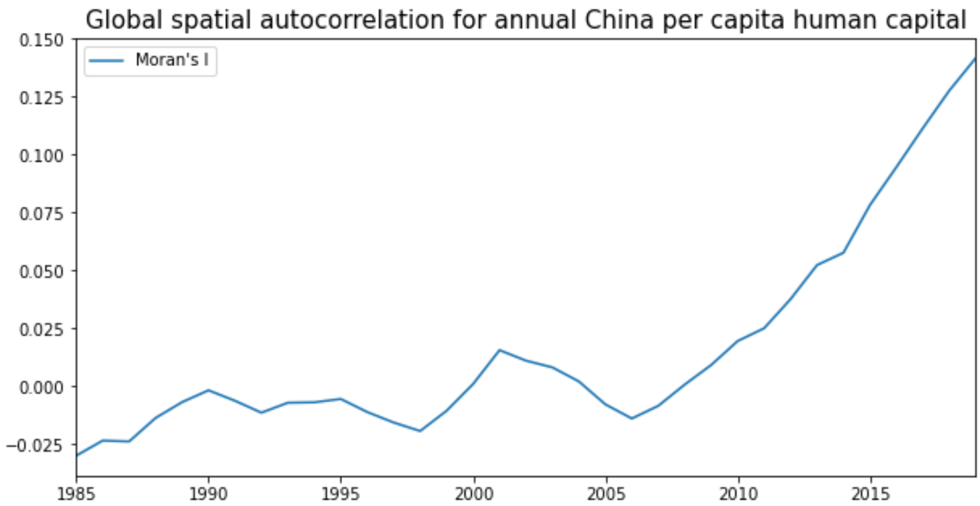
Figure 2 Change in Human Capital Index Distribution



Source: Author's calculation using data from China Center for Human Capital and Labor Market Research (CHLR)

Figure 2 shows the regional differences and evolution of human capital in China. Based on the quintiles of the human capital index, Chinese provinces are divided into five levels, with darker colors indicating higher human capital. Some features are concluded from these evolution maps. First, the regions with higher human capital index are concentrated in the developed eastern coastal provinces, and the neighboring regions are also at a relatively high level. Second, regions in western China have lower human capital indexes, with Xinjiang province dropping to the lowest level in the latest graph. Third, China's northeastern region drops from the higher level to the lowest level. Although the human capital levels of Chinese provinces are in dynamic change, the pattern of higher in the east and lower in the west is unchanged. It implies that the gap between the eastern and western provinces does not decrease over time, and there is an increasing spatial clustering of human capital.

Figure 3 Yearly Moran's I of Human Capital Index[2]



Source: Author's calculation using data from China Center for Human Capital and Labor Market Research (CHLR)

From the maps in Figure 2, we found the sign of spatial clusters. Then we can use Moran's I as a measure to describes the global spatial correlation of variables. Figure 3 shows Moran's I of human capital from 1985 to 2019. We can see the spatial correlation in human capital has experienced an increase from negative to positive during this period, although the change in values is small. This implies positive neighboring effects in human capital are increasing.[3] It also implies that we should consider spatial effects in researching on the changes of regional human capital.

4.2 Markov Chain

Based on the way of Rey (2001), we divided Chinese provinces into five quintile levels according to the human capital per capita index, namely poor, lower, middle, upper and rich. The level of each region is changeable over time. Markov chains can record the frequency of these changes in each year and calculate the probability of moving from one state to another. This is an important message in a dynamic change. I only used the full sample in this research since it requires a large sample in Markov chain analysis. For example, Rey (2001) applied 3120 samples in his research. My paper only has 1085 samples, if I divide full sample into several sub-groups, movement between different groups will be less captured.

