

Pu, Yingxia; Ge, Ying

**Conference Paper**

## Multilateral mechanism analysis of interprovincial migration flows in China

56th Congress of the European Regional Science Association: "Cities & Regions: Smart, Sustainable, Inclusive?", 23-26 August 2016, Vienna, Austria

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Pu, Yingxia; Ge, Ying (2016) : Multilateral mechanism analysis of interprovincial migration flows in China, 56th Congress of the European Regional Science Association: "Cities & Regions: Smart, Sustainable, Inclusive?", 23-26 August 2016, Vienna, Austria, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/174656>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

---

## Multilateral Mechanism Analysis of Interprovincial Migration Flows in China

Yingxia PU<sup>a\*</sup>, Ying GE<sup>b</sup>, Hongling HAN<sup>c</sup>, Fanhua KONG<sup>d</sup>,

<sup>a</sup> School of Geographic and Oceanographic Sciences, Nanjing University, Nanjing 210023, China. Email address: [yingxiapu@nju.edu.cn](mailto:yingxiapu@nju.edu.cn);

<sup>b</sup> School of Earth Science and Engineering, Hohai University, Nanjing 210097, China. Email address: [geying@hhu.edu.cn](mailto:geying@hhu.edu.cn);

<sup>c</sup> College of Business, Linyi University, Linyi 276000, China. Email address: [hanhongling@lyu.edu.cn](mailto:hanhongling@lyu.edu.cn);

<sup>d</sup> International Institute for Earth System Science, Nanjing University, Nanjing 210023, China. Email address: [fanhuakong@163.com](mailto:fanhuakong@163.com).

**Abstract:** Population migration flows between different regions are related to both origin- and destination-specific attributes and to migration flows to and from neighboring regions. Intuitively, changes to the characteristics of a single region will impact both inflows and outflows to all other regions as well as the region itself. In order to explore these multilateral effects driving increasing population migration flows in China, this paper builds an unconstrained spatial origin-destination (OD) model of interprovincial migration using the Sixth National Population Census data and other related socio-economic data. First, migration flows show significant network autocorrelation effects among origin and destination regions, which mean that the migration behaviors of migrants in one region are influenced by those of other regions. Second, multilateral effects of regional economic and social factors through the spatial network system lead to clustered migration flows across interrelated regions. The distance-decay effect plays the most influential role in shaping migration flows among all the factors, and the negative spillover effect further exacerbates the friction of distance. In terms of destinations, the influences of wage level and migration stocks are beyond that of GDP, and the positive spillover effects of these factors enhance the attraction of neighboring regions. The spillover effects of unemployment rates and college enrollment in higher education are significantly negative, while destination population is not significant. In terms of origins, the multilateral effects of unemployment rates

---

are negative, indicating that interprovincial mobility capability tends to slow down with the decrease of employment rate. Finally, changes to regional characteristics will potentially affect the whole migration system.

**Key words:** China; multilateral effects; network autocorrelation; population migration flows; spatial mechanism analysis; spatial OD model

## 1 Introduction

Along with the integration of the global economy and development of urbanization processes, the increase in population migration among regions following different spatial scales has had far-reaching influences on the redistribution of the population and socio-economic development of countries (or regions). These trends have attracted much attention from governments, academia, and societies globally (IOM, 2008). Population migration modeling methods can be classified into micro and macro schools, according to their respective theoretical perspectives and methodologies (Stillwell, & Congdon, 1991). The former identifies different factors that influence migration decisions for individuals or households, including, sex, age, size and structure of a family, and employment, wage or house value of a destination, and decisions are based on discrete choice theory and maximum utility (Smith, 1975; Sen, & Smith 1995). The latter draws on classical economy and statistical equilibrium methods, and investigates the impacts of macro socio-economic structures by locating population migration within the context of country or regional labor market, and identifies the “push” and “pull” forces driving migration (Lee, 1966; Wilson, 1967; Harris, & Todaro, 1970; Massey, 1990; Roy, & Thill, 2004). Due to dearth of micro survey data about individual or household migration decisions, research on establishing macro relationship between migration and influential factors is more extensively available (Stillwell,

---

2005).

From a systematic perspective, migration flows among different regions constitute a complex migration network system. A region can be regarded as an origin (O) or a destination (D) node, and origin-destination (OD) flows between nodes are not only related to the origins and destinations individually, but also to the multilateral spillover effects from neighboring regions (LeSage, & Pace, 2008, 2009). However, traditional macro migration modeling mainly based on gravity models (or spatial interaction models) works on the assumption of independence, meaning that the magnitude of migration flows is determined by “push or pull” factors of origins or destinations (namely, population, income or employment rates), as well as separation distances between regions. This implies that a change of particular factor in one region only influences the  $2 \times (n-1)$  flows that are directly relevant to this region, and the other flows that do not take this region as an origin or a destination will not experience any change.

Since the 1970s, the practice of testing the independence assumption in gravity models and making corresponding adjustments has laid the foundation of the formal expression of multilateral effects. Curry (1972) firstly points out that distance variables alone cannot avoid the spatial dependence among residuals in gravity models. Griffith and Jones (1980) further present the spatial lag of dependent variables or error terms can capture the dependence in residuals. Goodchild (1987) argues for autocorrelation among arcs when there is a similarity between the attributes of arcs in adjoining nodes in a network system. Black (1992) applies global Moran's *I* statistic to verify the existence of network autocorrelation among migration flows. Until recently, LeSage and Pace (2008) specify 9 spatial autoregressive models based on different restrictions on three kinds of dependence in a network system, and provide their maximum likelihood estimations.

