

## Research Article

# Spatial Patterns and Driving Factors of Interprovincial Rural Migration: A Case Study of Henan Province, China

Ge Gao\* 

School of Languages and Communication Studies, Beijing Jiaotong University, Beijing, China

## Abstract

The large-scale mobility of the rural population plays a significant role in shaping China's economic and social development, highlighting the need for deeper research into its patterns and influencing factors. Based on data from China's Seventh National Census, this study explores the spatial patterns of interprovincial migration among the rural population, taking Henan Province as a case study. Henan Province is selected because of its agricultural significance and its role as a major source of rural migrant outflow in China. By carrying out statistical analysis, spatial analysis and using a multivariate linear regression model, the study finds that the cross-provincial rural migrants from Henan Province are scattered throughout the country, but they are relatively concentrated in the eastern and southern coastal areas. Between the Sixth and Seventh National Censuses, the overall number of migrants increased significantly, while the spatial migration patterns remained largely unchanged. Meanwhile, it is found that the distance between the outflow and the inflow areas has a complex impact on interprovincial rural migration. At present, economic factors outweigh geographical distance in importance, especially given the advancements in China's transportation infrastructure and communication technologies. Regional language differences have little impact on the rural population's migration across the country. More importantly, the factor of migration stock primarily affects the interprovincial migration of the rural population. Migration inertia functions as a key mechanism in determining migration destination choices. Employment opportunities and wage income levels are also fundamental factors guiding rural migrants' destination choices. These findings provide insights into understanding rural migration in urbanization and contribute to regional development policies.

## Keywords

Rural Population, Interprovincial Migration, Spatial Pattern, Driving Factors, Henan Province

## 1. Introduction

The large-scale rural population migration within China has an important impact on its economic and social development. According to data from the Seventh National Census, China's rural migrant population is 346 million, accounting for 24.54% of the country's total population [1]. Migrant workers constitute the main body of the migrant rural population. In 2022, the total number of migrant workers in the country reached 296 million,

including 124 million local migrant workers and 172 million migrant workers with cross-township mobility (including 71 million who moved across provinces) [2]. At present, China is in a stage of high-speed urbanization. The rural population flowing into cities (rural-urban migrant population) accounts for 75.19% of the total urban inflow [3], and the rural migrant population has become the main force of urbanization. The

\*Corresponding author: gog1209@163.com (Ge Gao)



rural population represents only part of China's overall migrant population. In fact, the total migrant population in China reaches 376 million [4], and the stock of immigrants from China to all parts of the world has also reached more than 10 million [5]. Migration is an important socio-economic phenomenon in China. This large rural migration has not only effectively promoted the development of the real estate industry and urban construction in the destination cities but also profoundly affected the rural and agricultural development of migrant-sending areas. Therefore, scientifically understanding the spatial law of rural population migration is highly practical.

In recent years, Chinese scholars have researched rural population migration from the perspective of geography, mainly focusing on the direction and causes of migration, multistage migration, return migration, and the impacts of migration. In 1978, China began to implement a family co-production contracting responsibility system, the agricultural productivity was greatly increased, and a large number of surplus rural labor appeared. In pursuit of higher incomes, farmers have initiated large-scale migration and mobility. Since the 1990s, the migration of the rural population in China has become increasingly concentrated. Guangdong, Shanghai, Jiangsu and Zhejiang in the eastern regions have become the main inflow locations for the interprovincial migration (referring to the migration from one provincial-level administrative region to another) of rural labor, whereas Anhui, Hunan, Jiangxi, and Henan in the central regions and Sichuan in the west are the main outflow regions [6, 7]. Job opportunities, determined by differences in the level of economic development, are the root driver of rural population migration, and the network of relationships plays a central role in the choice of migration destination [8, 9]. In the structure of the migration space, the rural population migration in the eastern regions is mainly intraprovincial migration, whereas interprovincial migration is observed mainly in the central and western regions [6]. However, in both the east and the central and western regions, the outflow rate of the rural population at the county level has shown a deepening trend, and migration from rural to urban areas is the basic feature of the migration of migrant workers [10]. Moreover, migrant workers establish cultural and identity recognition through language, and individuals' adaptability to the dialects of destination regions also influences rural population mobility.

Due to various factors, particularly the household registration system, the migration of migrant workers is not stable, presenting the characteristics of multiple migrations and multistage migration. Changes in migrant work locations are very frequent, and the regional viscosity index of enterprises in the work location is relatively low [8]. Migrant workers frequently engage in secondary cross-regional migration, and regional mobility has become the fundamental form of their migration [11]. The choice of migration destination is related to chain mobility [8, 11]. With the advancement of industrial transfer and the development of the county economy in the central and western regions, the return of migrant workers to

their hometowns has become an important path for relocation under the new economic normal. Employment, entrepreneurship and home purchases after return have become important forces driving the local urbanization of migrant-sending areas [12]. The places of return are mainly urban areas [13] and can also be the villages, townships, and county governments at migrants' places of origin. Counties and villages have become the main locations where returnees start businesses [12]. Multidimensional aspects of distance and proximity have a significant effect on individuals' willingness to return to their hometowns [14]. At the same time, the establishment of cultural identity and status identity by migrant workers through language, as well as their adaptability to the language and place of immigration, will also affect the migration of rural population [15].

Researchers outside China delve deeper into rural–urban migration, its motivations and effects, with specific regional case studies as the main focus. These studies have focused mainly on developing countries because rural–urban migration is the main type of population migration. In contrast, relevant research in developed countries has focused on urban–rural migration, that is, migration from cities to rural areas. Among the drivers of rural–urban migration, economic factors are the most important [16]. Migration occurs mainly due to individuals' aim to obtain high wage income in cities [17]. The low income and instability associated with rural agriculture and the lack of employment opportunities are the main drivers of migration. Urban conveniences and superior infrastructure are also key factors in the rural–urban migration [18]. In long-term analyses of economic drivers, classic theories of population migration from rural to urban areas, such as the Lewis model and the Harris–Todaro model, have been developed. In recent years, involuntary rural–urban migration has attracted increased interest among scholars. The reasons for passive migration include social factors [19] but more importantly climate and environmental factors, such as drought and extreme weather events [20]. In addition, recent studies have considered the impact of personal characteristics, such as age, gender, education and risk awareness [21], and family characteristics, such as the family size and the dependency ratio, on rural–urban migration [22].