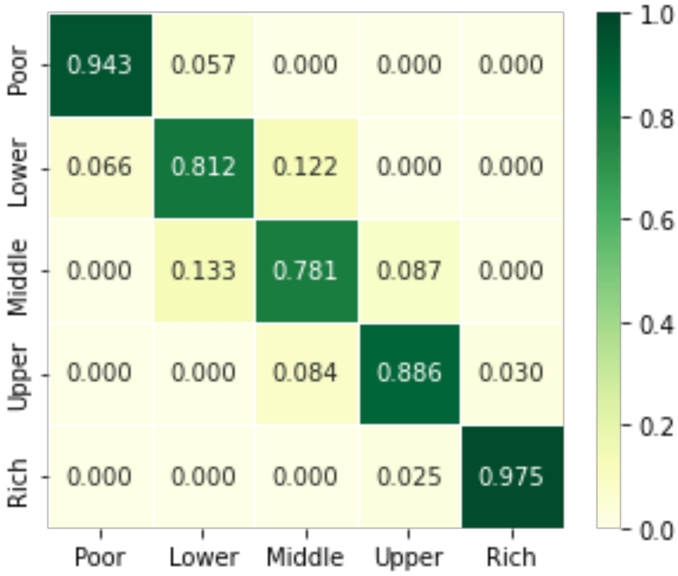
Table 4 Classic Markov Transition Frequency

	Poor	Lower	Middle	Upper	Rich
Poor	230	14	0	0	0
Lower	14	173	26	0	0
Middle	0	26	153	17	0
Upper	0	17	179	6	0
Rich	0	0	0	5	194

Note: This frequency records the transition of 31 regions in China from 1985-2019. The transition probability of Figure 4 is calculated based on this frequency.

Source: Author's calculation using data from China Center for Human Capital and Labor Market Research (CHLR)

Figure 4 Classic Markov Transition Matrix



Source: Author's calculation using data from China Center for Human Capital and Labor Market Research (CHLR)

Table 4 and Figure 4 record the transfer frequency and transfer probability of sample points among the five levels, respectively. The transfer probability is derived from the transfer frequency, so they convey the same information. For example, in Table 4, we focus regions on Poor class in the first period. 230 sample points are still on Poor class in the next period, and 14 sample points enter the Lower class. No points enter the higher levels of Middle, Upper and Rich. Combining the results with the other rows we can see that all transfers are made in the level adjacent to the starting point and no jumps across levels occur. The values on the diagonal are much larger than the other values, which indicates that most of the points stay in their original level, and even Middle, which is the least stable, has a 78.1% probability of staying in its original position in the next period. In summary, the transition matrix in Figure 4 implies that richest and poorest areas (Rich class and Poor class in Figure 4) are very difficult to change, while Middle class has a relatively higher probability of making changes.

4.3 Spatial Markov Chain

Table 5 Spatial Dependance Test

Number of classes: 5		
Number of transitions: 1054		
Number of regimes: 5		
Test	Likelihood Ratio	χ^2
Stat.	47.599	46.883
DOF	29	29
p-value	0.037	0.043

Source: Author's calculation using data from China Center for Human Capital and Labor Market Research (CHLR)

Likelihood Ratio test (hereafter referred to as the LR test) and the χ^2 test are used to test whether the regional context has a significant effect on the estimation results of the Markov Chain. The null hypothesis of this test is that the variables are spatially independent. In Table 5, P-value shows that both the LR test and χ^2 test reject the null hypothesis, which indicates that the variables are not spatially independent. In other words, we cannot ignore the interaction of variables in space, which is why we use the spatial Markov Chain.

Figure 5 Comparison of Classic and Spatial Markov Transition Matrix

Source: Author's calculation using data from China Center for Human Capital and Labor Market Research (CHLR)

Figure 5 shows both the classical Markov probability matrix and the spatial Markov probability matrix. A classical Markov chain does not consider the influence of neighbors, while a spatial Markov chain considers this spatial influence from neighbors. In Figure 5, I divide the neighbors into 5 classes from small to large by quintiles. Level 1-5 correspond to 'Poor', 'Lower', 'Middle', 'Upper', and 'Rich', indicating the stratum in which the neighbors' human capital levels are located.

