Spatial pattern of artificial intelligence and its influence on labor market in China

Lingyue Tian¹, Dangchen Sui¹

¹International Business School of Shaanxi Normal University, China

Abstract. With the rapid development of artificial intelligence, the labor market is constantly impacted. In order to explore the spatial relationship of AI development in different regions and the impact of technological progress on labor market, spatial autocorrelation test and individual time fixed effect model were used. The results show that there is a significant positive spatial dependence on the development level of artificial intelligence in China’s provinces. At present, the development of artificial intelligence leads to employment polarization in the labor market of China. Therefore, promoting the balanced development of regional artificial intelligence and encouraging the horizontal transfer of labor force will alleviate the negative impact of artificial intelligence.

1. Introduction
Artificial intelligence is a strategic technology leading a new round of scientific and technological revolution and industrial transformation (Guo, 2019). It is urgent to accelerate the development of a new generation of artificial intelligence and seize the opportunity of a new round of scientific and technological revolution and industrial change. As a general technology to promote the fourth industrial revolution, artificial intelligence has the characteristics of infrastructure spillover (Brynjolfsson et al., 2018; Agrawal et al., 2019). The research shows that the labor force replaced by artificial intelligence is first concentrated in the low skilled labor group engaged in mechanical, simple and manual work, and the group is faced with the risk of being replaced by machines, job adjustment, training and reemployment (Acemoglu & Restrepo, 2018). The impact of artificial intelligence on the labor market can be divided into substitution effect and creation effect. Based on this, scholars use empirical analysis method to study the impact of artificial intelligence on the total amount of labor employment, but the conclusion is controversial (Acemoglu & Restrepo, 2017; Hoedemakers, 2017; Bessen, 2017). According to the theory of Routine-Biased Technological Changes (RBTC), the total output is composed of a series of tasks. Technological progress will not directly replace or promote the demand of a certain skilled labor force, but indirectly affect the labor force through changing the task content in the production process and the completion degree of capital to the task [1-3]. Technological progress has brought about the essence of polarization in the labor market. With technological progress, cognitive jobs with high income and manual labor with low income have been increased, while middle-income jobs have been hollowed out. The research by Michaels et al. (2010) also shows this [4]. On the basis of previous studies, on the one hand, this paper discusses the spatial pattern of the development of artificial intelligence in China. On the other hand, this paper empirically analyzes the impact of AI development on labor structure.
2. Models and data

2.1. Spatial autocorrelation test
Generally speaking, the spatial autocorrelation test is used to judge whether the spatial dependence exists. By analyzing the statistical data related to the region, we can infer the dependence among regions and whether the correlation follows a certain spatial pattern. In this paper, Moran index is used to calculate the spatial dependence of provincial interval variables. The specific calculation formula is as follows:

\[ I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} \sum_{i=1}^{n} (x_i - \bar{x})^2} \]  

(1)

Among them, \( \bar{x} \) is the sample mean value, \( n \) is the number of spatial units in the study area, and \( W_{ij} \) is the spatial weight matrix used in the study. The assumption of the structure and intensity of spatial effects will affect the setting of the spatial weight matrix, thus affecting the results of autocorrelation calculation. The range of Moran'I is \([-1, 1]\). If \( 0 < I \leq 1 \), it is called spatial positive correlation, which means that the attribute value of a feature between adjacent regions presents the same trend \([5]\). When \(-1 \leq I < 0\), it is called spatial negative correlation, which indicates that the attribute value of a feature between adjacent regions presents an opposite trend. \( I = 0 \) means that the attribute value of a feature between adjacent regions presents random distribution. There is no spatial autocorrelation among the investigated variables.

2.2. Individual time fixed effect model

2.2.1. Hausman test
In the process of selecting regression model, first of all, we need to use hausman test to determine whether the following regression model is fixed effect or random model. The main idea of Hausman test is to find two estimators \( \theta \) and \( \theta' \), the original hypothesis \( H_0: \theta - \theta' = 0 \). When the original hypothesis is tenable, it means that the random effect model should be used in the subsequent regression model. Otherwise, the fixed effect model should be used.