In terms of the impact of rural–urban migration, studies have focused mainly on the impact on economic and social development in rural areas, and only a few studies have mentioned the impact on urban areas [23]. The impacts on rural areas involve mainly the livelihoods of farmers, the left-behind population, social development, and agricultural land use. Both negative and positive impact analyses are conducted. The loss of the young population has led to a downward spiral in rural development [24], the lack of labor has left land and agricultural facilities in villages idle, and the increase in remittances has led to an increase in the use of fertilizers and pesticides and a decrease in crop diversity. The outmigration of parents has had adverse effects on the education and behavior of left-behind children [25]. However,

rural–urban migration has reduced the incidence of rural poverty [26], and rural immigrants have accumulated skills and knowledge in cities and gained an increased sense of responsibility with regard to their hometowns. Migration has pushed an increase in the intensity of adoption of agricultural technology [26], objectively promoting agricultural modernization.

In summary, though progress has been made in existing research, deficiencies remain. First, most analyses of the direction of interprovincial rural population migration are limited to the general regional level, and there still lacks a detailed characterization of the direction of specific regional migration flows, especially with the use of accurate census data. Second, in current research on the dynamics of rural population migration, analyses of the migration stock that plays a key role in the selection of migration destinations are still insufficient. Taking Henan Province as the case area, we analyze the spatial structure characteristics of the interprovincial migration of the rural population based on data from the Seventh National Census of China and quantitatively analyze migration causes to understand the laws of interprovincial migration of the rural population in typical agricultural areas. We thus provide a reference for the formulation of rural population migration policies and regional development policies.

## 2. Data and Methods

### 2.1. Selection of Henan Province as the Case

Henan Province was selected as a case area for research because of its typicality and representativeness as an agricultural and migrant-outflow province in China. China has 23 provinces, 5 autonomous regions, 4 municipalities and two special administrative regions, with a total of 34 provincial administrative regions. Henan Province is one of them and is located in the central region of China. According to statistics, the added value of Henan's primary industry reached 581.78 billion yuan (RMB, same below) in 2022, ranking the third in the country. It has long been a large agricultural province in China. Serving as an important agricultural production base in the country, it ranks at the forefront of the country regarding the output of main agricultural products. Among them, Henan's output of grain is 67.894 million tons, ranking the second in the country; its output of oil is 6.840 million tons, ranking the first in the country; and its output of meat is 6.600 million tons, ranking the third in the country. Moreover, the urbanization rate of Henan Province is relatively low, at only 57.07% in 2022, such that it ranks the sixth last in the country and has a significantly lower value than the national average of 65.22%. Therefore, its rural population is large, reaching 42.3805 million [27]. Correspondingly, the number of migrant workers is very large, reaching 31.8202 million in 2022 [28], accounting for approximately 10.76% of the country's migrant population [2]. In addition, Henan Province is located in the central plains and has convenient transportation; thus,

there are many destinations for migration. Migrants from Henan are distributed in all directions, including the eastern and southeast coastal regions, the Beijing–Tianjin region in the north, the western region, the southwest region and the Xinjiang Uygur Autonomous Region (henceforth Xinjiang). The migration of migrant workers in Henan Province has diverse and typical characteristics.

### 2.2. Data Source

The data on rural population migration in Henan Province are derived from the data of the Seventh National Census [1] and the Sixth National Census [29]. The Seventh National Census was conducted in 2020, and the long table data contain the population living in a place other than that of their hukou. It should be specified that hukou, a specialized Chinese term, means the household registration in China. The migration population in this paper refers to the population whose hukou is inconsistent with their current place of residence. The place unit here refers to the township. The rural migrant population refers to the migrant population from areas outside urban built-up areas. In this study, we define this group as the population whose hukou is associated with the neighborhood committee of a township or town or the village committee of a town. Since the long table is sampled at 10%, expanding the sampled data by 10 times yields the overall migration data. The wage level, GDP, urbanization rate difference, nonfarm employment and other data for the model variables come from the *China Statistical Yearbook* (2021) [30]. The straight distance between provincial capital cities comes from measurements based on Baidu maps, and the social security expenditure data come from the *China Fiscal Yearbook* of the same year.

### 2.3. Methods

#### 2.3.1. Factor Selection

According to the analysis based on the push–pull theory of population migration, population pressure theory, geographic neighbor effect theory, with reference to previous studies [31–33], six factors are selected for analysis: GDP, wage level (WL, the same below), urbanization rate difference (URD), per capita social security expenditure (PCSSE), number of primary and secondary schools (NPSS), and migrant stock (MS), and the linear distance factor is selected as the control variable for robustness analysis.

In theory, economically developed areas provide many jobs and high wages, offer relatively comprehensive social services and social security, and present a high level of urbanization. These areas are attractive to rural laborers, which yields a positive effect on migration inflows. Therefore, GDP, wage level, urbanization rate difference, and per capita social security expenditure are selected as the influencing factors for analysis. With the continuous migration of rural migrant workers in China, family migration becomes an important trend, and the issue of children's admission has become a

factor migrants consider. Therefore, the ease of school enrollment affects the destination of rural population migration. Because of this, the number of primary and secondary schools at the destination is selected for analysis. As an economic and social phenomenon, population migration has historical inheritance, namely, inertial effects and channel effects. Therefore, the amount of migrants flowing into a certain area in the early stage (migration stock from the Sixth National Census in 2010) should have a positive effect on current migration. Thus, migrant stock is also selected to analyze the direction of rural population migration.

According to the first law of geography, the closer the distance is between two places, the more likely migration is. Therefore, the geographic distance between two cities (taking the straight distance between provincial capital cities as the proxy variable) has a negative effect on rural population migration. Moreover, according to the theories of immigration integration and cultural integration, language use is an important tool for the integration of migrant populations into local society. The degree of language difference is an important factor in immigrants' choice of destination, and language differences are related to geographical distance. Therefore, the distance between two places is also an important indicator for characterizing language differences.

Among the above factors, in addition to the indicators with clear physical significance, the meanings of the other indicators are as follows. The wage level is the average salary of private urban units (unit: yuan/year). The private sector is chosen because most of the rural population works in the private sector. The urbanization rate difference refers to the urbanization rate of the destination area minus the urbanization rate of Henan Province. Social security expenditure is the sum of the final accounts of social security and employment expenditure and the budget expenditure of the social insurance fund (unit: 100 million yuan). Per capita social security expenditure is the social security expenditure divided by the total population (unit: yuan/person).

### 2.3.2. Statistical Analysis and Multivariate Linear Regression

Statistical analysis and spatial analysis are used to analyze the quantitative structure and spatial structure of the migrant population, and the multivariate linear regression is used for genesis analysis. We carried out spatial global autocorrelation analysis of the migration of the rural population in Henan Province using the Geoda1.20.0 software. According to the results, the global Moran index was 0.105, and the significance level was 0.113 (the pseudo p value generated by 999 permutations), which failed the test; thus, the classic OLS model was used for analysis.

