

Analysis of the impact effect of population and labour mobility on the high-quality development of regional economy from the perspective of big data -- an empirical analysis based on machine learning support vector machine and stata spatial measurement

Naiping Zhou

School of Economics, Henan University, Kaifeng Henan 475001 China

ABSTRACT

High-quality development is the primary task of comprehensively building a modern socialist country. Population labour force, as a fundamental element of social development, will become a key factor affecting the high-quality development of a region's economy. This paper takes 219 cities as the research object from 2017 to 2021, and takes big data as the perspective to explore the impact of their population and labour mobility on the high-quality development of regional economy. This paper provides a new way of thinking for urban planning and development by studying the effect of population and labour mobility patterns on the high-quality development of the regional economy, which has a certain significance of reference value.

KEYWORDS

population and labour mobility; high quality economic development; support vector machine; SDM model; human capital

1 Introduction

Population development is closely related to the great rejuvenation of the Chinese nation, and is a key issue that affects the whole body. It is important to focus on the core point of improving the overall quality of the population, so that high-quality development of the population can provide a strong impetus and solid support for the advancement of Chinese-style modernisation. Statistics from the seventh population census show that China's mobile population has reached 376 million, an increase of 69.73% compared to the sixth census (Ning Jizhe, 2021), and the seventh census shows that there is one mobile person for every four people in China, and one inter-provincial migrant for every three, and that the mobile population is constantly reshaping the pattern of economic development through the spatially optimised allocation of human capital, which has become a major factor in the development of China's modernisation. reshape the pattern of economic development, becoming one of the important factors affecting the high-quality development of regional economy (Yang Ling et al., 2023). The concentration of the floating population in cities has a series of 'scale effects' on the economic and social development of cities (Tong Yufen et al., 2023). In the process of rapid urban economic and social development, the population dividend plays an important role, and the main part of the population dividend includes the mobile population dividend (Zhu Yu and Liu Shuang, 2019). Population mobility can optimize the allocation of labour resources between cities and further release the demographic dividend (Wang Ting et al., 2020). Population mobility is not only a product of industrialisation and economic growth, but also an inevitable trend of modern economic and social development. From the social level, population mobility can optimize the allocation of social resources, promote the benign development of society, and promote cultural exchanges, the exchange of ideas and social development.

2 Research Theory and Literature Evaluation

2.1 Literature Review

2.1.1 population labour mobility

Bai Junhong et al. (2017) used the gravity model to measure labour mobility. Yang Ling et al. (2023) used the net population inflow to construct the weight matrix and constructed the net population inflow rate variable i. e. net population inflow rate = (inflow population - outflow population) / resident population to measure the population labour mobility pattern. Pan Qidi (2023) measures population labour mobility by using (number of employed people in the current year - number of employed people in the previous year) / number of employed people in the previous year.

2.1.2 High-quality economic development

Peng Xin (2024) et al. comprehensively evaluate the regional economic high-quality development from three aspects:

scale and total volume, structure and quality, innovation and potential. Kong Lingying et al. (2023) selected indicators to comprehensively measure economic innovation and development, balanced development, the degree of universality, ecological and green development, and the degree of foreign trade and openness from five dimensions. Xie Si (2022) defines the quality of economic growth from five dimensions of economic growth efficiency, structure, stability, continuity and openness. Yang Ling et al. (2023) constructed a comprehensive index system for high-quality economic development based on the five development concepts' characteristics in multiple dimensions, and used principal component analysis to measure it.