---

Chun (2008) employs the Eigenvector Spatial Filtering method to deal with network autocorrelation among migration flows. Though spatial filtering can decrease the bias of parameter estimates, it may filter out some useful information, such as spillover effects (Pace, LeSage, & Zhu, 2013). LeSage and Thomas-Agnan (2015) present the concept of total effects based on partial derivatives to reinterpret the coefficient estimates of spatial OD models, which provides a useful tool for comparing multilateral effects with coefficient estimates of gravity models.

With the huge amount of migration population flows since the open door policy, a lot of empirical work has been done to identify push and pull forces based on gravity models (Yang, Chan, & Liu, 1999; Cai, & Wang, 2003; Wang, Pan, & Lu, 2012). Though gravity models provide a bilateral framework for analyzing the main push and pull forces during the migration process, they do not account for the dependence among migration flows. Yu, Pu, Chen, & Wang (2012) analyze the effects of spatial dependence on parameter estimates. However, interpretation of these coefficient estimates of explanatory variables requires further investigation.

The structure of this paper is as follows. Section 2 introduces the methodology and provides the interpretation of multilateral effects based on the concept of partial derivatives. Section 3 clarifies the sixth National Census migration data and related socio-economic variables, and establishes the unrestricted spatial OD model to analyze the multilateral effects among Chinese interprovincial migration flows. Simulation of chain effects on interprovincial migration flows due to 5 percent increase of GDP in Jiangsu Province is further explored in Section 4, which provides a reasonable explanation of the spatial mechanism driving the process of interprovincial migration. Conclusions are drawn in the last section to provide some guidelines for future modeling.

---

## 2 Methodology

### 2.1 Bilateral effects and gravity models

Gravity models (including various restricted versions) are widely applied in migration modeling cases. Specifically, the characteristics of origins and destinations are used to express the “push and pull” forces at play during the migration process, and the distance variable is included to explain the decay effect seen across space. The log-form of gravity model is illustrated as follows (LeSage, & Pace, 2008):

$$y = \alpha \iota_N + X_d \beta_d + X_o \beta_o + \gamma g + \varepsilon \quad (1)$$

where  $y$  stands for the logged vector of interprovincial migration flows;  $\iota_N$  is a vector of ones;  $X_d$  and  $X_o$  represent explanatory variable matrices of the destinations and origins;  $g$  is an  $N$  by 1 vector of distances between origins and destinations;  $\alpha$  denotes the constant term parameter on  $\iota_N$ ;  $\beta_d$ ,  $\beta_o$  are  $k$  by 1 parameter vectors;  $\gamma$  reflects the effect of distance  $g$ ;  $\varepsilon$  is an  $N$  by 1 error term vector, and we assume  $\varepsilon \sim N(0, \sigma^2 \mathbf{I}_N)$ .

After the logarithm transformation, gravity models can be estimated by ordinary least squares (OLS), and variable parameters can be treated as *bilateral effects or elasticities*, meaning if some explanatory variable increases or decreases by 1%, the dependent variable will change at the amount of the coefficient of this variable. If the parameter estimate for the  $r$ th destination characteristic ( $\beta_d^r$ ) is positive, it indicates that there is greater attraction to the destination for migrants. Similarly, if the parameter estimate for the  $r$ th origin characteristic ( $\beta_o^r$ ) is positive, more migrants will be from this origin location. Therefore, the related origin and destination characteristics correspond to “push” and “pull” factors. LeSage and Thomas-Agnan (2015) further argue that the partial derivative of the change of the dependent variable ( $y$ ) due to the change of

---

the independent variable ( $X_d$  or  $X_o$ ) can be used to evaluate the influence of the changes to various regional characteristics on migration flows.

## 2.2 Multilateral effects and spatial OD models

In fact, regional migration flows are dependent not only on the bilateral effects of origin and destination characteristics, but also on the neighboring regions of these flows. [LeSage and Pace \(2008\)](#) define spatial dependence as larger flows, which stem from origin A to destination B and are accompanied by: (1) larger flows from regions nearby origin A to destination B, which they label as *origin-dependence*, (2) larger flows from origin A to regions neighboring destination B, which they label as *destination-dependence*, and (3) larger flows from regions that are neighbors to origin A to regions that are neighbors to destination B, which they label as *origin-to-destination dependence*. Three kinds of network matrices ( $W_d$ ,  $W_o$ ,  $W_w$ ) are accordingly set up to express the above destination-, origin-, and origin-to-destination dependence. Based on the restrictions of these matrices, several spatial OD models are built by incorporating different spatial lags of dependent variables into traditional gravity models. [Chun \(2008\)](#) provides social behavior underpinnings for the network weight matrices by relating “intervening opportunities” and “competing destinations” with origin-based and destination-based matrices. The following is an unrestricted expression of the spatial OD model ([LeSage, & Pace, 2008](#); [LeSage, & Thomas-Agnan, 2015](#)).

$$y = \rho_o W_o y + \rho_d W_d y + \rho_w W_w y + \alpha t_N + X_d \beta_d + X_o \beta_o + \gamma g + \varepsilon \quad (2)$$

where  $W_o y$  is the origin-based spatial lag of dependent variable, indicating the average migration flows from the neighboring region of each origin to each of the destinations;  $W_d y$  is the destination-based spatial lag of dependent variable, indicating the average migration flows from

---

each origin to the neighboring region of each of destination;  $W_w y$  is the origin-to-destination based spatial lag of dependent variable, indicating the average migration flows from the neighboring region of the origin to neighboring region of the destination;  $\rho_o$ ,  $\rho_d$  and  $\rho_w$  measure the strengths of network autocorrelation effects among origins, destinations and flows.