The general form of a multivariate linear regression model is as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \mu \quad (1)$$

Where  $Y$  is the dependent variable,  $X_j (j = 1, 2, \dots, k)$  is the explanatory variable,  $\beta_j (j = 0, 1, 2, \dots, k)$  is the regression coefficient,  $k$  is the number of explanatory variables, and  $\mu$  is the random error term.

The performance and fit degree of the multivariate linear regression model are evaluated by the multiple determination coefficient  $R^2$  [34], the expression is as follows:

$$R^2 = \frac{SSR}{TSS} \quad (2)$$

Among them,  $SSR$  represents Sum of Squares for Regression, and  $TSS$  refers to Total Sum of Squares.  $R^2$  is a value between 0 and 1. A higher fit indicates the better performance of the model.

$F$  (F-statistical) is used to evaluate the overall fit of a regression model. It is obtained by comparing the ratio of the sum of squares between groups and the sum of squares within the group in the regression model. The expression is as follows:

$$F = \frac{ESS/k}{RSS/(n-k-1)} \quad (3)$$

Among them,  $ESS$  refers to Explained Sum of Squares,  $RSS$  is short for Residual Sum of Squares,  $n$  is the number of samples and  $k$  is the number of independent variables. When  $F > F_{0.05}$ , the explanatory ability of the model can pass the test.

In multivariate linear regression, collinear diagnosis is usually carried out using the variance Inflation Factor (VIF, Variance Inflation Factor). The calculation formula of VIF is as follows:

$$VIF = \frac{1}{1 - R_k^2} \quad (4)$$

Among them,  $R_k^2$  is the determination coefficient between the  $k$ -th independent variable  $x_k$  and the remaining independent variables. The value of VIF is greater than 1. The closer the VIF value is to 1, the lighter the multiple collinearity, and vice versa.

## 3. Results

### 3.1. Spatial Patterns of Rural Population Migration

#### 3.1.1. High Concentration and Relative Stability of Migration Destinations

The rural migrant population from Henan Province is scattered across provinces throughout the country, but it is concentrated mainly in Zhejiang, Guangdong, Jiangsu, Shanghai, Beijing, Xinjiang, and Fujian (Figure 1). According to the Seventh National Census data, the interprovincial rural

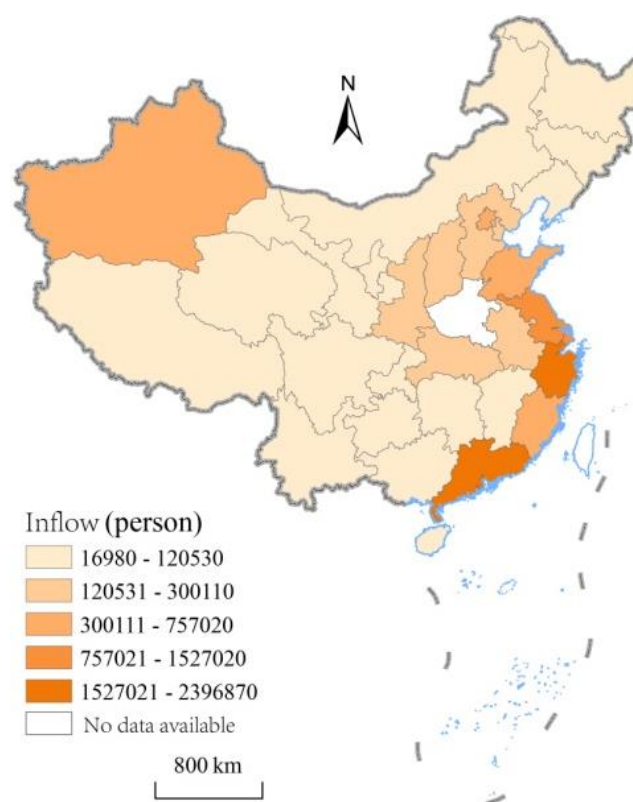
migrants from Henan Province are distributed in the other 30 provinces and regions of the country (there are no statistics for Taiwan, Hong Kong or Macau), but their number varies greatly, and migrants are concentrated mainly in the above-mentioned seven provinces and regions (accounting for 76.25% of the total). Among them, 2.3969 million people flowed into Zhejiang Province, 2.2301 million people flowed into Guangdong Province, and 1.5270 million people flowed into Jiangsu Province, accounting for 20.40%, 18.99% and 13.00%, respectively, of the total outflow from Henan Province, occupying 52.39% in total [1]. The above three provinces have become the main importers of the rural population of Henan Province.

In addition, Shanghai on the eastern coast and Fujian on the southeast coast present large inflows of the rural population of Henan Province, with migrant workers reaching 1.0606 million and 412,200 [1]. On the whole, the economically developed eastern and southeastern coastal areas have become the main inflow areas of rural migrants from Henan Province. The reason for this is the large number of jobs and the high wage income associated with the highly developed economies. For example, in 2020, Zhejiang Province ranked the fourth in GDP, the sixth in the number of employed people, and the sixth in the average salary of urban private units. Many rural migrants have moved from Henan Province to Beijing because of Beijing's status as the capital, its relatively developed economy, and the relatively short distance between the two provinces. Xinjiang is another main inflow area for the rural population in Henan because of its rich natural resources and special historical factors.

Except for a significant increase in total migration, the spatial structure of the destinations of rural migrants from Henan Province has not undergone substantial changes since the Sixth National Census (Figure 1 and Figure 2). According to the Seventh National Census, the total population outflow from Henan reached 11.7468 million, an increase of 4.3322 million or 58.43% since 2010, showing an obvious growing trend. The number of migrants to major destinations has increased significantly. For example, the number of migrants to Guangdong Province increased from 1.6466 million to 2.2301 million, and the number of migrants to Zhejiang Province increased from 1.2346 million to 2.3969 million [29]. However, the relative proportions changed only slightly. The abovementioned seven provinces and cities remain the top seven destinations attracting a total proportion of more than 75% of rural migrant workers from Henan Province, but their relative positions have changed slightly. Zhejiang has become the largest inflow region of rural migrants from Henan, replacing Guangdong that used to have the largest in

