DEMOGRAPHIC RESEARCH

Research Article

A spatial dynamic panel approach to modelling the space-time dynamics of interprovincial migration flows in China

Yingxia Pu

Xinyi Zhao

Guangqing Chi

Jin Zhao

Fanhua Kong

Contents

A spatial dynamic panel approach to modelling the space-time dynamics of interprovincial migration flows in China

Yingxia Pu[1]

Xinyi Zhao[2]

Guangqing Chi[3]

Jin Zhao[4]

Fanhua Kong[5]

Abstract

BACKGROUND
Migration plays an increasingly crucial role in regional growth and development. However, most empirical studies fail to simultaneously capture the spatial and temporal aspects of migration flows, and thus fail to reveal how space-time dynamics shape path-dependent migration processes.

OBJECTIVE
This study attempts to incorporate space-time dimensions into a traditional gravity model and to measure the impact of regional socioeconomic changes on migration flows. Doing so allows us to better understand the space-time dynamics of complex migration processes.

METHODS
We construct a spatial dynamic panel data model for dyadic migration flows and decompose origin, destination, and spillover effects into their contemporaneous, short-, and long-term components. We then apply this approach to panel data of interprovincial migration flows in China from 1985 to 2015.

[1] School of Geography and Ocean Science, Nanjing University, Nanjing, China.
Email: yingxiapu@nju.edu.cn.
[2] School of Geography and Ocean Science, Nanjing University, Nanjing, China.
[3] Department of Agricultural Economics, Sociology, and Education, Population Research Institute, and Social Science Research Institute, Pennsylvania State University, University Park, PA, USA.
[4] Department of Mathematics, Nanjing University, Nanjing, China.
[5] International Institute for Earth System Science, Nanjing University, Nanjing, China.

RESULTS
The empirical results indicate that space-time interactions are crucial to the formation and evolution of migration flows. Population size and age structure play an important role in the Chinese interprovincial migration process, particularly dominated by significant and positive spillover effects in the contemporaneous period. The relationship between development and migration tends to take an inverted U-shaped curve, a fact not revealed in nonspatial models.

CONCLUSIONS
This fine-grained decomposition of origin, destination, and spillover effects over contemporaneous, short-, and long-term periods can help understand the spatial and temporal mechanisms of migration and its driving factors.

CONTRIBUTION
We propose a spatial dynamic panel data model for migration flows and estimate space-time effects of regional explanatory variables. This method could be applied to model other flow data such as trade and transportation flows.

1. Introduction

In recent times, migration has played an increasingly prominent role in shaping regional population growth, socioeconomic transition, and regional resilience (Bijak 2011; White 2016; Biagi et al. 2018). Regional economic, social, and cultural factors affect the relative propulsion and attractiveness of origins and destinations, as well as the highly selective and asymmetric migration processes (Lee 1966; Williamson 1988; Rogers 1990; Stillwell 2008; Shen 2016, 2017). Thus, measuring the impact of regional origin- and destination-specific characteristics on migration flows can not only help to identify important determinants and dynamic mechanisms of migration, but also to understand the vulnerability and resilience of this complex network (Griffith and Chun 2015; Crown, Jaquet, and Faggian 2018).

   More importantly, migration is a path-dependent process: Inter-personal relations, through network and herd effects, facilitate current and subsequent migration. When previous migrants provide information about shelter and work assistance or generally reduce the difficulty and stress of relocation to a foreign environment, pioneer migration leads to network externalities in migration (Petersen 1958; Lee 1966; Church and King 1993; Baláž and Williams 2007). In the face of uncertainty or imperfect information about destinations, potential migrants are more likely to follow others, thus exhibiting stronger herd behaviours (Epstein 2008). It is these endogenous feedback mechanisms operating on the network that give migration processes their own

momentum (De Haas 2010a). However, modelling the dynamics of migratory phenomena over space and time has not been investigated extensively (Mitze 2012, 2016; Mavroudi and Nagel 2016; Bernard 2017; Beenstock and Felsenstein 2019).