According to the explanation of spatial regression models discussed in [LeSage and Pace \(2009\)](#), the associated data generation process (DGP) for the spatial OD model shown in (2) is as follows:

$$y = (I - \rho_o W_o - \rho_d W_d - \rho_w W_w)^{-1} (\alpha I_N + X_d \beta_d + X_o \beta_o + \gamma g + \varepsilon) \quad (3)$$

In order to understand the multilateral effects and their sources during the migration process, transformation needs to be exerted on the inverse matrix shown in (3). Under the restriction assumption of spatial stationarity, and the row standardization of the three network weight matrices ( $W_o$ ,  $W_d$ ,  $W_w$ ), the above inverse matrix can be expanded into equation (4) according to the Taylor series as follows:

$$(I - \rho_o W_o - \rho_d W_d - \rho_w W_w)^{-1} = I + \rho_o W_o + \rho_d W_d + \rho_w W_w + (\rho_o W_o + \rho_d W_d + \rho_w W_w)^2 + (\rho_o W_o + \rho_d W_d + \rho_w W_w)^3 + \dots \quad (4)$$

where identity matrix  $I$  means that the change of regional factors only influences the migration flows related to this region; the term  $\rho_o W_o + \rho_d W_d + \rho_w W_w$  implies that the change of regional factors can impact the neighbors of the origin or the neighbors of the destination and their corresponding migration flows through the network connectivity relationship; and the term  $(\rho_o W_o + \rho_d W_d + \rho_w W_w)^2$  reflects the influence of change of regional factors on the migration flows from the neighbors of neighbors to the origin or flows to the neighbors of neighbors to the destination by the second order of network matrices, and so on. Change will occur in the entire migration system because of this series of spatial spillover effects arising from changes experienced in regional factors.



The above network effects seen in traditional gravity models can be neglected by assuming  $\rho_o$ ,  $\rho_d$  and  $\rho_w$  zeros, and equation (4) under the assumption can be simplified as an identity matrix, which means that the change of regional factors in one region will not alter migration flows from other origins or to other destinations. More generally, when all the network effects ( $\rho_o$ ,  $\rho_d$ ,  $\rho_w$ ) are not zeros, equation (4) will be not a simple identity matrix, and the elements in the non-diagonal will be not zeros. If the  $r$ th factors of region  $i$  changes, it will impact migration flows to or from this region, but also have an impact on other migration flows through the neighboring regions to the origin or destination, and pass the impacts of factor changes onto other migration flows. The multilateral effects occurred can be expressed as follows (LeSage, & Thomas-Agnan, 2015):

$$TE = \begin{pmatrix} \partial y_1 / \partial x_1^r \\ \partial y_2 / \partial x_2^r \\ \dots \\ \partial y_n / \partial x_n^r \end{pmatrix} = (I - \rho_o W_o - \rho_d W_d - \rho_w W_w)^{-1} \begin{pmatrix} Jd_1 \beta_d^r + Jo_1 \beta_o^r \\ Jd_2 \beta_d^r + Jo_2 \beta_o^r \\ \dots \\ Jd_n \beta_d^r + Jo_n \beta_o^r \end{pmatrix} \quad (5)$$

where  $Jd_i$  is zero matrix with  $n \times n$  dimensions, and the elements of the  $i$ th row are 1;  $Jo_i$  is a zero matrix with  $n \times n$  dimensions, and the elements of the  $i$ th column are 1;  $TE$  is an  $n^2 \times n$  partial derivative matrix, indicating the total effects of the change of the  $r$ th variable on migration flows; partial derivative  $\partial y_i / \partial x_i^r$  indicates total effects of the change of the  $r$ th variable of region  $i$  on all migration flows.

### 3 Multilateral effects in the Chinese interprovincial migration system

#### 3.1 Data Sources and Variable Choices

Interprovincial migration data used in this study are aggregated from the “population that has lived in the other provinces which are different from his *hukou* registration place” in the Sixth National Population Census of China (NBS, 2012). In the notice of “New standards to the city classification system” published by the State Council on Nov. 20, 2014, those people currently living in the

---

town (township or sub-district), but who have also left their original registered residence more than 6 months ago constitute the resident population (State Council, 2014). According to this notice, we define those people who left their original registered residence more than 6 months ago and crossed the provincial boundary as interprovincial migrants. There were a total of 930 interprovincial migration flows that occurred between 31 provinces, including autonomous areas and municipalities (Hong Kong, Macau, and Taiwan are not included, and intra-provincial migration flows are also excluded).



Figure 1 Location map of the study area

Population size as the regional mass characteristic of origins and destinations was introduced into gravity law by Zipf (1946) laying the foundation of quantitative migration research. Research work on migration has been actively developing since the time of Lee's (1966) conceptualization of migration and the "push and pull" theory. New classical economics, dual labor markets, and cumulative causes further complement and improve migration theories, with regional economic development levels, employment rates, individual incomes and development opportunities being considered as the "push and pull" factors that influence migration (Harris, & Todaro, 1970; Massey, 1990). These factors are influential at different temporal and spatial scales, and have been

empirically tested and compared over time (Greenwood, 1969; Cai, 1999; Fan, 2005; Shen, 2012). Following the previous studies, we choose population size, regional gross domestic production (GDP), registered unemployment rate, real wage rate, proportion of college enrollment in higher education, migration stock and distance variables to interpret the multilateral impacts of macro socio-economic conditions on the variation of interprovincial migration flows, as shown in Table 1. As each region can be treated as an origin or a destination, we use “O\_” or “D\_” to indicate this with each explanatory variable (e.g., O\_GDP, D\_GDP) to differentiate the role that the same explanatory variable plays in the migration models. If different variables are chosen for a region as an origin or a destination actor, it implies that prior assumptions are exerted on the same region, which may cause bias due to omitted or missing variables (Thomas-Agnan, & LeSage, 2015). We expect that higher GDP, real wage rate, college enrollment and migration stock at a destination will attract more immigrants, and the more population, higher unemployment rate and more migration stock at an origin will push more people to leave their original place of residence. The other variables will exert negative impacts on interprovincial migration activities.