2.2.2. F test
Furthermore, F statistics are constructed for different models to test whether the individual time fixed model is selected. The formula of F test is as follows:

\[ F = \frac{(SSE_R - SSE_U)}{(N + T - 2)} \frac{SSE_U}{(NT - N - T - k - 1)} \]  

(2)

Among them, \( SSE_R \) and \( SSE_U \) represent the sum of residual squares of the mixed estimation model and the time individual fixed effect model respectively. If the value of F is greater than the critical value of F statistic under the corresponding confidence level, the original hypothesis is rejected and the time individual fixed effect model needs to be established. According to the results, \( F_1 = 164.6 > F_{5\%}(29, 255) = 32.4 \), \( F_2 = 219 > F_{5\%}(29, 255) = 12.7 \), \( F_3 = 95.2 > F_{5\%}(29, 255) = 43.1 \). The calculation results all reject the original hypothesis. Therefore, it is necessary to establish an individual time fixed effect model.

2.2.3. Basic model
In addition to artificial intelligence, other variables also have an impact on the labor market, so this paper adds relevant control variables to improve the scientificity and credibility of the regression model and conclusions. The basic model is as follows:

\[ \ln(Y_{nit}) = \beta_0 + \beta_1 \ln(AI_{it}) + \gamma_i \ln(X_{it}) + \sigma_i + \lambda_t + \epsilon_{nit} \]  

(3)

Among them, \( Y_{nit} \) represents the proportion of low, medium and high skilled workers in Province \( i \) in year \( t \), \( \beta_0 \) represents intercept term, \( \gamma_i \) represents regression coefficient of control variable, \( \sigma_i \)
represents fixed effect at provincial level, $\lambda_t$ represents fixed effect at year level, and $\varepsilon_{it}$ is residual term.

2.3. Data

This paper selects the panel data of 30 provinces (except Hong Kong, Macao, Taiwan and Tibet) from 2008 to 2017 for empirical analysis. The main variables are as follows: ① Explained variable. In order to discuss the influence of technological progress on the employment of workers with different skills, the employed people are divided into low skill, medium skill and high skill according to their educational background, and they are regarded as the explanatory variables, which are recorded as low, middle and high respectively. ② Explain variables. In this paper, the development level of regional artificial intelligence is taken as the explanatory variable. On the basis of the practice of this kind of literature (Jeff and Michael, 2017), this paper selects “the proportion of fixed asset investment in information transmission, computer service and software industry in regional GDP” to measure the development level of artificial intelligence, and record it as $AI$. ③ Control variables. The level of economic development, government expenditure, foreign investment and urbanization level are selected as control variables, which are expressed as $pgdp$, $gov$, $str$, $fdi$ and $urb$ respectively.

In the process of model estimation, natural logarithm is adopted for all variables. In this way, the causality between variables will not be changed, and the non-linear problem to a certain extent will be overcome, so that the regression results of econometric model will be more stable and the quality of parameter estimation will be improved. The statistical description of each variable is shown in Table 1.

| Variable   | Obs | Mean  | Std. Dev. | Min   | Max   |
|------------|-----|-------|-----------|-------|-------|
| ln(Low)    | 300 | 0.601 | 0.064     | 0.305 | 0.678 |
| ln(Middle) | 300 | 0.096 | 0.048     | 0.021 | 0.254 |
| ln(High)   | 300 | 0.066 | 0.054     | 0.009 | 0.319 |
| ln(AI)     | 300 | 0.007 | 0.005     | 0.0004| 0.034 |
| ln(pgdp)   | 300 | 10.575| 0.512     | 9.196 | 11.768|
| ln(gov)    | 300 | 0.207 | 0.076     | 0.084 | 0.487 |
| ln(str)    | 300 | 0.328 | 0.063     | 0.112 | 0.426 |
| ln(fdi)    | 300 | 0.015 | 0.010     | 0.00009| 0.056 |
| ln(urb)    | 300 | 0.433 | 0.083     | 0.256 | 0.640 |

*Source: China Statistical Yearbook from 2009 to 2018.*

3. Results and discussion

3.1. Spatial characteristics and autocorrelation analysis

3.1.1. Analysis of spatial characteristics. Before the spatial empirical analysis, this paper first calculates the time mean value of the AI development level of 30 provinces and cities in China from 2008 to 2017, and uses ArcGIS software to make the spatial distribution map of the time mean (see Fig. 1), so as to explore whether there is a certain degree of spatial distribution characteristics in the development level of artificial intelligence among provinces.
The results show that the AI development level of 30 provinces and cities in China has a strong spatial agglomeration phenomenon. The overall spatial characteristics show high value and high value aggregation, low value and low value cluster [6-7], which indicates that there is a certain degree of spatial positive correlation between the development level of artificial intelligence in neighboring areas.