Different neighbors have a significant effect on the movement of human capital levels across classes. If a region is surrounded by neighbors with rich human capital, it will have a higher probability of staying in the same class or moving to a higher class, while if it is surrounded by neighbors with less human capital, the region will have a higher probability of falling into a lower level. For example, the middlemost grid in each graph indicates the probability that a 'Middle' region will remain in 'Middle' in the next stratum. When we do not consider the effect of spatial factors, this probability is 0.781. when surrounded by poor areas, this probability decreases to 0.615. Conversely, if one region is surrounded by rich areas, the probability of remaining in this stratum is 0.875. In addition, 'Middle' regions with poor neighbors have a 0.308 probability of falling into the 'Lower' class, which is much higher than if space is not a factor (0.133) and if they have rich neighbors (0.021)

Cases in other levels also support neighbor’s effects. – neighbor’s level will affect the transition probability of moving up and dropping down. It implies the role of spatial effect. According to Tobler's First Law of Geography, geographical things are correlated with each other in terms of spatial distribution. Therefore, the results of this study are consistent with Tobler's First Law.

5 Discussion

The regional distribution of human capital in China is uneven, and this distribution is changing over time. Most references describe this change based on the Gini coefficient and trend over time without considering the spatial factor. Li et al. (2013) and Li et al. (2009) found the rapid increase in human capital, by gender and urban-rural distinction, but not discussing regional difference. Fraumeni et al. (2019) divided China into four regions: West, Central, East, and Northeast, and discussed regional patterns as well as growth differences. However spatial effects are not addressed in this study. On the other hand, articles studying human capital spillover effects generally use spatial models to capture the role of spillover effects. Wang (2013) found the spillover effect from human capital will speed up the beta convergence process. Very few articles use dynamic exploratory data analysis methods to study human capital issues with spatial effects.

This paper did a dynamic analysis by applying Markov Chain framework. Results show that regional context influences regional human capital changes. This phenomenon can be explained by talent migration and urbanization. As the carrier of human capital, people are mobile, though there are some limitations from Hukou system. Migration is an essential factor of human capital increase. In China, human capital and labor are flowing to developed regions and big cities to pursue better job opportunities. The attractiveness of large cities themselves has exacerbated the uneven development of human capital in China.

High-quality education plays an important role in human capital development. There are 42 "first-class" universities in China, and Beijing and Shanghai account for 12 of them, while other provinces only have 1-2 first-class universities at most. Inequality in regional education resources is a significant problem in China. Inequality in regional education resources and access to education can have a direct impact on local economic development (Qiu and Wen 2010). After finishing study, graduates are more likely to work locally. To solve the problem of regional imbalance in the distribution of human capital, apart from the central government's conscious policy inclination towards the less developed regions, it is also essential for the resources of universities to be inclined towards the less developed regions.

The expansion of university enrolment in less developed regions is one way to alleviate the problem of regional imbalances in human capital. Zhou et al. (2017) and He et al.ß (2020) found expansion of higher education enrolment can contribute to human capital accumulation. If talent attraction policies in less developed regions could be coordinated with university expansion policies, more graduates would choose to study and work in these regions, which would alleviate the polarization issue in the distribution of human capital in China.

6 Concluding Remarks

This article aims to explore the characteristics of regional human capital dynamics in China and compare the difference between considering and not considering spatial effects. To calculate the transition probabilities, this paper applies Markov chains and spatially extended Markov chains methods in regional human capital index. Results indicate that China's human capital development is uneven, with some wealthy coastal regions having much higher per capita human capital than others. Results imply that the dynamic change may be correlated to the increasing spatial effects. From 1985 to 2019, the degree of spatial autocorrelation increased. It turned from an insignificant negative correlation to a significant positive correlation.