**Table 1 Descriptions and expected effects of major explanatory variables**

Variable name	Descriptions	Expected effects	
		Variable at an origin	Variable at a destination
<i>pop</i>	Total population in 2010 (person)	+	-
<i>GDP</i>	Gross domestic production in 2010 (billion yuan)	-	+
<i>uemp</i>	Registered unemployment rate in urban area and town in 2010 (%)	+	-
<i>wage</i>	Real wage rate in urban area and town in 2010 (%)	-	+
<i>stu</i>	Proportion of college enrollments in higher education in 2010 (%)	-	+
<i>flows</i>	Interprovincial migration stock in 1995-2000 (person)	+	+
<i>Dist</i>	Railroad distance between capitals of provinces (kilometer)	-	

Source: NBS, 2002, 2011, 2012.

In this paper, we treat interprovincial migration flows as a dependent variable, and others as

---

explanatory variables. The data of population size, social and economic variables are drawn from the China Statistical Yearbook (NBS, 2011), and migration stocks are collected from the Fifth National Population Census (NBS, 2002).

### 3.2 Spatial OD model for the Chinese interprovincial migration system

The spatial patterns of interprovincial migration flows have been rather stationary in recent years. This means that less developed central and western populations migrate to developed eastern coastal areas (Ding, Liu, Cheng, & Zou, 2005; Deng, Liu, Cai, & Lu, 2014; Li, Liu, & Tang, 2015). According to the possible spatial dependencies that exist among origins, destinations and flows, we define three types of network weight matrices ( $W_o$ ,  $W_d$ ,  $W_w$ ) following LeSage & Pace (2008). Specifically,  $W_o = W \otimes I_n$ ,  $W_d = I_n \otimes W$ ,  $W_w = W \otimes W$ , in which  $W$  is an  $n \times n$  spatial weight matrix for the  $n$  regions, and  $\otimes$  represents a *Kronecker product*. The construction of spatial weight matrix  $W$  is based on the criteria of the common boundary. Considering that Hainan Province is an island, we define Guangdong Province as its neighbor because Hainan was under the jurisdiction of Guangdong before 1988. Since intra-provincial migration flows are larger than interprovincial migration flows, we do not consider the intra-provincial flows in this study. Accounting for the explanatory variables selected and three kinds of network weight matrices developed for this research, the specific spatial OD model of Chinese interprovincial migration flows is as follows:

$$\begin{aligned}
 y = & \rho_o W_o y + \rho_d W_d y + \rho_w W_w y + \alpha t_N + a_1 D\_pop + a_2 D\_GDP + \\
 & a_3 D\_wage + a_4 D\_uemp + a_5 D\_stu + a_6 D\_flows + b_1 O\_pop + \\
 & b_2 O\_GDP + b_3 O\_wage + b_4 O\_uemp + b_5 O\_stu + b_6 O\_flows + \gamma Dist
 \end{aligned} \tag{6}$$

Where  $y$  represents logged interprovincial migration flows between 2005 and 2010.  $W_o y$ ,  $W_d y$  and  $W_w y$  are the three spatial lags of the dependent variable corresponding to different network matrices. The descriptions of the other explanatory variables are listed in Table 1, and are

compared with equation (2),  $\beta_d = [a_1, a_2, \dots, a_6]$ ,  $\beta_o = [b_1, b_2, \dots, b_6]$ .

We first estimated the coefficients of the gravity model and spatial OD model. In order to compare the results from these two models, we then estimated the multilateral effects of the explanatory variables and conducted significance tests using the Monte Carlo simulation method. Under the MATLAB environment, the results of these two models are compared in Table 2.

**Table 2 Estimates for the gravity and spatial OD models of interprovincial migration flows in China (2005–2010)**

Variables	Gravity model		Spatial OD model			
	Bilateral effects	<i>t</i> value	Coefficients	<i>t</i> value	Multilateral effects	<i>t</i> value
<i>const</i>	-1.0026	-0.5529	-2.6240*	-1.6877	-3.2204	-1.4902
<i>D_pop</i>	-0.2527	-1.7736	-0.0142	-0.1232	-0.0314	-0.7387
<i>D_GDP</i>	0.5587***	3.9851	0.2076*	1.8285	0.4579***	5.2578
<i>D_wage</i>	1.0361***	4.7452	0.5564***	3.1511	1.2269***	6.7115
<i>D_ump</i>	-0.5202***	-3.3278	-0.3058**	-2.4435	-0.6744***	-3.8168
<i>D_stu</i>	-0.4109***	-3.2120	-0.3004***	-2.9186	-0.6626***	-6.5182
<i>D_flows</i>	0.5262***	9.0630	0.3834***	7.8925	0.8454***	12.2943
<i>O_pop</i>	0.9158***	6.3453	0.4422***	3.5926	0.9751***	13.0867
<i>O_GDP</i>	-0.3522***	-3.0305	-0.0813	-0.8283	-0.1792**	-2.2690
<i>O_wage</i>	-0.5262**	-2.2819	-0.5571***	-2.9173	-1.2285***	-6.8218
<i>O_ump</i>	0.1458	0.9810	-0.1164	-0.9590	-0.2567	-1.4547
<i>O_stu</i>	-0.0478	-0.3876	-0.1515	-1.5279	-0.3340***	-3.4703
<i>O_flows</i>	0.3550***	7.3797	0.1823***	4.5848	0.4020***	9.8791
<i>Distance</i>	-1.1374***	-23.3270	-0.6375***	-13.4866	-1.4058***	-19.9182
$\rho_o$			0.3796***	14.4158		
$\rho_d$			0.3705***	14.2417		
$\rho_w$			-0.2030***	-5.9458		
<i>Log-LIK</i>	-822.86		-652.70			
<i>AIC</i>	1673.72		1339.40			