Most empirical migration studies have adopted the traditional gravity models to examine the push and pull forces of origins and destinations, but such studies fail to reveal the intrinsic mechanisms of migration systems (Griffith and Jones 1980; Chun 2008; LeSage and Pace 2008, 2009; Griffith 2009; Chun and Griffith 2011). Fortunately, recent developments in spatial econometric interaction models have laid a solid foundation for modelling origin-destination (OD) flows through various network structures (Patuelli and Arbia 2016). This provides an opportunity to capture the spatial dependence of migration flows and to investigate the spillover effects implicit in migration processes (LeSage and Pace 2008, 2009; LeSage and Fischer 2016). However, the network and other types of spatial effect are always interpreted as outcomes of a long-run equilibrium: Time plays no explicit role in such cross-sectional models (LeSage and Thomas-Agnan 2015).

In this paper, we present a spatial dynamic panel data model to study migration flows that introduces an individual time lag (capturing temporal dependence), various spatial lags (accounting for spatial dependence), and cross-product terms, thereby integrating space-time dependence and traditional gravity models. This allows us to explore the path-dependent migration process: how a change in a single region's characteristics will produce contemporaneous and future responses to that region's inflows and outflows as well as to those of other regions.

We then apply our proposed model with its derived space-time effects to analyse the dynamic mechanisms behind interprovincial migration flows in China from 1985 to 2015. Since the start of the reform and opening-up policy in late 1978, internal migration has been on the rise, contributing enormously to China's 'economic miracle' and rapid urbanization (United Nations 2012; Liang and Song 2016). According to the 1990 census and the 2015 1% population sample survey, the number of interprovincial migrations soared from 11.06 million between 1985 and 1990 to 53.28 million from 2010 to 2015, roughly a fivefold increase during the three decades (NBSC 1993, 2016).

The determinants driving large-scale interprovincial migration in China have received much attention (Cai and Wang 2003; Li 2004; Fan 2005; Shen 2016, 2017). Institutional factors (i.e., the relaxation and reform of the household registration or *hukou* system) had a large impact on the initial increase in migration in the 1980s, while the subsequent increase since 1990 has been largely driven by China's rapid and unbalanced economic development (Shen 2015). In addition, the global phenomenon of social networks plays an important role in people's decision to undertake migration and in the choice of specific destinations (Fan 2005).

To conduct a detailed space-time analysis of interprovincial migration flows in China, we choose four explanatory variables to characterize regional socioeconomic conditions at the origin and destination locations – population size, gross domestic product (GDP), age structure, and real wage level –, and use railway travel time between different provinces to measure improved transportation infrastructure. We also consider temporal, spatial, and space-time dependencies among migration flows to characterize the potential existence of a path-dependent process. Moreover, we compare the results from the spatial dynamic panel data model with those of a nonspatial model to learn more about how space-time dynamics influence the evolution of a migration system, and identify the most significant factors dominating interprovincial migration in China from 1985 to 2015. The empirical results may have important implications for understanding the evolution of regional migration systems and for pursuing sustainable regional development planning.

This paper proceeds as follows. Section 2 reviews recent developments in migration theories and migration modelling methods. Section 3 introduces our general dynamic panel data model and spatial dynamic panel data model for migration flows and derives the space-time effects for our spatial model. In section 4 we conduct an empirical space-time analysis to uncover the mutual dynamics of China's increasing interprovincial migration flows from 1985 to 2015. Section 5 concludes the manuscript and proposes directions for future research.

2. Literature review

Migration is a path-dependent process across space and over time (Petersen 1958). Not only do pioneer migrants (or migration stocks) help break the ice and clear the way for later migrants, but neighbouring migration flows also influence potential migrants' decision-making. The forces behind this semi-automatic migration process can be further ascribed to two types of effects (Bauer, Epstein, and Gang 2007; Epstein 2008: 568): network effects, or "I will go where my people are, since they will help me," and herd effects, or "I will go where I have observed others go, because those who went before me cannot be wrong, even though I would have chosen to go elsewhere."

The existence of positive network externalities explains migration clusters, when previous migrants provide information and help for followers (Bauer, Epstein, and Gang 2007; Epstein 2008). Over time, increase in migration stocks may aggravate competition for jobs, leading to negative externalities. However, even if network effects are not present, potential migrants may follow the herd and move to the places recent migrants have been observed to choose, thus producing mass migration. Herd effects

may be more pronounced when people believe that others will follow prior migrants. That said, herd and network effects can interact with each other in migration processes.