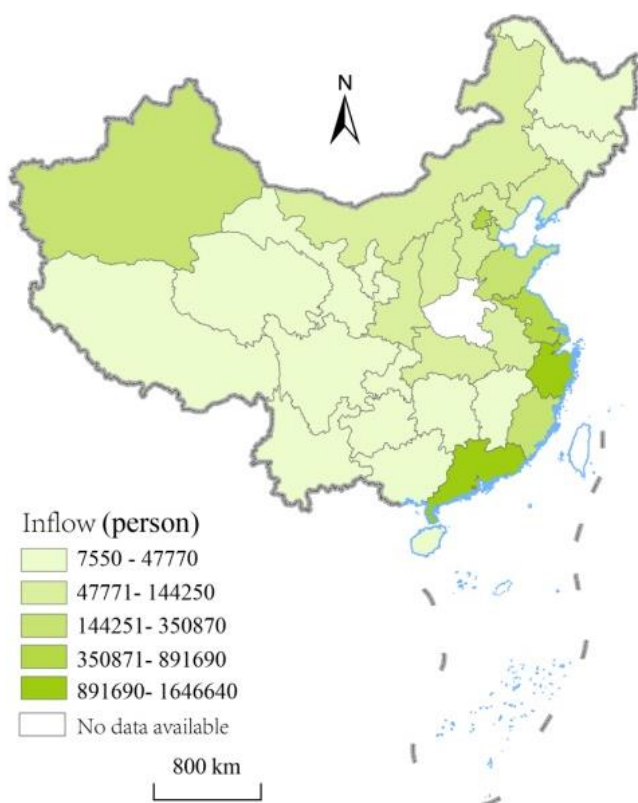
2010, while Shanghai surpassed Beijing to become the fourth largest inflow area. The remaining relative positions did not change.

Guangdong Province has long been the main inflow region of rural migrants from Henan Province, but its relative status has been declining. In the Fifth National Census of China in 2000, it accounted for 35.65% of Henan's migrants (ranked the first), and in the Sixth National Census in 2010, the proportion fell to 22.21% (ranked the first). By the time of the Seventh National Census in 2020, the figure had fallen further to 18.99% (ranked the second) [1]. This may be related to the rapid economic development of the Yangtze River Delta region, which is spatially close to Henan. This change reveals that migration destinations are in the process of dynamic change. Labor migration is the result of "voting with one's feet" to maximize the income.



**Figure 1.** Distribution map of rural population migration of Henan Province in 2020.

On the basis of the review drawing number No. GS (2020) 4619 (produced by the Ministry of Natural Resources of China). The foundation map has not been modified.



**Figure 2.** Distribution map of rural population migration of Henan Province in 2010.

On the basis of the review drawing number No. GS (2020) 4619 (produced by the Ministry of Natural Resources of China). The foundation map has not been modified.

### 3.1.2. Complex Impacts of Distance and Language Difference on Rural Population Migration

The relationship between distance and interprovincial rural migration is not linear. On the one hand, interprovincial rural migrant generally concentrated within a relatively close-range of 1120 km. On the other hand, the correlation analysis of distance and migration reveals no correlation. According to statistics and calculations, the number of rural migrants from Henan Province living within a distance of 500 km is 1.5713 million, accounting for 13.38% of the total; the number of migrants within 800 km is 6.7248 million, accounting for 57.25%; and the number of migrants within 1120 km is 10.7287 million, accounting for 91.33%. These figures show that few people move within 500 km. The bulk of migration is concentrated within the range of 500~1120 km, which accounts for 77.95% of all migrants from Henan and covers the most economically developed coastal provinces and cities in China. According to the correlation analysis of linear distance and migrants, the Pearson correlation coefficient is -0.123 and is not significant, which indicates that there is no significant correlation between the two.

Proximity has an important impact on the outflow of the rural population in Henan Province. Shandong, Hubei, Hebei, Anhui, Shaanxi, Shanxi and other neighboring provinces are

home to large rural migrant populations from Henan. Among them, Shandong Province has an inflow of 402,000, ranking the first among neighboring provinces and accounting for 3.42% of the total outflow from Henan Province; Hubei Province has an inflow of 300,100, accounting for 2.55% of the total and ranking the second. The neighboring province with the smallest inflow is Shanxi Province, but the inflow also reached 196,100, accounting for 1.67% of the total. In the ranking of outflows from Henan Province, the neighboring provinces are second only to the abovementioned seven provinces and regions with the largest outflows, which shows that geographical distance has an important impact on rural population migration but to a lesser degree than economic factors. The main driving factors of rural population migration is to increase the economic income; thus, areas with abundant job opportunities and high wage incomes can attract many rural people. However, population migration is also affected by the first law of geography. The closer the spatial distance is, the smaller the language differences, the more social connections, and the greater the probability of migration is.

### 3.1.3. Chain Migration as the Mechanism for Destination Selection

Chain migration has an important effect on rural migrants' choice of destination. For migrant workers, who represent the main body of rural population migration, the first consideration is whether they can find jobs and obtain higher incomes, and the second consideration is how to safely implement the migration process. Among them, guidance offered through individuals' relationship network often plays a central role, and early migrants often guide new migrants to their destination, which is called chain migration. Chain migration represents a channel effect; through the migration channel, later migrants realize the change in spatial position. Xinjiang is 2440 km away from Henan and is the farthest region from Henan, but with an inflow of 571,900 rural migrants, it is one of the main inflow areas of Henan's rural population, ranking the sixth, higher than Fujian Province and Shandong Province and significantly higher than the surrounding neighboring provinces. This is mainly a result of the migration channel. Before the liberation (1949), most of the migration from Henan Province occurred west along the Longhai Railway. In the early liberation period, Henan supported the development and construction of the frontier. Large-scale population migration also occurred mainly along the Longhai Lanxin Railway (the main east–west traffic artery connecting Henan and Xinjiang). Moreover, migration is related to Xinjiang's rich resources and Chinese government's policies supporting the western region's development. A comparison of the numbers of interprovincial migrants in 2010 and 2020 reveals a high degree of correlation; that is, the provinces and regions with large inflows in 2010 also had large inflows in 2020. The correlation coefficient between the two reaches 0.974 (the significance level 0.01, two-tailed method). This shows that migration destinations

have stability as well as inertia, and highlights the importance of chain migration.

## 3.2. Analysis of Driving Factors

### 3.2.1. Model Operation Results

Taking the number of interprovincial rural migrants flowing out of Henan Province (migration volume) as the explained variable and the above factors as the dependent variables, we carried out the multivariate linear regression anal-

ysis (OLS method) in SPSSAU and used the robust standard error regression method. The regression results are shown in Table 1. In the model,  $R^2=0.966$ , adjusted  $R^2=0.957$ ,  $D-W=1.92$ , and the model passed the F test ( $F=127.457$ ,  $p=0.000<0.05$ ). The statistical indicators of the model thus meet the needs of this study. In the collinearity test of influence factor, the maximum factor variance inflation factor (VIF) is 5.316, and the minimum tolerance is 0.188. As a VIF value  $>10$  is the standard for collinearity, this result indicates that there is no collinearity between the factors.