Note: \* stands for "significant" at the significance level of 10%, \*\* at the significance level of 5%, and \*\*\* at the significance level of 1%.

From the above table, the preliminary findings are as follows: First, there was significant spatial dependence among Chinese interprovincial migration flows between 2005 and 2010. The estimates of the three network autocorrelation effects ( $\rho_o, \rho_d, \rho_w$ ) are statistically significant from zero, which implies that the independence assumption ( $\rho_o = \rho_d = \rho_w = 0$ ) in the gravity model is not

---

valid. The network effects of origins and destinations are significantly positive, which means that there are positive multilateral effects which exist among origins and destinations. Therefore, the migrants departing from the same origin will likely concentrate on a particular destination and the neighboring regions of that destination. Similarly, the migration flows that arrive at the same destination will in all likelihood originate from similar origin locations and their neighboring regions. For example, by using Shanghai as a destination, it can be seen that the amount of the population originating from Anhui, Jiangsu and Zhejiang Provinces is 2,232,000 (reflecting 45% of the total immigrant population of 4,934,000) during the period of 2005 and 2010. Similarly, by using Shandong Province as an origin, it is shown that the amount of migrants who arrive at Beijing and Tianjin is 576,000, and to the Yangtze River Delta (Jiangsu, Shanghai and Zhejiang) is 700,000 (reflecting 63% of the total emigrant population of 2,015,000) during the period of 2005 and 2010. From the perspective of migration selectivity, the network autocorrelation effects can be understood as spatial mimicking behavior. That is to say, the migrants adjacent to particular origins tend to arrive at the same destinations, and the migrants adjacent to these destinations tend to originate from the same origin locations.

At the same time, the flow-based effect ( $\rho_w$ ) is significantly negative, which implies that the migration flows originating from origin A to destination B are likely to decrease the migration flows from the neighboring regions of origin A to the neighboring regions of destination B. This is mainly due to the existence of the network autocorrelation effects of origins and destinations. For example, the number of emigrants from Anhui to Jiangsu is 1,546,000 in the period of 2005 to 2010. The migrating population from Henan (a neighbor of Anhui) to Shanghai (a neighbor of Jiangsu) is 473,000, and the population from Hubei (another neighbor of Anhui) to Shanghai is

---

only 247,000. Similarly, the flow-based network effects can be understood as spatial competition behavior, which implies that the more that migration flows from origin A to destination B, the less migration will flow from neighboring regions of origin A to the neighbors of destination B.

Second, the multilateral effects of regional socio-economic factors result in the spatial concentration of migrants in origins and destinations. When spatial dependencies are expressed explicitly in the spatial OD model, the total effects of most variables (in absolute values) have increased to some extent, which further strengthens the impacts of socio-economic variables on migration flows. Some developed regions hold more advantages in receiving migration flows, and other less developed regions in sending migration flows, which therefore adopt characteristics of Matthew effects (Wang, Pan, & Liu, 2012).

The influence of population size at the origin ( $O_{pop}$ ) on interprovincial migration flows is in line with the expected effect. On average, if the population size of a particular region increases by 1%, the population migrating to other provinces will increase 0.98%, which is 0.06% higher than that shown in the gravity model. This implies that the increase of the regional population will create positive spillover effects and drive neighboring populations to emigrate from their native homes. The effect of the population size of the destination ( $D_{pop}$ ) is negative and not significant, which indicates that population size at the provincial scale does not create attraction effects. This finding is in contrast with the current studies conducted by Wang, Pan, & Lu (2012), which may be due to the difference between the models and variables selected.

The effects of regional domestic gross production ( $GDP$ ) are in line with the outlined expectation. On average, if regional GDP increases by 1%, immigration flows will increase by 0.46%, and emigration flows will decrease by 0.18%, which are lower than the estimates of the

---

gravity model. The spillover effect of the GDP at destinations ( $D\_GDP$ ) is positive and the effects of the GDP at origins ( $O\_GDP$ ) is negative, which indicates that the increase in regional GDP will attract more foreigners and neighboring regions will attract more migrants. At the same time, the number of emigrants from origin locations will decrease and so do the numbers from neighboring regions. There is an asymmetric relationship in terms of origins and destinations.

The impact of real wage rates on migration flows are in accordance with the expectation. On average, if the real wage rate increases by 1%, immigration flows will increase by 1.23%, and emigration flows will decrease by 1.23%. This variable shows a symmetric characteristic when the region acts as an origin and a destination, which is different from the results of other variables and gravity models. The spillover effects of the real wage rate at destinations are significantly positive, which indicates that the regional real wage will improve neighboring regions' ability to attract foreigners, as well as the attractiveness of the destination. Similarly, the spillover effects of the real wage rate at origins are significantly negative, which shows that the increase of regional real wages will slow down migration flows from neighboring regions to an origin, as well as its own radial effects. The effect of this variable extends beyond that of the GDP variable.