A popular approach to testing the hypothesis that past migration may influence present migration patterns is to introduce migration stock as an explanatory variable in traditional gravity models (Greenwood 1969, 1970; Kau and Sirmans 1979; Fan 2005). The positive coefficient of migration stock supports the idea that information from past migrants (friends and relatives) stimulates current migration, and the failure to include this variable could result in other variables being overestimated or obscured (Greenwood 1969, 1970). Further research indicates that recent migrants exert a dominant influence on future migrants, and migration stock may exert a negative influence on current migration in the adjustment process towards equilibrium (Laber 1972; Kau and Sirmans 1979). While gravity models have been widely employed in migration modelling since the 1940s (Zipf 1946; Sen and Smith 1995; Stillwell 2008), they mainly focus on the push and pull forces of origins and destinations over a given period. The dynamic relationship implied in migration processes needs to be modelled explicitly.

With the growing availability of space-time migration flow data, panel data analysis provides an opportunity to investigate the evolution mechanisms of observed migration patterns (Odland and Shumway 1993; Rainer and Siedler 2009; Etzo 2011; Bellos and Subasat 2012; Chort and de la Rupelle 2016; Mátyás 1997, 2017). The temporal dependence structures among migration flows at different time periods can be accommodated as a time-lagged dependent variable in dynamic panel gravity models. For example, Davidescu et al. (2017) identify the main determinants of Romanians' migration and confirm that previous migration has a positive impact on later flows. Moreover, panel data analysis can improve the precision of model estimates with larger sample sizes as well as with the expression of time effects (Elhorst 2010, 2012, 2014). However, the feedback effects among migration flows are largely ignored because of the independence assumption of interrelated spatial units implicit in panel data models.

The clustered patterns of migration flows can be measured in terms of spatial or network autocorrelation in spatial statistics and spatial econometrics (Plane and Rogerson 1986; Fotheringham 1991; Rogers et al. 2002; Cushing and Poot 2003; Anselin, Le Gallo, and Jayet 2008; Zimmer 2008; Chun 2008; LeSage and Pace 2008, 2009; Chun and Griffith 2011; Griffith and Fischer 2013, 2017; Lamonica and Zagaglia 2013; Bavaud, Kordi, and Kaiser 2018). If the number of outflows generated by origins (or inflows received by destinations) is enhanced by neighbouring origins (or neighbouring destinations), positive spatial autocorrelation among migration flows occurs (Griffith and Jones 1980; Black 1992). LeSage and Pace (2008) further define three types of spatial dependence to distinguish different migratory phenomena: Any migration flow is influenced by (1) flows from neighbouring origins to the same

destination, or origin-based dependence; (2) flows from the same origin to neighbouring destinations, or destination-based dependence; and (3) flows from neighbouring origins to neighbouring destinations, or origin-to-destination-based dependence. The herd and network effects provide a rationale for the observed spatial dependence among origins and/or destinations (Bauer, Epstein, and Gang 2007; Epstein 2008).

Today, there are two approaches to accommodating spatial dependence structures in migration modelling. The first is eigenvector spatial filtering, which filters origin/destination dependence by abstracting eigenvectors from a spatial weight matrix as additional covariates within Poisson-type regression models (Chun 2008; Griffith and Chun 2015). The second is spatial econometric interaction modelling, which explicitly expresses spatial dependence in spatial econometric interaction models (LeSage and Pace 2008, 2009; LeSage and Fischer 2016). In this study we focus on the second approach because of its ability to quantify spatial dependence and associated spillover effects among migration flows.

The incorporation of space and time dimensions into migration studies is promising. Mitze (2012) discusses several methods for spatially updating dynamic panel models, including spatial lag and spatial Durbin model specifications, and demonstrates the potential role of spatial autocorrelation in the analysis of German internal migration flows since re-unification. Subsequently, Mitze (2016) applies a spatial dynamic panel data model to link outflows with regional labour market signals and highlights the importance of mutual dynamic features in German internal migration between 1993 and 2009. Although the extension of spatial dynamic panel models to the case of dyadic flow data is straightforward, the interpretation of effects estimates of explanatory variables is not explicit, due to the existence of numerous responses triggered by changes in regional characteristics in an interwoven network of migration.