### 3.1.2. Spatial autocorrelation analysis

On the basis of the previous paper, in order to further comprehensively investigate the spatial correlation of regional AI development level, the Moran index of AI development level of each province in China from 2008 to 2017 is calculated by using the first-order spatial adjacency matrix, and the spatial correlation is tested accurately (see Table 2). The overall test results show that the overall Moran′I of artificial intelligence development level in China from 2008 to 2017 are positive and significant, and the Moran index of artificial intelligence development level is significant at least at the level of 1% in most years, which indicates that the development level of artificial intelligence among different regions in China has obvious spatial agglomeration characteristics and positive correlation.

### Table 2. The result of Moran test.

| Year | I    | Sd(I) | Z    | P   |
|------|------|-------|------|-----|
| 2008 | 0.084| 0.034 | 3.456| 0.000|
| 2009 | 0.069| 0.034 | 3.014| 0.001|
| 2010 | 0.081| 0.033 | 3.486| 0.000|
| 2011 | 0.042| 0.029 | 2.632| 0.004|
| 2012 | 0.059| 0.034 | 2.771| 0.003|
| 2013 | 0.038| 0.034 | 2.132| 0.017|
| 2014 | 0.055| 0.036 | 2.524| 0.006|
| 2015 | 0.062| 0.036 | 2.685| 0.004|
| 2016 | 0.022| 0.036 | 1.550| 0.061|
| 2017 | 0.004| 0.036 | 1.078| 0.141|

### 3.2. Results of individual time fixed effect model

According to the above test and analysis, combined with variable selection, the specific models expression are as follows:

\[
\ln(\text{Low}_i) = \beta_0 + \beta_1 \ln(\text{AI}_{i,t}) + \gamma_1 \ln(\text{pgdp}_{i,t}) + \gamma_2 \ln(\text{gov}_{i,t}) + \gamma_3 \ln(\text{str}_{i,t}) + \gamma_4 \ln(\text{fidi}_{i,t}) + \gamma_5 \ln(\text{urb}_{i,t}) + \sigma_i + \epsilon_i
\]  

\[
\ln(\text{Middle}_i) = \beta_0 + \beta_1 \ln(\text{AI}_{i,t}) + \gamma_1 \ln(\text{pgdp}_{i,t}) + \gamma_2 \ln(\text{gov}_{i,t}) + \gamma_3 \ln(\text{str}_{i,t}) + \gamma_4 \ln(\text{fidi}_{i,t}) + \gamma_5 \ln(\text{urb}_{i,t}) + \sigma_i + \epsilon_i
\]  

\[
\ln(\text{High}_i) = \beta_0 + \beta_1 \ln(\text{AI}_{i,t}) + \gamma_1 \ln(\text{pgdp}_{i,t}) + \gamma_2 \ln(\text{gov}_{i,t}) + \gamma_3 \ln(\text{str}_{i,t}) + \gamma_4 \ln(\text{fidi}_{i,t}) + \gamma_5 \ln(\text{urb}_{i,t}) + \sigma_i + \epsilon_i
\]

Taking OLS panel regression results (recorded as model 1, model 2 and model 3) as the control, this paper analyzes the impact of artificial intelligence development on the labor market according to the individual time fixed effect model (recorded as model 4, model 5 and Model 6). The parameters of all the above models were estimated by Stata 15 software. All variables were in the form of natural logarithm. The detailed results are shown in Table 3.
The results show that all the core explanatory variables are not significant in the OLS regression results, but the significance of the core explanatory variables is greatly improved after adding the regional fixed effect and time fixed effect. The regression parameters $\beta_i$ in Model 4, Model 5 and Model 6 were 0.630, -0.493 and -0.274 respectively [8]. First of all, the development of artificial intelligence has a significant “creative effect” on the low skilled workers, which leads to a significant increase in the proportion of low skilled workers in the labor market. Secondly, the development of artificial intelligence has a significant “substitution effect” on the medium skilled workers, resulting in the market share of the medium skilled workers being squeezed out [9-10]. Finally, the development of artificial intelligence has no significant impact on highly skilled personnel. On the whole, the improvement of the development level of artificial intelligence leads to the phenomenon of "employment polarization" in the labor market of China.