The effect of the unemployment rate at origins on migration flows ( $O\_uemp$ ) is not in accordance with the expectation. On average, if the regional unemployment rate increases by 1%, immigration flows will decrease by 0.67%, and the emigration flows will decrease by 0.26%. This result is different from the expectation and the gravity model. The spillover effect of this variable is negative, which indicates that the increase of the unemployment rate will significantly decrease the attractiveness at destinations and their neighboring regions, but also decrease the ability to emigrate from origin locations and their neighboring regions.



---

The effects of college enrollment in the destination ( $D_{stu}$ ) on migration flows are not in line with the expectation. On average, if the proportion of college enrollment increases by 1%, the immigration flows will decrease by 0.66%, and emigration flows will decrease by 0.33%. The increase of the proportion of college students at destinations will improve the threshold of interprovincial migration, which acts in a similar manner to unemployment rate. Similarly, the increase of the proportion of college students at origin locations will also decrease migration flows to other regions. The spillover effects of the two variables are negative, which indicates that the number of college students will not only decrease the attractiveness of destinations and their neighboring regions for foreigners, but also prevent the population from moving outwards from origin locations and their neighboring regions.

The effects of migration stocks are in line with the expectation. On average, if the regional migration stocks increases by 1%, immigration flows will increase by 0.85%, and emigration flows will increase by 0.40%. The spillover effects of these variables are positive, indicating that the social network will improve bidirectional flows between different regions. It not only improves the attraction ability of destinations and their neighboring regions, but also drives the population of origin locations and their neighboring regions to move across provincial boundaries. The total effect of migration stocks at destinations ( $D_{flows}$ ) has increased by 60% compared with that of the gravity model, which is nearly twofold of that of the GDP variable.

The effect of distance on migration flows is as expected. On average, if the distance between two regions increases by 1%, migration flows will decrease by 1.41%. Compared with the estimates of the gravity model, the decay effect of distance has been enhanced in the spatial model, and ranks first among all the socio-economic variables. In the current stage of regional

---

development, distance is still the most important limiting factor during the process of migration. According to existing studies, the impact of distance on population migration will eventually decrease (Cai, & Qian, 2013). It seems reasonable from the coefficient estimate of the spatial model, but the explanation of its role in migration flows is yet not correct.

#### **4 Simulation of the impacts of the change in regional characteristics on interprovincial migration flows**

Intuitively, changes in the characteristics of a single region will influence flows into and out of this region, as well as flows into and out of the other regions through network connectivity. However, with the increase of the order of network matrices, the spillover effects of the region will attenuate in a geometric manner, and the main influence is restricted to the center of oscillation and its neighbors. In order to illustrate the impact of a particular change in the characteristic of a region on migration flows, we simulated the impacts of 5% increase of the GDP variable in Jiangsu on the entire interprovincial migration system. We list the main origins and destinations whose outflows and inflows have changed greatly due to change of GDP in Jiangsu in Table 3.

Influenced by the increase of the GDP in Jiangsu Province, migration flows from Jiangsu to other provinces have decreased to various extents. According to the simulation results, Shanghai will experience the most significant decrease in migration flows among all the destinations. Specifically, the number of migration flows from Jiangsu to Shanghai will decrease by 11,365 persons, occupying half of the total reduction. The effect on the emigration flows from Jiangsu to Tibet is the least significant, since it is located further away from Jiangsu. Concurrently, the emigration structure of the entire system has, to some extent, changed. The neighbors of Jiangsu,

---

(including Shanghai, Zhejiang, Anhui, Shandong, and Henan) have experienced the most significant changes. In general, most of the flows to other provinces will decrease to some extent, except in the case of the increase of flows to the Yangtze River Delta. Inevitably, those regions experiencing the most change will be concentrated on the neighborhoods of Jiangsu Province.

If the GDP of Jiangsu Province increases by 5%, the immigration flows arriving at Jiangsu from other provinces around the whole country will increase, to some extent. Specifically, the number of total increasing flows from Jiangsu to other provinces is 124,975, which is fivefold more than that of the total reduction due to the GDP increase in Jiangsu. Among all the origins, the number of emigration flows from Anhui Province is the largest, and is beyond one third of the total increase, followed by Henan and Sichuan. At the same time, the population arriving at the Yangtze River Delta will moderately increase. Zhejiang and Shanghai are experiencing the greatest influx of migrating populations, and the inflow to Shandong and Fujian will increase to some degree too. However, migration flows arriving at Anhui from other provinces cannot counteract the emigration flows that are exiting from this area. From the perspective of spillover effects, the neighbors of Jiangsu (except Anhui) have increased their attraction to foreigners, and this is due to the GDP increase in Jiangsu.