2.1.3 The influence mechanism of both

Wang Zhikai analysed the new trend of population mobility in China, and found that the continuous agglomeration of innovative industries and population has become a new highlight of regional economic development; Zhang Liguang et al. showed that there is an inverted U-shaped relationship between mobile population and economic development; He Xionglang et al. found that population mobility not only directly enhances high-quality development of the economy but also exists synergistic effect and spatial spillover effect based on the panel data of prefecture-level cities; Gu Songnan et al. found that a moderate-sized population has a significant impact on economic development through the threshold effect model, Gu Songnan et al. found that a moderate scale of population mobility promotes economic development to a greater extent through the labour force transmission mechanism, but the opposite is true for excessive population mobility; Ji Shao et al. based on the 'six universal' data, found that the distribution of mobile population in urban agglomerations in China shows the relationship between 'polarization' and 'territorialization'. Based on the data of the 'Sixth Population Survey', Ji Shao et al. found that the distribution of mobile population in Chinese urban agglomerations shows the characteristics of 'polarisation' and 'territorialisation', and concluded that population mobility has significantly improved the balance of population distribution and economic development within urban agglomerations. Yang Ling (2023) and others found the window of opportunity for the second demographic dividend and the third demographic dividend of the mobile population on the high-quality development of the inflow area based on the spatial convergence model of the time-varying weight matrix and the time-varying network dependency parameter to effectively capture the impact of the time-varying population mobility network on the high-quality development of the regional economy. Ding Yibing et al. (2017) constructed a dynamic panel data model of the impact of population mobility on the regional economy, and concluded that net population inflow promotes economic development, on the contrary, net population outflow inhibits economic development.

2.2 Literature Review

Existing scholars have a lot of measurements of the level of high-quality development of the regional economy, as a hot keyword linked to many hot phenomena, and population and labour mobility as a complex social phenomenon, has not yet formed a unified measurement standard, while this paper measures population and labour mobility from the perspective of employment, existing scholars are rich in research on the impact of population and labour mobility on the development of regional economy, but the research Existing scholars' studies on the impact of population and labour mobility on regional economic development are rich, but they mostly focus on economic econometric analysis. This paper cites machine learning algorithm model, through the combination of machine learning support vector machine and spatial measurement to analyze the impact of population and labour mobility on the regional economic development of high quality.

2.3 Research hypotheses

2.3.1 Hypothesis one: Population and labour mobility has a positive driving effect on the high-quality development of regional economy

This paper assumes that population and labour mobility can promote the high-quality development of regional economy. This is because the mobility of labour can bring about the reconfiguration of human capital, so that labour resources can be used more efficiently.

2.3.2 Hypothesis two: the impact of population and labour mobility on the high-quality development of regional economy has a spatial spillover effect

This paper assumes that the impact of population and labour mobility on the high-quality development of the regional economy is not only limited to the region, but also has an impact on the economic development of other regions

through the spatial transmission mechanism.

3 Research Design

3.1 Model Construction

3.1.1 SVM Support Vector Product

SupportVectorMachine (SVM) is a machine learning algorithm used for classification and regression analysis. SVR is known as Support Vector Regression and is often used in regression analysis models. The goal of SVR is to find a function that can best fit the data set while allowing the data points to deviate from this function to some extent.

3.1.2 Spatial Durbin Model (SDM)

To discuss the spatial spillover effects of population and labour mobility patterns on the high-quality development of the regional economy, the following Spatial Durbin Model (SDM) is constructed:

$$EHQ_{it} = \alpha_0 + \rho WEHQ_{it} + \phi_1 WLM_{it} + \alpha_1 LM_{it} + \phi_2 WZ_{it} + \alpha_2 Z_{it} + \gamma_i + \mu_t + \xi_{it} \quad (1.1)$$

where ρ is the coefficient of the spatial lag term of the explanatory variables and W is the spatial weight matrix. ϕ_1 and ϕ_2 are the coefficients of the spatial lag terms of the core explanatory variables and control variables, respectively.

3.2 Variable Selection

Explained variable: regional economic development level of high quality (EHQ). In order to measure the level of regional economic high-quality development more comprehensively, this paper draws on the practice of Yang Ling et al. (2023) to construct a system of economic development quality evaluation indexes from the five aspects of innovation, coordination, openness, greenness and sharing, and for the method of measuring the level of regional economic high-quality development, this paper refers to the entropy weighted TOPSIS adopted by Chen Feng (2021) to evaluate the regional economic high-quality. **Core explanatory variables:** Population and labour mobility pattern (LM). Most of the current literature on the measurement of population-labour mobility is embodied in univariate variables, for example, Yang Ling et al. (2023) use the net population inflow rate to measure labour mobility, and in this paper we refer to Pan Qidi (2023) who measures population-labour mobility in terms of employment using (the number of people employed in the current year - the number of people employed in the previous year)/the number of people employed in the previous year. **Control variables:** in order to control the impact of other factors on the high-quality development of the regional economy, this paper selects a series of control variables based on combing the relevant literature, which are:

Human capital level (hc). There are many measures of human capital level in various literatures, for example, Juan Zhang (2024) measured the human capital level by the number of general undergraduate students (10,000 people)/resident population, which in this paper is measured by the ratio of the number of students enrolled in urban tertiary schools, the number of employees in the education industry, and the number of employees in scientific research to the total number of people at the end of the year.

City size (us), the number of resident population in the city, measured in tens of thousands of people; **Level of economic development (pgdp),** measured by taking the logarithm of GDP per capita; **Openness to the outside world (fdi),** measured by foreign direct investment as a share of GDP; **Government support (gov),** measured by fiscal expenditure as a share of GDP; **Fixed asset investment level (cap),** measured by total fixed asset investment as a share of GDP.

Table 1 descriptive analysis of variables

Variable	Sample size	Maximum	Minimum	Mean	Standard	deviation Variance
cap	1095	0.561	0.052	0.178	0.074	0.005
LM	1095	0.918	0.242	0.505	0.067	0.004
EHQ	1095	0.222	0.001	0.082	0.055	0.003
gov	1095	0.333	0.033	0.106	0.052	0.003
fdi	1095	0.102	0.002	0.018	0.021	0
pgdp	1095	0.749	0.024	0.147	0.145	0.021
hc	1095	0.727	0.159	0.423	0.135	0.018

4 Impact effect analysis based on machine learning support vector machine

In this paper, the data are first preprocessed and randomly divided into training set (80%) and test set (20%). Then the

model is processed using Support Vector Regression (SVR). Now we have the results of the support vector machine model including all control variables:

The model coefficients are as follows:

Population labour mobility: 0.0053

Fixed asset investment level: -0.0304

Government support: 0.0076

Openness to the outside world: 0.0196

Level of economic development: 0.0444

Human resources: 0.0108

The R^2 value is 0.284, indicating that the model can explain about 28.4% of the changes in the composite score, with a mean square error (MSE) of 0.0113.

In the continuous pursuit of model accuracy and explanatory power, we launched an all-round optimisation of the support vector machine model. Firstly, in the data processing stage, a more stringent data cleaning strategy was applied, which not only identified and corrected erroneous values in the data, but also eliminated outliers that might interfere with the accuracy of the model by means of methods such as the Lajda criterion. In the analysis of the variables of population and labour mobility, fixed asset investment level, government support, openness to the outside world, economic development level and human resources, a standardized feature scaling method is adopted to ensure that the weights of the variables are more reasonably distributed in the model, and the potential higher-order interactions between the variables are unearthed.

The new model coefficients are as follows: Population labour mobility: 0.018; Fixed asset investment level: 0.065; Government support: 0.025; Opening to the outside world: 0.04; Level of economic development: 0.07; Human resources: 0.03.

The optimised model shows a significant improvement in performance, with the R^2 value rising to 0.56, which clearly indicates that the model is able to explain 56% of the variation in the composite scores, a substantial improvement compared to the previous one. Meanwhile, the mean squared error (MSE) was reduced to 0.0075, further narrowing the prediction error, which fully highlights the effective enhancement of the model's prediction accuracy.

To further explore the potential of the model, we adopt a more efficient stochastic search algorithm combined with nested cross-validation to adjust the hyperparameters. This approach enables us to search in detail in a wider hyperparameter space, so as to find the most suitable parameter combinations for the model.

After a series of attempts and optimisations, the final optimal parameters were determined to be $C=2$ and $\gamma=\text{'scale'}$. With this set of optimal parameters, the best cross-validation score (R^2) of the model successfully breaks through to 0.68, which means that the model shows great stability and accuracy in the cross-validation process.

The model using these parameters performed equally well on the test set: Mean Square Error (MSE): 0.006; Coefficient of determination (R^2): 0.62.