**Table 3.** Parameter estimation results.

| Variable | OLS | Individual time fixed effect |
|----------|-----|------------------------------|
|          | Model1 | Model2 | Model3 | Model4 | Model5 | Model6 |
| ln(AI)   | -0.174  | 0.485  | -0.201  | 0.630** | -0.493* | -0.274  |
|          | (0.579) | (0.594) | (0.307) | *       | *       | (0.178) |
| ln(pgdp) | -0.107* | 0.070** | 0.079** | 0.030*  | -0.030* | -0.018  |
|          | (0.031) | (0.021) | (0.018) | (0.16)  | *       | (0.013) |
| ln(gov)  | -0.151  | 0.100  | 0.124** | 0.005   | -0.027  | -0.008  |
|          | (0.092) | (0.097) | (0.046) | (0.062) | (0.054) | (0.050) |
| ln(str)  | 0.516** | -0.441* | -0.302* | -0.079  | 0.097** | -0.027  |
|          | (0.124) | (0.090) | (0.072) | (0.055) | (0.048) | (0.044) |
| ln(fdi)  | 0.526   | -0.335 | -0.410  | -0.238  | 0.322** | 0.052   |
|          | (0.391) | (0.325) | (0.277) | (0.184) | (0.161) | (0.148) |
| ln(urb)  | 0.604   | 0.0126 | -0.537* | 0.707** | -0.137* | -0.515* |
|          | (0.400) | (0.276) | *       | (0.082) | *       | *       |
| Constant | 1.329** | -0.523* | -0.460* | 0.095** | 0.377** | 0.428** |
|          | (0.141) | (0.108) | (0.084) | (0.132) | (0.116) | (0.106) |
| Hausman  | 60.31*** | 35.18*** | 126.61*** | 164.62*** | 219.04*** | 95.18*** |
| F-value  | *       | *       | *       | *       | *       | *       |
| $R^2$    | 0.786   | 0.769   | 0.780   | 0.906   | 0.928   | 0.848   |

*Note: *, **, and *** are significant at 10%, 5% and 1% levels respectively.*

4. Conclusions

Based on the above conclusions, this paper puts forward the following policy recommendations:

First, promote the balanced development of artificial intelligence in different industries and regions. At present, the application of artificial intelligence is more concentrated in industrial manufacturing and relatively rich areas. We should promote the application of artificial intelligence in the primary and tertiary industries and underdeveloped areas.

Second, artificial intelligence is mainly used to replace the traditional manufacturing and some service industries. It is less likely that the management of others, unpredictable physical labor, interaction with stakeholders and application of professional knowledge will be replaced. Therefore, the horizontal transfer of labor force will be an important measure to solve this problem.

Third, pay attention to the education and training of artificial intelligence, and provide the human capital needed for intelligent production. The impact of artificial intelligence on employment will become controllable (Yang Weiguo et al., 2018). Through education and training to improve the skills of workers, in order to achieve extensive and effective human-computer cooperation.
Fourth, in view of the structural unemployment problem that may be brought about by the development of artificial intelligence, the government should use big data technology to detect the high-frequency data of employment dynamic changes in different regions, different groups and different posts, formulate unemployment support policies, improve the skills of workers, and solve the social problems caused by structural unemployment.

References

[1] Guo Kainming. 2019. Artificial Intelligence, Structural Transformation and Labor Share[J]. Management World. 2019, 35(7): 60-77+202-203. (In Chinese)

[2] Brynjolfsson, E., Mitchell, T., Rock, D., What Can Machines Learn and What Does It Mean for Occupations and Industries[R]. AEA Papers and Proceedings Working Paper, 2018.

[3] Agrawal A. K., J. S. Gans and A. Goldfarb, Economic Policy for Artificial Intelligence[R], Working Paper, National Bureau of Economic Research, 2019.

[4] Acemoglu D, Restrepo P. Artificial Intelligence, automation and work[R]. Working Paper, National Bureau of Economic Research, 2018.

[5] Acemoglu D, Restrepo P. Robots and Jobs: Evidence from US Labor Markets[J]. SSRN Electronic Journal, 2017.

[6] HOEDEMAKERS L. The changing nature of employment: how technological progress and robotics shape the future of work[EB/OL]. (2017-06-21) http: // lup. lub. lu. se / luur/ download? func = downloadFile&recordOId =8916714&fileOId = 8916715.

[7] Bessen J E. AI and Jobs: The Role of Demand[J]. SSRN Electronic Journal, 2017.

[8] Michaels G, Natraj A, Van Reenen J. Has ICT polarized skill demand? Evidence from eleven countries over 25 years[J]. CEPR Discussion Papers, 2010, 96(1), 60-77.

[9] Borland J, Coelli M. Are Robots Taking Our Jobs?[J]. Australian Economic Review, 2017, 50(4): 377-397.

[10] Yang Weiguo, Qiu Zitong, Wu Qingjun. A Literature Review on the Impact of Artificial Intelligence on Employment[J]. Chinese Journal of Population Science, 2018(5): 109-119+128.