**Table 3 Simulation of migration flow changes from and to major regions following a 5% GDP increase in Jiangsu**

Province	Origins						Destinations					
	Jiangsu	Shanghai	Zhejiang	Anhui	Shandong	Henan	Jiangsu	Shanghai	Zhejiang	Anhui	Fujian	Shandong
Beijing	-1261	-17	-54	-142	-285	-20	236	55	9	4	0	15
Tianjin	-468	-5	-19	-46	-240	-7	147	31	6	2	0	15
Hebei	-241	-2	-12	-18	-65	4	1409	161	45	18	1	250
Shanxi	-227	-1	-11	-15	-22	-8	851	108	27	7	1	63
Inner Mongolia	-215	-1	-10	-13	-33	-4	346	57	16	4	1	94
Liaoning	-285	-3	-13	-25	-57	-4	611	151	31	6	1	93
Jilin	-92	-2	-6	-7	-21	-1	671	141	35	6	1	222
Heilongjiang	-123	-2	-5	-8	-30	-1	1156	237	60	8	2	487
Shanghai	-11365	0	2258	5464	703	1845	1431	0	-47	-1	-4	2
Jiangsu	0	1431	4706	44101	7010	17109	0	-11365	-2505	-799	-272	-972
Zhejiang	-2505	-47	0	2176	114	1295	4706	2258	0	32	-41	51
Anhui	-799	-1	32	0	12	106	44101	5464	2176	0	-50	164
Fujian	-272	-4	-41	-50	-15	20	2328	646	234	23	0	42
Jiangxi	-169	-7	-26	-19	-7	3	4455	1200	2215	43	167	36
Shandong	-972	2	51	164	0	501	7010	703	114	12	-15	0
Henan	-209	-3	-11	-22	-33	0	17109	1845	1295	106	20	501
Hubei	-247	-5	-21	-27	-13	4	6711	967	957	53	31	93
Hunan	-117	-3	-14	-9	-8	-1	3391	547	806	28	24	39
Guangdong	-818	-13	-68	-147	-72	-47	745	189	50	11	9	19
Guangxi	-92	-2	-16	-10	-7	-2	761	118	151	7	7	10
Hainan	-62	-1	-5	-7	-4	-1	98	23	6	2	1	4
Chongqing	-104	-3	-11	-5	-7	-1	2986	546	642	14	43	29
Sichuan	-199	-6	-21	-10	-15	-2	10866	1499	1350	38	88	103
Guizhou	-100	-2	-14	-7	-6	-1	4602	355	1630	28	50	29
Yunnan	-103	-3	-20	-8	-7	-2	2357	168	448	25	11	53
Tibet	-9	0	-1	-1	-1	-0	31	3	1	1	0	1
Shaanxi	-217	-3	-15	-14	-20	-8	3456	302	182	14	6	58
Gansu	-121	-2	-10	-5	-7	-2	1699	228	68	6	2	50
Qinghai	-84	0	-4	-4	-5	-2	231	27	9	2	0	15
Ningxia	-61	0	-4	-6	-7	-2	117	24	6	1	0	9
Xinjiang	-335	-5	-13	-28	-27	-16	358	69	16	3	1	28
Total	-21873	1290	6600	51250	6824	20757	124975	6756	10033	-296	87	1603

Note: “+” stands for the increase of migration flows; “-” stands for the decrease of migration flows; “0” indicates no change or intra-provincial migration flows; the “province” column stands for 31 destinations (corresponding to ‘origins’ column); or 31 origins (corresponding to ‘destinations’ column).

---

## 5 Conclusions

In this paper, we constructed an unrestricted spatial OD model for interprovincial migration flows based on the data extracted from the Sixth National Population Census of China. By incorporating three different network dependencies, we extended the “bilateral effects” of origins and destinations in the gravity model to include “multilateral effects.” In doing so, the network autocorrelation effects among interprovincial migration flows between 2005 and 2010 were estimated, and the impacts of the changes in regional socio-economic factors on the entire migration system were simulated. Several preliminary findings were drawn:

First, there are significant spatial dependencies that exist among interprovincial migration flows. From the perspective of migration selectivity, positive dependence can be understood as spatial imitation behavior. That is, the migrants adjacent to particular origins tend to arrive at the same destinations, or the migrants adjacent to these destinations tend to originate from the same origin locations. Otherwise, the negative dependence can be regarded as spatial competition behavior.

Secondly, multilateral effects of regional socio-economic factors will induce spatial concentration characteristics of interprovincial migration flows. The absolute total effects of most variables increase (to some extent) when spatial autocorrelation effects are expressed explicitly, which further strengthen the influences of socio-economic factors on migration flows. The distance decay effect in absolute value ranks first among all the explanatory variables. Real wage rates at origins and destinations show symmetric characteristics, which are different from other variables. The impact of migration stock at destinations is beyond that of the GDP variable. If the unemployment rate decreases, migration flows will increase, indicating that active labor markets

---

will urge the population to move across space. The effect of GDP is manifested in attracting foreigners to destinations or their neighbors. The increase of the proportion of college enrollments will decrease migration opportunities by improving the threshold for moving. This is an area which requires further investigation.

Finally, changes to the characteristics of a single region will potentially influence the entire migration system. The most significant changes to migration flows occurred around the centre of oscillation and its neighbors. From the simulation results generated through the 5% GDP increase in Jiangsu Province, all emigrating flows from Jiangsu to the other provinces decreased to some extent, and all immigrating flows from the whole country to Jiangsu increased to some degree. Among all the related regions, the most significant impacts due to GDP change in Jiangsu occurred in Shanghai and Anhui. In addition, migration structures from and to the neighboring regions of Jiangsu changed greatly.

### **Acknowledgments**

This research was supported by the National Natural Science Foundation of China (Grant Number: 41271388), Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), and Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application. The authors thank Yaping WANG who helped to collect the data and conduct necessary calculations.

### **References**

- Black, W. R. (1992). Network autocorrelation in transport network and flow systems. *Geographical Analysis*, 24(3): 207-222.
- Cai, F. (1999). Spatial patterns of migration under China's reform period. *Asian and Pacific Migration Journal*,

---

8(3): 313-327.

Cai, F., & Wang, D. (2003). Migration as marketization: What can we learn from China's 2000 census data ? *China Review*, 3(2): 73-93.

Cai, G., & Qian, J. (2013). The origin of spatial correlation: A theoretical model and empirical evidence. *China Economic Quarterly*, 12(3): 869-893. [in Chinese]

Chun, Y. (2008). Modeling network autocorrelation within migration flows by eigenvector spatial filtering. *Journal of Geographical Systems*, 10(4): 317-344.