Table 1. Model operation results.

Factors	Model 1	Model 2
Constant	453295.095* (-1.97)	542105.554 (-1.777)
GDP	3.476** (-2.923)	3.038* (-2.206)
MS	1.527** (-5.759)	1.573** (-5.671)
WL	-5.716 (-1.521)	-5.606 (-1.556)
URD	-2270.751 (-1.350)	-4457.457 (-1.219)
PCSSE	-10.11 (-1.739)	-9.873 (-1.610)
NPSS	-20.658* (-2.376)	-23.208* (-2.146)
Control variable		-61.139 (-0.871)
Sample size	30	30
$R^2$	0.966	0.967
Adjust $R^2$	0.957	0.957
F	$F=127.457$ , $p=0.000$	$F=76.105$ , $p=0.000$
D-W	1.92	1.802

Note: The dependent variable is the number of migrants;  $t$  values are in brackets; \*  $p<0.05$  \*\*  $p<0.01$ .

### 3.2.2. Robustness Test

We use the control variable method to test the robustness of the base model. After the linear distance control variable is added to the base model (see Model 1, Table 1), the significance factor and direction of action of the model basically do not change (see Model 2, Table 1), which reflects the robustness of the results calculated via the above model. After we add the control variables, the significance level of the model is 0.000,  $R^2=0.967$ , adjusted  $R^2=0.957$ ,  $D-W=1.802$ , and  $F=76.105$ . Therefore, the indicators meet the analysis requirements. In addition, the maximum VIF value of the collinearity test of each factor of the model is 5.572, and the minimum tolerance is 0.179, which indicates the absence of collinearity.

### 3.2.3. Factor Effect Analysis

GDP has had an important effect on the choice of destina-

tion of rural migrants from Henan Province. In Model 1 (Table 1), the regression coefficient value of GDP is 3.476, and it is significant at the 0.01 level ( $t=2.923$ ,  $p=0.003<0.01$ ), which means that GDP has a significant positive effect on the number of migrants. Under the condition that other factors remain unchanged, for every 100 million yuan increase in GDP, the number of rural migrants from Henan Province flowing into a place increases by 3.476 persons. In fact, GDP is an iconic indicator of the degree of economic development of a region. It is closely related to various measures, such as jobs, employment opportunities, income levels, and social security. Therefore, high-GDP areas often become migrants' first-choice destinations. Many studies in China and abroad have confirmed the crucial impact of GDP on the migrant population. This also shows that economic factors are important factors in rural population migration.

Migration stock has a significant effect on the outflow of the rural population from Henan Province, and the migration

of the rural population has an inertial effect. In Model 1, the coefficient of the migration stock is 1.527, the significance level is 0.000, and the sign is positive, indicating that the migration stock significantly positively affects the number of migrants. The results of the model show that when other factors remain unchanged, for every increase of one migrant in the migration stock, the number of rural migrants flowing into a place increases by 1.527 persons. The early migration stock has a multiplier effect on the current number of migrants. Moreover, based on the analysis of standardized coefficients, the proportion of influence of migration stock also reaches 66.10%. If this factor is used for unary linear analysis, it can explain 95% of the change in the number of migrants (the indicators of the model meet the analysis requirements). This means that the migration stock has an extremely important impact on the current number of migrants and represents the main mechanism of rural population migration.

In fact, in the process of the long-distance migration of rural populations, chain migration is an important mechanism, and the guidance of early migrants is an important way for rural populations to achieve inter provincial migration. Moreover, the concentration of the rural population in an inflow region in the previous period reflects the ability of the region to attract the rural population, which is related to its developed manufacturing and tertiary industries and has a directional influence on later rural population migration. When we analyze the data from the Sixth National Census, we find a similar pattern; that is, the provincial distribution of rural population outflows in Henan Province in 2010 is closely related to the regional distribution of outflows in 2000. The direction and quantity of rural population migration, in the absence of mutations in external factors, continues to develop along the previously formed inertial direction. This is the inertial law of rural population migration.

The number of primary and secondary schools has an important influence on migrants' destination choices. In Model 1, the factor reaches the 0.05 significance level, but the coefficient is negative, indicating that the number of primary and secondary schools has a significant negative effect on rural population migration. A possible reason is that the number of schools does not necessarily reflect the ease of admission for rural migrants' children. The degree of convenience is more likely to depend on the local attitude toward immigration and the admission policy regarding children who accompany their migrant parents. The negative coefficient may also be explained by the interaction between factors during the calculation of the model.

The linear distance factor does not reach significance in Model 1, which indicates that this factor has no influence on the number of rural migrants. With the development of China's transportation industry, especially with the rapid development of high-speed railway, the time and physical costs of travel have been greatly reduced. Migrant workers no longer pay much attention to ticket costs and travel time but are more concerned about job opportunities and wage incomes.

Moreover, the penetration rate of Mandarin in China has reached a high level, and the regional language differences have gradually decreased. The immigration barriers caused by language are negligible, which may contribute to the insignificance of the spatial distance factor.

In short, GDP and migration stock have a significant positive effect on the number of migrants at a destination, and the number of primary and secondary schools has a significant negative effect on the number of migrants. However, the differences in per capita social security expenditures, wage levels, and urbanization rates do not affect the number of migrants, and neither do straight-line distances or language differences.

## 4. Discussion

The rural population constitutes the majority of China's population migration and presents different migration characteristics than the urban population. Due to data constraints, previous research on the spatial structure of interprovincial migration among rural populations is limited. By mining data from China's Seventh National Census data, we obtain data on the interprovincial migration of the rural population. On this basis, the spatial structure characteristics and factors of the interprovincial migration of the rural population in typical agricultural areas in Henan Province are analyzed, which enriches the research on the spatial migration of labor. This study suggests that rural migrants' choice of destination provinces and regions depends mainly on employment opportunities. The migration inertia is prominent, and distance has a complex impact. We provide an objective analysis of the patterns of rural population migration and innovative insights are offered.

In terms of migration direction, China's rural population flows mainly on an east–west axis, with Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang and Guangdong Provinces as the main destinations; while the central and western regions as the main source areas [7]. Among them, Guangdong, Shanghai, Jiangsu, and Zhejiang in the eastern regions are the main destination provinces, and Anhui, Hunan, Jiangxi, Henan, and Sichuan in the central and western regions are the main outflow provinces [6]. Henan Province is located in the central region and is an important export area of the rural population in China. The rural migrants from Henan account for approximately 10% of the national migration population. This study reveals that the main destinations of the rural migrants from Henan Province are Zhejiang, Guangdong, Jiangsu, Shanghai, Beijing, Xinjiang, and Fujian Provinces (accounting for 76.25%). The migration direction is basically consistent with that observed in previous research [6, 7], but Henan also has its own unique characteristics; that is, the group of rural migrants flowing westward into Xinjiang is also relatively large, reaching 571,900 (accounting for 4.87%) and exceeding the number in Fujian Province, which is in the southeastern coastal area.