To better understand the feedback mechanisms behind migration processes, we should analyse and interpret models' parameter estimates correctly and roundly. Recently, scholars have expressed some caveats regarding employing parameter estimates to reflect the responses of the dependent variable to changes in explanatory variables over space and time (Debarsy, Ertur, and LeSage 2012; Elhorst 2012, 2014; LeSage and Thomas-Agnan 2015; LeSage and Fischer 2016). As such, the coefficient estimates for dynamic space-time panel models cannot be directly interpreted as partial derivatives that relate the responses of the dependent variable completely to changes in the explanatory variable (Debarsy, Ertur, and LeSage 2012). Instead, these models adopt direct and indirect effects, which are the own- and cross-partial derivatives associated with explanatory variables, as well as contemporaneous, short-, and long-term effects. To interpret spatial econometric dyadic flow models, LeSage and Thomas-Agnan (2015) derive origin, destination, intraregional, and spillover effects (or network effects) arising from changes in regional characteristics. Considering the expression of

space-time relationships in general panel data models and various types of effect estimates from spatial econometric interaction models, we need to make more effort to explore the mutual space-time effects of migration flows in the space-time dimension.

To address these challenges in modelling and interpreting migration processes, this study accomplishes the following: (1) We present a spatial dynamic panel data model of migration flows to uncover the space-time dynamics underlying path-dependent migration processes. Temporal, spatial, and space-time dependencies among migration flows can be accommodated in our proposed model. (2) We derive the expression of origin, destination, and spillover effects of explanatory variables as well as their corresponding decomposition of contemporaneous, short-, and long-term effects. These effects can be used to quantify the magnitude of herds and network effects implicit in migration processes, thus bridging the gap between migration theories and applied methodologies. (3) We apply our model and its associated effect estimates to explain the space-time dynamics in China's interprovincial migration from 1985 to 2015, which can be extended to similar OD flow studies as well as to other countries.

3. Methods

3.1 General dynamic panel data model

Traditional gravity or spatial interaction models focus on the impacts of origin- and destination-specific characteristics and the distance-decay effect on observed flows (Sen and Smith 1995; Mátyás 1997). Given this starting point, this study considers temporal dependence between adjacent time periods. For simplicity, we assume that migration flows in a given time period are dependent only on the immediate past period values, regardless of subsequent periods, as shown in Figure 1.

Introducing a first-order time lag of the dependent variable into the traditional gravity model allows us to determine a dynamic panel data model for migration flows as shown in Equation (1) (Parent and LeSage 2010, 2012):

$$Y = \phi Y_{-1} + \alpha \iota_{N \times T} + X_o \beta_o + X_d \beta_d + \gamma G + \eta,$$
$$\eta = \mu_N + \varepsilon_{N \times T}, \tag{1}$$

where $Y = (y_1', \dots, y_T')'$ consists of T time periods of migration flows and $y_t = (y_{1t}, \dots, y_{Nt})'$ is an $N \times 1$ $(N = n^2)$ column vector of regional gross flows constructed by stacking the columns of the $n \times n$ flow matrix for n regions at the tth time period; $Y_{-1} = (y_0', \dots, y_{T-1}')'$ represents the first-order time lag of Y in the same manner; $\iota_{N \times T}$ is an $NT \times 1$ column vector of ones; $X_o = (x_{o1}', \dots, x_{oT}')'$ and

$X_d = (x_{d1}', \ldots, x_{dT}')'$ are the explanatory variables at origins and destinations covering T time periods where $x_{ot}(x_{dt})$ is formed by the Kronecker product $x_t \otimes \iota_n$ ($\iota_n \otimes x_t$), x_t is an $n \times k$ explanatory variable matrix for n regions at the tth time period, resulting in an $n^2 \times k$ block matrix as the origin (destination) variable for all the flows; and $G = (g_1', \ldots, g_T')'$, where g_t represents the distance between all origins and destinations at the tth time period, in the form of time-invariant Euclidean distances or time-variant migration costs. The error term vector η is decomposed into random individual effects μ_N and the disturbance vector $\varepsilon_{N \times T}$. The $N \times 1$ column vector μ_N captures individual heterogeneities among N migration flows with $\mu_i \sim N(0, \sigma_\mu^2)$, which change only with the individuals and are constant over time. The $\varepsilon_{N \times T}$ is assumed to be uncorrelated with μ_N and is identically distributed with zero mean and variance $\sigma_\varepsilon^2 I_N$. α, β_o, β_d, γ, and ϕ are the parameters to be estimated: α is the intercept, β_o and β_d reflect the impacts associated with the origin- and destination-specific characteristics, γ denotes the impact of distance between different regions, and ϕ represents the magnitude of temporal dependence.