Curry, L. (1972). A spatial analysis of gravity flows. *Regional Studies*, 6(2): 131-147.

Deng, Y., Liu, S., Cai, J., & Lu, X. (2014). Spatial pattern and its evolution of Chinese provincial population and empirical study. *Acta Geographica Sinica*, 69(10): 1473-1486. [in Chinese]

Ding, J., Liu, Z., Cheng, D., Liu, J., & Zou, J. (2005). Areal Differentiation of Inter-provincial migration in China and Characteristics of the flow field. *Acta Geographica Sinica*, 60(1): 106-114. [in Chinese]

Fan, C. C. (2005). Modeling interprovincial migration in China, 1985-2000. *Eurasian Geography and Economics*, 46(3): 165-184.

Goodchild, M. F. (1987). *Spatial autocorrelation: Concepts and Techniques in Modern Geography*. Norwich: Geo- Books.

Greenwood, M. J. (1969). An analysis of the determinants of geographic labor mobility in the United States. *The Review of Economics and Statistics*, 51(2): 189-194.

Griffith, D. A., & Jones, K. (1980). Explorations into the relationships between spatial structure and spatial interaction. *Environment and Planning A*, 12: 187-201.

Harri,s J. R., & Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *American Economic Review*, 60(1): 126-141.

- 
- IOM [International Organization for Migration]. (2008). *World Migration Report 2008*. Geneva: IOM.
- Massey, D. S. (1990). Social structure, household strategies, and the cumulative causation of migration. *Population Index*, 56(1): 3-26.
- Lee, E. S. (1966). A theory of migration. *Demography*, 3(1): 47-57.
- LeSage, J. P., & Pace, R. K. (2008). Spatial econometric modeling of origin-destination flows. *Journal of Regional Science*, 48(5): 941-967.
- LeSage, J. P., & Pace, R. K. (2009). *Introduction to Spatial Econometric*. Boca Raton, FL: CRC Press, Taylor & Francis Group.
- LeSage, J. P., & Thomas-Agnan, C. (2015). Interpreting spatial econometric origin-destination flow models. *Journal of Regional Science*, 55(2): 188-208.
- Li, Y., Liu, H., & Tang, Q. (2015). Spatial-temporal patterns of China's interprovincial migration during 1985-2010. *Geographical Research*, 34(6): 1135-1148. [in Chinese]
- National Bureau of Statistics (NBS). (2002). *Tabulation on the 2000 population census of the People's Republic of China*. Beijing: China Statistics Press. [in Chinese]
- National Bureau of Statistics (NBS). (2011). *China Statistical Yearbook*. Beijing: China Statistics Press, <http://www.stats.gov.cn/tjsj/ndsj/2011/indexce.htm>. [in Chinese]
- National Bureau of Statistics (NBS). (2012). *Tabulation on the 2010 population census of the People's Republic of China*. Beijing: China Statistics Press. [in Chinese]
- Pace, R. K., & LeSage, J. P., & Zhu, S. (2013). Interpretation and computation of estimates from regression models using spatial filtering. *Spatial Economic Analysis*, 8(3): 352-369.
- Roy, J. R., & Thill, J. (2004). Spatial interaction modeling. *Papers in Regional Science*, 83(1): 339-361.
- Sen, A., & Smith, T. E. (1995). *Gravity Models of Spatial Interaction Behavior*. Berlin-Heidelberg:



- 
- Springer-Verlag.
- Smith, T. E. (1975). A choice theory of spatial interaction. *Regional Science and Urban Economics*, 5(2): 137-176.
- Shen, J. (2012). Changing patterns and determinants of interprovincial migration in China 1985-2000. *Population, Space and Place*, 18(3): 384-402.
- Stillwell, J. C. H., & Congdon, P. (1991). *Migration Models: Macro and Micro Approaches*, London: Belhaven Press.
- Stillwell, J. (2005). Inter-regional migration modeling: A review and assessment. Paper prepared for the 45<sup>th</sup> Congress of the European Regional Science Association, The Netherlands, 23-27 August.
- The State Council, People's Republic of China. (2014). China to apply new city classification standards. [http://english.gov.cn/policies/latest\\_releases/2014/11/25/content\\_281475015213546.htm](http://english.gov.cn/policies/latest_releases/2014/11/25/content_281475015213546.htm).
- Thomas-Agnan, C., & LeSage, J. P. (2014). Spatial econometric OD-flow models. *Handbook of Regional Science*, Berlin-Heidelberg: Springer-Verlag, 2014, 1653-1673.
- Wilson, A. G. (1967). A statistical theory of spatial distribution models. *Transportation Research*, 1(3): 253-269.
- Wang, G., Pan, Z., Lu, Y. (2012). China's inter-provincial migration patterns and influential factors: Evidence from year 2000 and 2010. *Chinese Journal of Population Science*, (5): 2-13. [in Chinese]
- Yang, Y., Chan, K. W., & Liu, T. (1999). Population migration in China: Multi-region model and empirical analysis. *Chinese Journal of Population Science*, (4): 20-26. [in Chinese]
- Yu, W., Pu, Y., Chen, G., & Wang, J. (2012). Spatial analysis of the patterns and mechanism of inter-provincial migration flows in China. *Geography and Geo-Information Science*, 28(2): 44-49. [in Chinese]
- Zipf, G. K. (1946). The P<sub>1</sub>P<sub>2</sub>/D hypothesis: On the intercity movement of persons. *American Sociological Review*, 11(6): 677-686.