Another important finding is regional context matters transition process, which is similar with Wang (2013) but in spatial Markov Chain method. Likelihood ratio test and the Chi-square test indicate that we cannot ignore spatial effects in regional human capital. Comparing Markov chain result and spatial Markov chain result, we find how spatial effects influence the transition process. If one region is surrounded by rich regions, it is more likely to move up to the upper level. On the opposite, regions surrounded by poor neighbors have a higher possibility of dropping down to lower level.

Finally, further research about dynamics in China's human capital can be extended in at least three fronts. First, to confirm the role of neighbors in human capital, more robustness check can be done by using other datasets. Second, alternative dynamic explorative analysis could be considered, like directional LISAs and local indicator of mobility association (LIMA). These methods can give more information about transition process. Finally, to measure the magnitude of neighbors' effect, spatial modelling for human capital is needed. For instance, Lu and Zhou (2014) used a spatial-extended Lucas model to estimate human capital spillover effect between provinces.

Acknowledgement

This work was financially supported by JST SPRING, Grant Number JPMJSP2125. The author (Initial) would like to take this opportunity to thank the "Interdisciplinary Frontier Next-Generation Researcher Program of the Tokai Higher Education and Research System."

Endnotes

1. Local spatial autocorrelation is a term that refers to neighboring effects.
2. This paper is using distance weight matrix, please see Appendix 1 for details.
3. Moran's I results have shown us that there is the phenomenon of spatial effects, but it cannot detect its impact on regional inequality. In this paper we do not discuss this issue.

References

Acemoglu, D. 2012. What does human capital do? A review of Goldin and Katz's The race between education and technology. *Journal of Economic Literature.* 50(2): 426-63.

Anselin, L., Sridharan, S., & Gholston, S. 2007. Using exploratory spatial data analysis to leverage social indicator databases: the discovery of interesting patterns. *Social Indicators Research.* 82(2): 287-309.

Barro, R. J., Sala-I-Martin, X., Blanchard, O. J., & Hall, R. E. 1991. Convergence Across States and Regions. Brookings Papers on Economic Activity. 1991(1): 107–182. https://doi.org/10.2307/2534639

Bode, E., & Nunnenkamp, P. 2011. Does foreign direct investment promote regional development in developed countries? A Markov chain approach for US states. *Review of World Economics.* 147(2): 351-383.

China Center for Human Capital and Labor Market Research(CHLR). Human capital index project 2020 http://humancapital.cufe.edu.cn/en/Human_Capital_Index_Project/Project_2020/Human_Capital_Brief_Report.htm

Chu, E., & Cao, C. 2019. Does Talent Flow Narrow the Regional Economic Disparities? ——Empirical Evidence from Technology Transfer. *Financial Science.* (9): 99-112.

Danny T. Quah. 1996. Twin Peaks: Growth and Convergence in Models of Distribution Dynamics. *The Economic Journal*. 106(437): 1045–1055. https://doi.org/10.2307/2235377

Fang, C., & Luo, Y. 2016. A study on the impact of educational human capital and its spillover effects on China's economic growth - a spatial econometric analysis based on Lucas model. *Education and Economics.* (4): 21-29.

Fraumeni, B. M., He, J., Li, H., & Liu, Q. 2019. Regional distribution and dynamics of human capital in China 1985–2014. *Journal of Comparative Economics.* 47(4): 853–866. https://doi.org/10.1016/j.jce.2019.06.003

Fu, Y. 2014. *A study on the contribution of human capital and its structure to China's economic growth.* Changchun: Jilin University (Doctoral dissertation).

Hammond, G. W. 2004. Metropolitan/non-metropolitan divergence: A spatial Markov chain approach. *Papers in regional Science.* 83(3): 543-563.

He, X. G., Luo, Q., & Chen, J. L. 2020. High-Quality Human Capital and Upgrading of Urban Industrial Structure in China: Evidence from Enrollment Expansion. *Economic Review.* (4): 3-19.

Kang, W., & Rey, S. J. 2018. Conditional and joint tests for spatial effects in discrete Markov chain models of regional income distribution dynamics. *The Annals of Regional Science.* 61(1): 73-93.