Figure 1:   The temporal dependence structure of migration flows
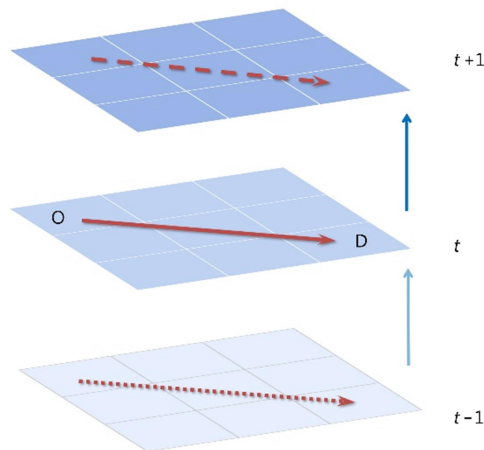
Note: The OD migration flow in the *t*th time period in a red solid line is influenced by the previous flow in time period *t*–1 in a red dotted line, and will exert a potential impact on the future flow in time period *t*+1 in a red dashed line.

A dynamic panel data model results in a greater degree of freedom and hence increases estimation efficiency (Elhorst 2012, 2014). Moreover, the model proposed

here accounts for more complicated temporal dependence and individual effects than the conventional gravity model can capture.

3.2 Spatial dynamic panel data model

As explained in Section 1, migration is a path-dependent process. Any migration flow is influenced not only by origin- and destination-specific characteristics, but also by the same flow during the previous period (temporal dependence), the neighbouring flows in the current period (spatial dependence), and the neighbouring flows in the previous period (space-time diffusion dependence) (Parent and LeSage 2010, 2012; Mitze 2016). In short, unlike unidirectional temporal dependence (Figure 1), space-time dependence is multidimensional and multidirectional in nature (Figure 2).

Figure 2:   The space-time dependence structure of migration flows
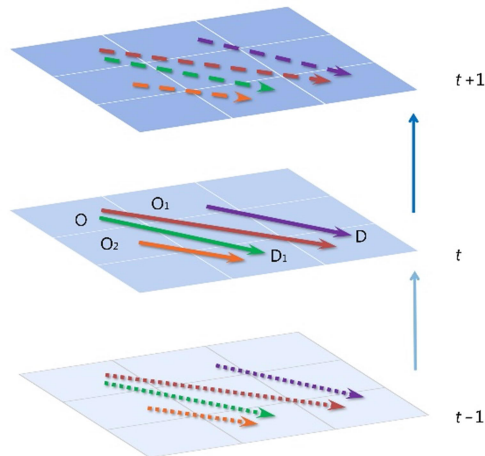
Note: The OD migration flow in the *t*th time period in a red solid line is influenced not only by the previous flow in time period *t*–1 in a red dotted line (capturing temporal dependence) and the current neighbouring flows (O_1D, OD_1 and O_2D_1) in purple, green, and orange solid lines (representing origin-, destination-, and origin-to-destination-based spatial dependence, respectively), but also by the previous neighbouring flows in purple, green, and orange dotted lines (representing origin-, destination-, and origin-to-destination-based space-time dependence, respectively). Furthermore, the current flows will exert potential impacts on the flows in future time period *t*+1 in dashed lines of different colours.

In addition to temporal dependence, our model incorporates origin-, destination-, and origin-to-destination-based spatial dependence, as proposed by LeSage and Pace

(2008, 2009). Assuming W is the $n \times n$ spatial weight matrix with zeros on the diagonal, these three types of spatial dependence are defined by the corresponding network weight matrices W_o, W_d, and W_w: $W_o = W \otimes I_N, W_d = I_N \otimes W$, and $W_w = W \otimes W$ (Chun 2008; LeSage and Pace 2008, 2009; LeSage and Llano 2013; Griffith, Fischer, and LeSage 2017).