A previous study of rural labor migration in Henan Province [35] disregarded this westward migration and only explored the migration to the three major metropolitan areas in the east (the Yangtze River Delta Metropolitan Area, the Beijing–Tianjin–Hebei Metropolitan Area, and the Guangdong–Hong Kong–Macao Greater Bay Area).

At a macro level, rural population migration is affected mainly by economic factors, but our research suggests that the role of migration inertia cannot be ignored. China's eastern coastal areas are the most developed regions and have a high concentration of manufacturing and service industries; in particular, the three major metropolitan areas attract many rural laborers from underdeveloped areas in the central and western regions to work and obtain wage income [3]. The outflow of the rural population in the central and western regions is also influenced by the inability of the local region to absorb a large number of surplus rural laborers. We believe that GDP, as an important indicator of the level of economic development, explains the direction of rural population migration in Henan Province to a large extent. This finding is similar to the conclusions of Ke et al. (2023) [31] and Liu (2017) [35]. However, when the migration stock factor is incorporated into the model, the influence of the GDP factor decreases significantly. This shows that the macrolevel migration stock factor explains the direction of rural population migration to a great degree. However, in current research, rural population migration is rarely explained from the perspective of migration stock. In this context, our results contribute to filling the gaps in this field of research.

Although we focus on the macrolevel, the mechanism of migration inertia revealed by the results is consistent with the findings of previous microlevel studies. As mobility and migration entail high costs, such as social risks, language barriers, travel costs and living expenses, and the psychological pain of leaving their families [18], the network of relationships is generally considered as an important factor in rural workers' choice of migration destination. In addition, the likelihood of securing employment, earning a wage, and appreciating the living conditions are fundamental pieces of information that migrants need to consider. It is helpful and convenient for migrants to obtain this information through a trusted and reliable system [9]; thus, the social relationship network offers important guidance regarding migrants' choice of destination. In the international population migration, language is a very important influencing factor [36] because it is directly related to communication with locals, and further connected with cultural integration that is another key variable for immigrants working for a long period in a foreign country [37]. In this study, migration inertia actually covers the total social relationship network at the individual level. After early immigrants work and settle in a certain city, they guide later immigrants to move into the city and complete the migration process through social networks. This results in the phenomenon of population migration to the previously estab-

lished locations.

Most studies posit that distance is an obstacle to population movement and migration. As the distance increases, the number of migrants decreases. As early as the end of the 19th century, Ravenstein (1889) [38] proposed the principle of the distance attenuation of population migration, arguing that as the spatial distance increases, the number of migrants decreases. Zipf (1946) [39] introduced Newton's theorem into the push–pull model of population migration and argued that the flow of population between regions is proportional to the size of the population of the two places and inversely proportional to the distance. In recent years, Chinese scholars have investigated the attenuation effect of distance between cities on population migration on the basis of big data [40]. Distance is generally believed to be an important factor affecting population migration, but the two do not necessarily present a linear relationship [41]. We believe that given the existence of developed modern transportation, the principle of distance attenuation is not fully applicable, and the traditional distance attenuation model for population migration needs to be updated and supplemented. In the present, owing to the significant reduction in transportation time and the high penetration rate of Mandarin, geographic spatial distance and dialects play a far less important role than economic factors, i.e., migration is more likely to occur over longer spatial distances.

Language and dialects are important factors affecting population migration. Language plays an important role in promoting or hindering population migration. Popularization of Mandarin can eliminate communication barriers between workers in different provinces, which is conducive to cross-provincial population migration [42]. With the increase of educational opportunities and the promotion of Mandarin, it has become more commonly-used among domestic migrants in China [43]. In the context of international migration, mastering the official language or common language of the target country (such as English) may also reduce the difficulty of migrants' adaptation, thereby promoting the occurrence of migration behavior [44]. Language significantly affects all aspects of immigration—from cultural adaptation and economic opportunity to educational success and crisis management [45]. However, dialects may inhibit the inter-provincial migration of labor. In areas with high dialect usage, residents are more inclined to stay locally than to migrate to other provinces [42]. Though the above studies lack direct evidence connecting the widespread use of Mandarin with the impact of reduced distance on domestic migration, it can be reasonably inferred that the widespread use of Mandarin is minimizing the impact of geographical distance on internal migration in China, constituting an important finding of this paper. This is also consistent with the global observation that standardized languages help integrate different populations [46].

By obtaining data on the migration of the rural population in Henan Province from the Seventh National Census of China,

this study conducts an in-depth analysis of the migration of the rural population in typical agricultural areas. We believe that the rural migrants from Henan Province are distributed across all of China but are concentrated mainly in certain regions. Geographical distance and language differences have complex and nonlinear effects on migration. The choice of migration destinations concerns mainly economic factors. The early migration stock has a key impact on population migration in the present. Chain migration based on social networks is an important method for external migration, and migration inertia is the basic law of rural population migration.

According to this inertial law, it is suggested that in the development of the labor economy in regions that export labor, attention should be given to the cultivation of “seed” migrant workers to encourage other rural laborers to go out to work and reduce risks related to migration. However, in the long run, out-migration is not the best choice for rural laborers. Instead, it is fundamental to address rural labor transfer by achieving local urbanization near potential migrants’ hometowns. In China, the economic development of counties in the central and western regions should be accelerated to promote the employment of nearby laborers. Moreover, a more effective new mechanism for coordinated regional development should be established to transfer labor-intensive industries to the central and western regions and reduce the cross-regional migration of labor.

There are still limitations in this study. First, this study uses provincial panel data on a large research scale. If municipal and county-level data could be obtained, research in this field could be deepened. Second, due to the lack of data on the microlevel drivers of rural population migration, the influencing factors are explored only at the macro level but not analyzed at the micro level. Further research from a micro perspective is needed in future. In addition, although the migration stock factor reveals the guiding role of the relationship network in the selection of rural population’s migration destinations, quantitative data for describing the relationship network are absent, precluding the possibility of conducting quantitative research. Finally, this study takes the representative Henan Province as the case, and the conclusions are mainly drawn on rural population migration data from this area. Whether the conclusions can be generalized requires further case studies for verification.

## 5. Conclusions

As the main body of population migration, rural population migration has had an important impact on China’s economic and social development. Drawing on data from the Seventh National Census, statistical analysis, spatial analysis and a multivariate linear regression model are used to study the spatial patterns and driving factors for the interprovincial migration of the rural population in Henan Province. The following conclusions can be drawn.