Le Gallo, J. 2004. Space-time analysis of GDP disparities among European regions: A Markov chains approach. *International Regional Science Review.* 27(2): 138-163.

Li, H., Fraumeni, B. M., Liu, Z., & Wang, X. 2009. Human Capital in China (Working Paper No. 15500; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w15500

Li, H., Jia, N., Zhang, X., & Fraumeni, B. M. 2013. Regional distribution and development dynamics of human capital in China. *Economic Research.* 48(7): 49-62.

Li, H., Liang, Y., Fraumeni, B. M., Liu, Z., & Wang, X. 2013. Human Capital in China, 1985-2008. *Review of Income and Wealth.* 59(2): 212-234. https://doi.org/https://doi.org/10.1111/j.1475-4991.2012.00517.x

Li, J., & Chen, Y. P. 2019. Regional polarization and dynamic evolution of human capital distribution: A measure based on the provincial dimension in China. Statistics and Information Forum. 34(6): 44-50.

Lu, J., & Zhou, H. M. 2014. An empirical analysis of the spatial spillover effects of human capital in Chinese provinces: Based on ESDA approach and spatial Lucas model. *Journal of Population.* 36(6): 48-61.

Lucas, R. E. 1988. On the mechanics of economic development. *Journal of monetary economics.* 22(1): 3-42. https://doi.org/https://doi.org/10.1016/0304-3932(88) 90168-7

Ma, H. 2021. The impact of human capital distribution structure on technological innovation: Also on the mediating and moderating effects of income disparity. *Journal of Central China Normal University (Humanities and Social Sciences Edition).* 60(1): 34-44.

Nie, J. X., & Liu H. L. 2018. Spatial Pattern and the Resulting Characteristics of Talent Flows in China. *Scientia Geographica Sinica.* 38(12): 1979-1987.

Peng, S. H. 2019. The dynamic evolution of regional human capital inequality and its spatial distribution in China. *Journal of Central University of Finance and Economics.* (11): 115- 128.

Qiu, J. P., & Wen F. F. 2010. Study on the regional distribution of higher education resources in China. *China Higher Education Studies.* (7): 12-16.

Rey, S. J. 2001. Spatial empirics for economic growth and convergence. *Geographical analysis.* 33(3): 195-214.

Rey, S. J., Kang, W., & Wolf, L. 2016. The properties of tests for spatial effects in discrete Markov chain models of regional income distribution dynamics. *Journal of Geographical Systems.* 18(4): 377-398.

Schultz, T. W. 1961. Investment in human capital. *The American economic review.* 51(1): 1-17.

Sun, H. B. 2017. *Research on the impact of human capital and its spatial distribution on industrial structure upgrading in China.* Changchun: Jilin University (Doctoral dissertation).

Tao, X. H., & Qi Y. W. 2013. Spatial-temporal Evolution Analysis of China's Regional Economy with Weighted Spatial Markov Chain Approach. *China Industrial Economics.* (5): 31-43.

VERGINER L, RICCABONI M. 2020. Cities and countries in the global scientist mobility network[J]. *Applied Network Science.* 5(1) : 1-16.

Wang, W. J. 2013. *Study on the role and convergence of human capital on regional economic growth.* Changchun: Northeast Normal University (Doctoral dissertation).

Wang, Q. 2014. *An empirical study on human capital and economic growth in Northeast China.* Changchun: Jilin University (Doctoral dissertation).

Xia, C. Y., & Zhang, X. 2021. *In-depth Analysis Report on the Comparison and Benefit Evaluation of Talent Policies in 31 Provinces and Cities in China in 2021.* Forward Publishing.

Xu, Y., & Li, A. 2020. The relationship between innovative human capital and interprovincial economic growth based on panel data model and spatial econometrics. *Journal of computational and applied mathematics.* 365:112381.