Based on the spatial dependence of migration flows, we then define another three types of space-time diffusion dependence. If a migration flow in a given time period is influenced by flows from neighbouring origins to the same destination in the time period immediately preceding it, this can be understood as origin-based space-time dependence. Similarly, destination- and origin-to-destination-based space-time dependence of migration flows can be accordingly defined.

We now extend the general dynamic panel data model with the above spatial dependence and space-time diffusion dependence to derive a spatial dynamic panel data model of migration flows (Equation 2):

$$
\begin{aligned}
Y = {} & \rho_o(I_T \otimes W_o)Y + \rho_d(I_T \otimes W_d)Y + \rho_w(I_T \otimes W_w)Y + \phi Y_{-1} + \theta_o(I_T \otimes W_o)Y_{-1} \\
& + \theta_d(I_T \otimes W_d)Y_{-1} + \theta_w(I_T \otimes W_w)Y_{-1} + \alpha\iota_{N \times T} + X_o\beta_o + X_d\beta_d \\
& + \gamma G + \eta,
\end{aligned}
$$
$$
\eta = \mu_N + \varepsilon_{N \times T}, \tag{2}
$$

where I_T is the T-dimensional identity matrix; the Kronecker products $I_T \otimes W_o, I_T \otimes W_d$, and $I_T \otimes W_w$ represent the space-time weight matrices for all migration flows over T time periods; $(I_T \otimes W_o)Y, (I_T \otimes W_d)Y$, and $(I_T \otimes W_w)Y$, accordingly, represent origin-, destination-, and origin-to-destination-based spatial lags of migration flows in corresponding time periods; $(I_T \otimes W_o)Y_{-1}$, $(I_T \otimes W_d)Y_{-1}$, and $(I_T \otimes W_w)Y_{-1}$ represent origin-, destination-, and origin-to-destination-based space-time lags of migration flows. Like the scalar parameter ϕ of temporal dependence, (ρ_o, ρ_d, ρ_w) and $(\theta_o, \theta_d, \theta_w)$ measure the strength of the spatial dependence and space-time diffusion dependence, respectively.

This study adopts the Bayesian Markov chain Monte Carlo (MCMC) sampling method to estimate the associated parameters, thus avoiding multidimensional numerical integration in maximum likelihood (ML) or quasi-maximum likelihood (QML) methods (Chib and Carlin 1999; LeSage and Pace 2009; Elhorst 2010, 2012, 2014). Specifically, we use Bayesian updating schemes based on the joint posterior distribution of the parameters of interest in the manner of sampling from their posterior distributions (Bhargava and Sargan 1983; Parent and LeSage 2010, 2012; LeSage 2014). The joint posterior distribution is derived as follows:

$$p\big(\alpha,\beta_o,\beta_d,\gamma,\sigma_\varepsilon^2,\sigma_\mu^2,\rho_o,\rho_d,\rho_w,\phi,\theta_o,\theta_d,\theta_w|Y\big) \propto f(Y|\Upsilon,\mu_N,\sigma_\varepsilon^2,\rho,\phi,\theta)p(\mu_N|\sigma_\mu^2)$$
$$p(\sigma_\mu^2)p(\Upsilon)p(\rho)p(\phi)p(\theta)p(\sigma_\varepsilon^2),\,(3)$$

where $\Upsilon = (\alpha,\beta_o{}',\beta_d{}',\gamma)$, $\rho = (\rho_o,\rho_d,\rho_w)$, and $\theta = (\theta_o,\theta_d,\theta_w)$; $f(Y|\cdot)$ represents the likelihood function; and $p(\cdot)$ represents the prior distribution of each parameter.

The samples drawn from the set of conditional posterior distributions by the MCMC sampling method are used to estimate the associated parameters and their significance. The specific MCMC sampling and updating scheme for the spatial dynamic panel model are presented in the Appendix.

3.3 Interpreting the space-time effects of migration processes

For a spatial dynamic panel data model, a change in regional characteristics not only has a direct influence on the outflows and inflows from and to this region, but also has an indirect effect on the flows to and from other regions in the migration network over space and time. Therefore, there is a need to quantify various responses in the network from a local perspective (Golgher and Voss 2016). Following LeSage and Thomas-Agnan (2015), we extend the scalar summary measures of origin, destination, intraregional, and spillover effects to a space-time context and decompose them into contemporaneous, short-, and long-term components to reveal the changing impacts over time.