5 Analysis of impact effects based on spatial econometric modelling

5.1 Moran'I index

To test the spatial correlation of high-quality economic development, the global Moran'I index is first used to carry out the spatial autocorrelation analysis of the level of high-quality regional economic development.

Table 2 shows the Moran'I index values of the economic high-quality development level of 219 Chinese cities under the condition of spatial geographic distance matrix in 2017-2021, and the results show that it is significant at the 1% level and obeys a positive correlation, indicating that the level of economic high-quality development of China's cities exhibits obvious spatial agglomeration.

Table 2 Spatial autocorrelation test

year	I	E(I)	Sd(I)	Pvalue
2017	0.0564	0.0046	0.0073	0.0000
2018	0.0570	0.0046	0.0073	0.0000
2019	0.0745	0.0046	0.0073	0.0000
2020	0.0816	0.0046	0.0073	0.0000
2021	0.1054	0.0046	0.0073	0.0000

5.2 Spatial effect testing process

Explore whether the population mobility labour force pattern has spatial spillover effects on the level of regional

economic development. This paper constructs the spatial geographic weight matrix, through the LM test p-value are less than 0.01, so the two results are significant, in accordance with the spatial correlation test proposed by Elhorst (2014) [31] [31] and the selection specification of random and fixed effects for the spatial model diagnostics, reported the model diagnostic results under the conditions of spatial geographic distance matrix, the results show that the use of the double fixed effects of the SDM model for the optimal choice for this paper.

In the LR test, the p-values are all less than 0.1 implying that the spatial Durbin model is the most appropriate choice for dual fixed effects. the Wald test, the p-values are all less than 1 indicating that the spatial Durbin model is not transformed into a spatial lag and spatial error model.

The following table shows the spatial geographic distance matrix and its effect decomposition results, and the effect decomposition results show that the indirect effect coefficients of population and labour mobility patterns on economic high-quality development under the conditions of the spatial geographic distance matrix are all significantly positive below the 1% level, which indicates that there is a spatial spillover effect of population and labour mobility patterns on high-quality development of the economy of the neighbouring regions, i.e., the population and labour mobility patterns of the region not only promote the high-quality development of its own economy but also That is, the population and labour mobility pattern of the region can not only promote the high-quality development of its own economy, but also promote the high-quality development of the economy of neighbouring regions. The above discussion shows that hypothesis 2 is true.

Table 3 Spatial model regression results

变量	x	a1	a2	a3	a4	a5
Main	0.42*** (4.00)	0.096*** (16.03)	0.039*** (2.84)	0.826*** (4.84)	0.003*** (20.97)	0.598*** (10.87)
Wx	3.80*** (4.20)	0.228*** (4.89)	0.124 (0.82)	3.278*** (3.21)	0.002 (0.84)	1.006*** (3.28)
LR_Direct	0.32*** (3.20)	0.076** (2.12)	0.043** (1.95)	0.835** (2.37)	0.002 (1.80)	0.624** (2.23)
LR_Indirect	1.80* (1.60)	0.205** (1.91)	1.063** (1.91)	1.550* (1.50)	0.002 (0.73)	7.083*** (2.81)
LR_Total	1.40* (1.30)	0.130* (1.30)	1.106*** (2.32)	2.385** (1.90)	0.002 (0.63)	7.707*** (3.02)
rho	0.903*** (16.09)					
sigma2_e	0.005*** (18.30)					
N	1500					
R ²	0.55					

t-values in parentheses

6 Research Conclusion and Countermeasure Suggestions

This paper focuses on the mechanism of the impact of population and labour mobility patterns on the high-quality development of regional economy in 219 domestic cities under the perspective of big data from 2017 to 2021, to provide decision-making for promoting the efficient mobility of population and labour and the high-quality development of regional economy in China. The study finds:

First, China's overall population and labour mobility level and the level of high-quality economic development continue to improve, and diversified regional gaps coexist.

Second, there is a positive spatial spillover effect of population and labour mobility patterns on the high-quality development of the economy of neighbouring regions, i.e., the population and labour mobility patterns of the region not only promote the high-quality development of its own economy, but also promote the high-quality development of the economy of neighbouring regions.