Yan, L. G., & Zeng, S. M. 2020. Why Eastern industries have difficulty moving to the Midwest: An explanation based on spatial differences in human capital. *Economic Geography.* 40(1): 125-131.

Zhang, K., & Huang, L. Y. 2020. A study on the spatio-temporal evolution characteristics of human capital structure in China. *Research in Quantitative Economics and Technology Economics.* 37(12): 66-88.

Zhou, M., Li, Y. N., Yao, X., & Lu, Y. 2019. Human Capital Accumulation and Urban Manufacturing Export Upgrading in China: Evidence from Higher Education Expansion.

Management World. (5): 64-77.

Appendix

1. Distance Weight Matrix

This paper is using distance weight matrix in calculating Moran's I and Markov Chain results. There are two main types of weight matrix—contiguity matrix and distance weight matrix. In the dataset, China has Hainan Province as an isolated island. I used distance weight matrix since it is not appropriate to use contiguity matrix with an isolated part in the dataset.

Distance can be calculated by:

$$d_{ij} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \tag{5}$$

for two points i and j, with respective coordinates $(x_i,\ y_i)$ and $(x_j,\ y_j)$.

2. Markov Steady State Distribution and Passage Time

Table 6 Classic Markov Steady State Distribution

Poor	Lower	Middle	Upper	Rich
0.223	0.195	0.179	0.185	0.218

Note: Results are calculated based on the definition of Markov steady state distribution: $\pi P = \pi$, where π is the steady state distribution, P is the transition probability matrix.

Source: Author's calculation using data from China Center for Human Capital and Labor Market Research (CHLR)

Table 7 Classic Markov Passage Time

	Poor	Lower	Middle	Upper	Rich

Poor	4.483	17.429	35.005	73.417	215.917
Lower	60.7	5.135	17.577	55.989	198.489
Middle	85.192	24.492	5.581	38.412	180.912
Upper	111.122	50.422	25.929	5.415	142.5
Rich	150.922	90.222	65.729	39.8	4.580

Note: Results are calculated based on the definition of mean first passage time from state i to state j (denoted by m_{ij}).

$$m_{ij} = p_{ij} \cdot 1 + \sum_{k \neq j} p_{ik}\left(1 + m_{kj}\right) = 1 + \sum_{k \neq j} p_{ik} m_{kj}$$

where p is the transition probability.

Source: Author's calculation using data from China Center for Human Capital and Labor Market Research (CHLR)

Table 6 shows the probabilities of the steady state distribution. It represents the proportion of the five levels in the steady state in China. If the distribution at time *t* is shown above, then any distribution afterwards is the same distribution. In addition, the Markov Chain can tell us the time required for traversal, that is, the time it takes to go from one state to another.

3. Spatial Markov Steady State Distribution

Table 8 Spatial Markov Steady State Distribution

	Poor	Lower	Middle	Upper	Rich
Poor Neighbors	0	0	0	0	1
Lower Neighbors	0.198	0.110	0.274	0.133	0.285

Middle Neighbors	0.259	0.171	0.208	0.236	0.125
Upper Neighbors	0	0	0	0.667	0.333
Rich Neighbors	0	0	0	0	1

Spatial Markov Chain steady state distribution results are shown in Table 8. There are five distributions that are corresponding five neighbor conditions. Classic Markov Chain steady state distribution in Table 7 shows the five classes are about 20%, it is almost equally distributed. Poor and Rich classes are slightly higher than other classes. It is consistent with the twin-peak distribution dynamics in Quah (1996). Quah (1996) found that distribution will polarize into twin peaks of rich and poor. We cannot find the same results in spatial Markov Chain distributions. All the conditions except Lower Neighbors condition show that there is only one peak distribution. The only twin-peak distribution is in Lower Neighbors condition, but the peaks are on the "Middle" and "Rich".

Poor and Rich Neighbors conditions show that all the regions will be in "Rich" class, no matter whether their neighbors are poor or rich. It is an unexpected result and requires further research to explore the potential reasons.