A change in a typical region at any time t will potentially bring about transitory impacts on the flows during the period $t + T$ to different degrees. Specifically, total effects (*TE*) are the impacts on all flows in the manner of $n^2 \times n$ partial derivatives shown in Equation (4) (Debarsy, Ertur, and LeSage 2012; LeSage and Thomas-Agnan 2015):

$$TE = \begin{pmatrix} TE_1 \\ TE_2 \\ \vdots \\ TE_n \end{pmatrix} = \begin{pmatrix} \partial Y_{1(t+T)}/\partial X_{1(t)}^r \\ \partial Y_{2(t+T)}/\partial X_{2(t)}^r \\ \vdots \\ \partial Y_{n(t+T)}/\partial X_{n(t)}^r \end{pmatrix} = D_T \begin{pmatrix} Jo_1\beta_o^r + Jd_1\beta_d^r \\ Jo_2\beta_o^r + Jd_2\beta_d^r \\ \vdots \\ Jo_n\beta_o^r + Jd_n\beta_d^r \end{pmatrix},$$
$$D_T = (-1)^T (B^{-1}C)^T B^{-1},$$
$$B = I_{n^2} - \rho_o W_o - \rho_d W_d - \rho_w W_w,$$
$$C = -(\phi I_{n^2} + \theta_o W_o + \theta_d W_d + \theta_w W_w),$$

(4)

where B represents the space filter and C the space-time diffusion filter, allowing producing the true elasticities of different factors after introducing space-time

interactions. Jo_i is an $n \times n$ matrix of zeros with the ith column equal to ι_n, and Jd_i is an $n \times n$ matrix of zeros with the ith row equal to ι'_n ($i = 1, \dots, n$). Thus, the rows of the matrix TE_i ($\partial Y_{i(t+T)}/\partial X^r_{i(t)}$) record the impacts on inflows to region i, and its columns record the impacts on outflows from region i. That is, each element corresponds to the change of different flows concerning region i in detail.

Accordingly, we can calculate the cumulative total effects (CTE) over all T time periods using Equation (5):

$$CTE = \begin{pmatrix} TE_1 \\ TE_2 \\ \vdots \\ TE_n \end{pmatrix} = \begin{pmatrix} \partial Y_{1(t \sim t+T)}/\partial X^r_{1(t)} \\ \partial Y_{2(t \sim t+T)}/\partial X^r_{2(t)} \\ \vdots \\ \partial Y_{n(t \sim t+T)}/\partial X^r_{n(t)} \end{pmatrix} = \sum_{s=0}^{T} D_s \begin{pmatrix} Jo_1\beta^r_o + Jd_1\beta^r_d \\ Jo_2\beta^r_o + Jd_2\beta^r_d \\ \vdots \\ Jo_n\beta^r_o + Jd_n\beta^r_d \end{pmatrix}. \tag{5}$$

A special case is the contemporaneous effect ($T = 0$), which is caused only by spatial interactions without temporal dependence. As time passes, T tends to be infinite and the cumulative long-term effects tend to be stable. In this circumstance, D_T approaches 0 and $\sum_{s=0}^{T} D_s = (B + C)^{-1}$.

Furthermore, the scalar summary measure of total effects (te), expressed by the average of the total impacts on all flows in the network, can be calculated as: $te = 1/n^2 \cdot \iota_{n^2} \cdot CTE \cdot \iota_n$.

Analogously, a scalar summary measure of origin effects (oe) can be based on $oe = 1/n^2 \cdot \iota'_{n^2} \cdot COE \cdot \iota_n$, where COE is the $n^2 \times n$ matrix defined in Equation (6):

$$COE = \begin{pmatrix} OE_1 \\ OE_2 \\ \vdots \\ OE_n \end{pmatrix} = \sum_{s=0}^{T} D_s \begin{pmatrix} \tilde{J}o_1\beta^r_o \\ \tilde{J}o_2\beta^r_o \\ \vdots \\ \tilde{J}o_n\beta^r_o \end{pmatrix}, \tag{6}$$

where $\tilde{J}o_i$ is the $n \times n$ matrix that adjusts the element (i, i) in Jo_i to zero for the purpose of separating the intraregional effect. The adjusted matrix allows for only ($n - 1$) changes in outflows from region i as an origin in the OD dyad.