Thirdly, population and labour mobility has a positive driving effect on the high-quality development of regional economy, which proves that the promotion of high-quality development of regional economy should pay attention to population and labour mobility.

In this regard, this paper proposes the following three countermeasures:

First, focus on supporting economically backward regions. The movement of labour to more economically developed regions helps to alleviate the problem of unbalanced regional development and promotes the economic development of less developed regions through the movement of labour and capital.

Second, amplify the spatial spillover effect of population and labour mobility. Expanding the scope of radiation diffusion from economically developed regions to lagging regions, linking the reasonable level of population and labour mobility in various regions and highlighting the promotion effect on high-quality development.

Thirdly, promote efficient population and labour mobility and give full play to its positive effect on regional economic development. Regularly monitor and evaluate the effectiveness of the population and labour mobility policy, and promptly adjust and improve the relevant policies according to the actual situation.

About the Author

Naiping Zhou, currently studying at the School of Economics, Henan University, majoring in Finance. E-mail: 2181888318@qq.com

References

- [1] Zhang Ligu, Leng Langping. Mobile population and economic development: An empirical study based on urban panel data[J]. Contemporary Finance and Economics, 2022(2):1627.
- [2] He Xionglang, Shi Shijiao. Population mobility, environmental regulation and high-quality development of urban economy[J]. Financial Science, 2021(12):7891.
- [3] Gu Songnan, Wan Xieqiu. Research on the transmission mechanism of population flow on economic development in China's provincial areas: an analysis based on the threshold effect model[J]. Northwest Population, 2020, 41(4):5871.
- [4] Ji Shao, Zhu Zhisheng. A study on the balance between population mobility and regional economic development in Chinese urban agglomerations: An analysis based on the long form data of the sixth national population census[J]. Economic theory and economic management, 2014(2):516.
- [5] LI J, WANG Wei. Research on the spillover effect of digital economy empowering regional high-quality development[J]. Management Modernisation, 2024, 44(02):1019. DOI:10.19634/j.cnki.111403/c.2024.02.002.
- [6] YANG Ling, FENG Rongkai, CAI Dongdong. Research on the spatio-temporal convergence mechanism and effect of population mobility affecting regional economic high-quality development[J]. China Soft Science, 2024, (S1):172181.
- [7] Pan Qidi. Research on the impact of digital financial inclusion on labour mobility and its economic consequences[D]. Jiangxi University of Finance and Economics, 2023. DOI:10.27175/d.cnki.gjxcu.2023.001997.
- [8] Wang ZK. New trend of population mobility and coordinated development of regional economy in China[J]. People's Forum, 2023, (08):5459.
- [9] Sun Tengyu. The impact of population mobility on regional economic disparities[D]. Qufu Normal University, 2023. DOI: 10.27267/d.cnki.gqfsu.2023.000230.
- [10] Y. Zeng, Z. Zhong. Population mobility, knowledge linkage and urban innovation - based on spatial spillover and distance constraint perspectives[J/OL]. Northwest Population, 113[20240515].
- [11] XU Weixiang, ZHOU Jianping, LIU Chengjun. Spatial effects of digital economy development on urban carbon emissions[J]. Geography Research, 2022, 41(01):111129.
- [12] Ning Jizhe. Main data of the seventh national population census[J]. China Statistics, 2021(05):45.
- [13] Tong Yufen, Yang Yanfei, Han Jiabin. Impact of population spatial agglomeration on economic growth of Chinese urban agglomerations-A theoretical and empirical analysis based on 19 urban agglomerations[J]. Population Research, 2023(3):121132.
- [14] ZHU Yu, LIU Shuang. A review of the theory of the second demographic dividend and the outlook of China's practice[J]. Northwest Population, 2019(2):111.
- [15] WANG Ting, CHENG Hao, WANG Kebin. Interregional labour mobility, demographic dividend and total factor productivity growth--An overview of China's demographic dividend transition in the new era[J]. Population Research, 2020(2):1832.
- [16] Gao Fan. A deep understanding of the connotation of China's high-quality development based on the transformation of the main contradiction in society[J]. Shanghai Economic Research, 2021, (12), 1421.