First, the cross-provincial migrants from Henan Province

are scattered throughout the country, but they are concentrated mainly in Zhejiang, Guangdong, Jiangsu, Shanghai, Beijing, Xinjiang, and Fujian. Since the Sixth National Census of China, the spatial structure of migration has not undergone substantial changes, but the total migration volume has increased significantly. The eastern and southern coastal provinces, which are highly developed economically, are migrants’ main destinations, and the Chinese capital Beijing is a large inflow area. For historical and resource reasons, Xinjiang is also an important destination. Areas with long geographical distances, large language differences, and relatively backward economies have the least inflow. On the whole, the level of economic development basically determines the migration direction of the rural population, and migration inertia is an important mechanism explaining rural population migration.

Second, distance has a complex impact on interprovincial population migration. On the one hand, the interprovincial outflow of the rural population is concentrated within a small range of 1120 km. On the other hand, outflow is not negatively correlated with distance; while adjacent locations in neighboring provinces can attract rural migrants. Owing to the development of modern and developed transportation and communication technologies, geographic distance has become less important than economic factors, but geographic space and language differences are still basic factors affecting population migration.

Third, the significant factors for the interprovincial migration of the rural population in Henan Province are GDP, migration stock and the number of primary and secondary schools. GDP and migration stock have positive impacts on the migration volume, and the number of primary and secondary schools negatively affects the migration volume. The long-distance migration of the rural population is the result of chain migration formed by reliance on relationship network. In a non-mutant economic environment, early population migration has a decisive effect on later migration. Migration inertia serves as an important law of migration, but the fundamental reason for the choice of migration destinations is the availability of employment opportunities and wage incomes, which are directly related to the level of economic development.

## Abbreviations

WL	Wage Level
URD	Urbanization Rate Difference
PCSSE	Per Capita Social Security Expenditure
NPSS	Number of Primary and Secondary Schools
MS	Migrant Stock

## Author Contributions

Ge Gao is the sole author. The author read and approved

the final manuscript.

## Data Availability Statement

The data is available from the corresponding author upon reasonable request.

## Funding

This work is supported by the Beijing Municipal Social Science Foundation (Grant No. 23YYC013).

## Conflicts of Interest

The author declares no conflicts of interest.

## References

- [1] Office of the Leading Group for the Seventh Census of the State Council. *China Census Yearbook 2020*. (China Statistics Press, 2022).
- [2] National Bureau of Statistics of China. *2022 Migrant Worker Monitoring Survey Report*, [https://www.stats.gov.cn/xxgk/sjfb/zxfb2020/202304/t20230428\\_1939125.html](https://www.stats.gov.cn/xxgk/sjfb/zxfb2020/202304/t20230428_1939125.html) (2023).
- [3] Wang, G. New trends in migration and urbanization in China: A preliminary investigation based on the Seventh Census Data. *Popul. & Econ.*, 36-55, <https://doi.org/10.3969/j.issn.1000-4149.2021.00.038> (2021).
- [4] National Bureau of Statistics. *Bulletin of the Seventh National Census (No. 7)*. [https://www.stats.gov.cn/sj/tjgb/rkpcgb/qgrkpcgb/202302/t20230206\\_1902007.html](https://www.stats.gov.cn/sj/tjgb/rkpcgb/qgrkpcgb/202302/t20230206_1902007.html) (2021).
- [5] IOM. *World Migration Report 2024*. <https://publications.iom.int/books/world-migration-report-2024> (2024).
- [6] Wang, G. & Lu, Q. Migration of rural population in recent years in China. *Sci. Geog. Sin.*, 630-635 (2007).
- [7] Jiang, Z., Wang, J., Yin, W. & Yuanyuan, Z. Spatio-temporal preferences and driving mechanism of migrant workers' mobility in China: Based on the data of the National Floating Population Dynamic Monitoring Survey from 2011 to 2018. *J. Huazhong Norm. Univ., Nat. Sci.* 56, 1011-1020, <https://doi.org/10.19603/j.cnki.1000-1190.2022.06.013> (2022).
- [8] Gao, G., Yang, H., Xu, J., Luo, Q. & Hai, B. Spatial change of employment destination from the first to the last of migrant workers. *Econ. Geog.* 36, 143-148, <https://doi.org/10.13959/j.issn.1003-2398.2016.02.008> (2016).
- [9] Gui, H. Analysis of the influencing factors of the flow selection for migrant workers going out: Taking Sichuan's household registration population as an example. *Contemp. Econ.*, 111-113 (2017).
- [10] Guo, Y., Zhou, Y. & Liu, Y. Spatial-temporal evolution of rural population outflow and its mechanism in China. *Sci. Geog. Sin.* 40, 50-59, <https://doi.org/10.13249/j.cnki.sgs.2020.01.007> (2020).
- [11] Gao, G., Yang, H., Xu, J., Hai, B. & Li, E. Study on the instability of migrant workers' working place: A case study of 33 villages in Henan Province. *Hum. Geogr.* 31, 43-49, <https://doi.org/10.13959/j.issn.1003-2398.2016.02.008> (2016).
- [12] Gao, G., Duan, X., Ji, H. & Zhu, J. Entrepreneurial location of returned migrant workers: A case study of 14 villages in Henan Province. *Geogr. Res.* 40, 2871-2884, <https://doi.org/10.11821/dlyj020191130> (2021).
- [13] Qian, H., Zhu, Y., Lin, L. & Ke, W. Regional return migration of China's floating population and the factors influencing their locational selections. *Geogr. Res.* 42, 2089-2103, <https://doi.org/10.11821/dlyj020230007> (2023).
- [14] Yin, J. *et al.* Return migration intention: Heterogeneous effects from multidimensional distances over space. *Acta Geogr. Sin.* 77, 3146-3161, <https://doi.org/10.11821/dlxb202212013> (2022).
- [15] Yu, W. The regional characteristics of the floating population's language situation in China and its influencing factors. *Appl. Linguist.*, 12-19, <https://doi.org/10.16499/j.cnki.1003-5397.2019.03.002> (2019).
- [16] Rudiarto, I., Hidayani, R. & Fisher, M. The bilocal migrant: Economic drivers of mobility across the rural-urban interface in Central Java, Indonesia. *J. Rural Stud.* 74, 96-110, <https://doi.org/10.1016/j.jrurstud.2019.12.009> (2020).
- [17] Al-Maruf, A., Pervez, A. K., Sarker, P. K., Rahman, M. S. & Ruiz-Menjivar, J. Exploring the factors of farmers' rural-urban migration decisions in Bangladesh. *Agriculture*. 12, 722, <https://doi.org/10.3390/agriculture12050722> (2022).
- [18] Selod, H. & Shilpi, F. Rural-urban migration in developing countries: Lessons from the literature. *Reg. Sci. Urban Econ.* 91, 103713, <https://doi.org/10.1016/j.regsciurbeco.2021.103713> (2021).
- [19] Sani Ibrahim, S., Ozdeser, H., Cavusoglu, B. & Shagali, A. A. Rural migration and relative deprivation in agro-pastoral communities under the threat of cattle rustling in Nigeria. *SAGE Open*. 11, 2158244020988856, <https://doi.org/10.1177/2158244020988856> (2021).
- [20] Dessalegn, M., Debevec, L., Nicol, A. & Ludi, E. A critical examination of rural out-migration studies in Ethiopia: Considering impacts on agriculture in the sending communities. *Land*. 12, 176, <https://doi.org/10.3390/land12010176> (2023).
- [21] Busso, M., Chauvin, J. P. & Herrera, N. Rural-urban migration at high urbanization levels. *Reg. Sci. Urban Econ.* 91, 103658, <https://doi.org/10.1016/j.regsciurbeco.2021.103658> (2021).

- [22] Duda, I., Fasse, A. & Grote, U. Drivers of rural-urban migration and impact on food security in rural Tanzania. *Food Sec.* 10, 785-798, <https://doi.org/10.1007/s12571-018-0788-1> (2018).
- [23] Thakur, G. M. & Guha, S. Petty services, profit - led growth and rural-urban migration in a developing economy. *Metroeconomica*. 70, 144-171, <https://doi.org/10.1111/meca.12233> (2019).
- [24] Camarero, L. & Oliva, J. Thinking in rural gap: Mobility and social inequalities. *Palgrave Commun.* 5, 1-7, <https://doi.org/10.1057/s41599-019-0306-x> (2019).
- [25] Das, P., Saha, J. & Chouhan, P. Effects of labor out-migration on socio-economic set-up at the place of origin: Evidence from rural India. *Child Youth Serv. Rev.* 119, 105512, <https://doi.org/10.1016/j.childyouth.2020.105512> (2020).
- [26] Fassil, E., Jema, H., Mengistu, K. & Abule, M. Impact of rural out-migration on agricultural technology adoption of rural households in Southern Ethiopia. *AGRIS P. Econ. Inform.* 15, <https://doi.org/10.7160/aol.2023.150101> (2023).
- [27] National Bureau of Statistics of China. *China Statistical Yearbook 2023*. (China Statistics Press, 2023).
- [28] Henan Provincial Bureau of Statistics. *Economic Overview of Henan Province*, <https://www.henan.gov.cn/2023/03-23/2712571.html?eqid=adc4486000078ea90000000264674914> (2023).
- [29] The Census Office of the State Council, the Department of Population and Employment Statistics of the National Bureau of Statistics. *China's 2010 Census data*. (China Statistics Press, 2012).
- [30] National Bureau of Statistics of China. *China Statistical Yearbook 2021*. (China Statistics Press, 2021).
- [31] Ke, W., Xiao, B., Lin, L., Zhu, Y. & Wang, Y. Interprovincial urban and rural floating population evolution of China and its relationship with regional economic development. *Acta Geogr. Sin.* 78, 2041-2057, <https://doi.org/10.11821/dlxb202308012> (2023).
- [32] Liu, T., Qi, Y. & Cao, G. China's floating population in the 21st century: Uneven landscape, influencing factors, and effects on urbanization. *Acta Geogr. Sin.* 70, 567-581, <https://doi.org/10.11821/dlxb201504005> (2015).
- [33] Yan, S. Inter-provincial migration and its determinants in the 1990's China. *Chin. Popul. Sci.*, 71-77 (2007).
- [34] Zhang, W., & Dong, W. *Advanced Tutorial on SPSS Statistical Analysis*. (Higher Education Press, 2018).
- [35] Liu, D. Analysis of population flow in Henan's new urbanization. *Contemp. County. Econ.*, 34-35, <https://doi.org/10.16625/j.cnki.51-1752/f.2017.12.012> (2017).
- [36] Adizovna, S. M. Language and migration, exploring linguistic integration in a globalized world. *Eur. J. Innov. Nonform. Edu.* 4, 184-189 (2024).
- [37] Akintayo, O. T., Atobatele, F. A. & Mouboua, P. D. The dynamics of language shifts in migrant communities: Implications for social integration and cultural preservation. *Int. J. Appl. Res. Soc. Sci.* 6, 844-860, <https://doi.org/10.51594/ijarss.v6i5.1106> (2024).
- [38] Ravenstein, E. G. The laws of migration. *J. Royal Stat. Soc.* 52, 241-305, <https://doi.org/10.2307/2979333> (1889).
- [39] Zipf, G. K. The  $P_1P_2/D$  hypothesis: On the intercity movement of persons. *Am. Sociol. Rev.* 11, 677-686, <https://doi.org/10.2307/2087063> (1946).
- [40] Gao, G., Deng, Y., Yuan, Y., Zhang, W., Zhang, J. & Jin, M. Return location of migrant workers: A case study of 14 sample villages in Henan Province, China. *Earth Sci.* 10(4), 145-156, <https://doi.org/10.11648/j.earth.20211004.11> (2021).
- [41] Dan, J. & Yin, J. Migration distance of floating population and regional differences of its influence on urbanization. *Acta Sci. Nat. Univ. Pekin.* 53, 487-496, <https://doi.org/10.13209/j.0479-8023.20175.043> (2017).
- [42] Huang, Z., & Yang, J. The impact of dialect on inter-provincial migration in China, *Popul. Res.*, 44(4): 89-101 (2020).
- [43] Lobanov-Rostovsky, S., He, Q., Chen, Y. *et al.* Growing old in China in socioeconomic and epidemiological context: Systematic review of social care policy for older people. *BMC Publ. Health*, 23, 1272, <https://doi.org/10.1186/s12889-023-15583-1> (2023).
- [44] Li, S. Immigration and language change: A study of a new variety of Putonghua by two generations of a Chinese community in York of UK. *China Lang. Strateg.*, (1): 175-188 (2024).
- [45] Bellofatto, A. How does language impact foreign investing in a multilingual country?. *SSRN Electron. J.*, 3001825, <https://doi.org/10.2139/ssrn.3001825> (2017).
- [46] Chen, S. C. Multilingualism in Taiwan. *Int. J. Sociol. Lang.* 205, 79-104, <https://doi.org/10.1515/ijsl.2010.040/html> (2010).