A scalar summary measure of destination effects (de) can be based on $de = 1/n^2 \cdot \iota'_{n^2} \cdot CDE \cdot \iota_n$, where CDE is the $n^2 \times n$ matrix defined in Equation (7):

$$CDE = \begin{pmatrix} DE_1 \\ DE_2 \\ \vdots \\ DE_n \end{pmatrix} = \sum_{s=0}^{T} D_s \begin{pmatrix} Jd_1\beta^r_d \\ Jd_2\beta^r_d \\ \vdots \\ \tilde{J}d_n\beta^r_d \end{pmatrix}, \tag{7}$$

where $\breve{J}d_i$ is the $n \times n$ matrix that adjusts the element (i, i) in Jd_i to zero without the intraregional effect. This matrix allows for only $(n-1)$ changes in inflows to region i as a destination.

A scalar summary measure for the intraregional effects (ie) can be based on $ie = 1/n^2 \cdot \iota'_{n^2} \cdot CIE \cdot \iota_n$, where CIE is the $n^2 \times n$ matrix defined in Equation (8):

$$CIE = \begin{pmatrix} IE_1 \\ IE_2 \\ \vdots \\ IE_n \end{pmatrix} = \sum_{s=0}^{T} D_s \begin{pmatrix} Ji_1(\beta_o^r + \beta_d^r) \\ Ji_2(\beta_o^r + \beta_d^r) \\ \vdots \\ Ji_n(\beta_o^r + \beta_d^r) \end{pmatrix}, \tag{8}$$

where Ji_i is an $n \times n$ matrix of zeros with a '1' in element (i, i).

The scalar summary measure for the cumulated spillover effects (se) can be calculated as $se \equiv te - oe - de - ie$. More generally, spatial effects can be expressed using $se = 1/n^2 \cdot \iota'_{n^2} \cdot CSE \cdot \iota_n$, where CSE is an $n^2 \times n$ matrix, by setting all the elements of the ith row and the ith column of each TE_i in TE to 0.

In the nonspatial dynamic panel model, (ρ_o, ρ_d, ρ_w) and $(\theta_o, \theta_d, \theta_w)$ are set to 0 and the spillover effects dissolve into nothingness.

4. Results

To illustrate our model of space-time dynamics within a migration network, we applied a panel data set of 5,580 gross migration flows between 31 mainland provinces in China (excluding intra-provincial migration flows) over six time periods between 1985 and 2015 (each time period equals five years). Migration flow data were collected from three population censuses (NBSC 1993, 2002, 2012) and three 1% population sample surveys (NBSC 1997, 2007, 2016). For consistency, migration flows to and from Chongqing during 1985–1990 and 1990–1995 were imputed based on its corresponding proportion of the population of Sichuan, since Chongqing was upgraded to a centrally administered municipality from a city of Sichuan province in 1997. Similarly, the missing migration flows to and from the Tibet Autonomous Region during 1985–1990 were also imputed from those of the adjacent time periods, since those data were not available in the 1990 census.

Besides the seven endogenous space and/or time lagged variables of migration flows previously discussed, to characterize regional socioeconomic status in determining how each variable influences interprovincial migration, we also chose regional population size, gross domestic product (GDP), age structure, real wages at the origin and destination, and a distance variable represented by time cost (Liu 2014; Shen

2016, 2017). Population size, as a typical gravity variable, has been demonstrated as effective in driving continued migration flows in China since the late 1970s (Fan 2005). The signs of population size at the origin and destination are expected to be positive, indicating push-pull forces between origins and destinations. GDP at the origin and destination is used to test the impact of regional economic disparity on stimulating migration flows. The expected signs of GDP at the origin and destination are negative and positive, respectively, suggesting people will move from a relatively poor region to a developed one. Age structure, measured by the proportion of working-age population (ages 15−64), reflects the predominance of labour migration. According to the 2015 1% population survey, labour population accounted for more than 85% of migrants during the 2010−2015 period (NBSC 2016). The signs of age structure at the origin and destination are expected to be positive and negative, respectively, indicating migrants are moving from labour-supplying regions to labour-demanding ones. Real wages as a measure of regional economic opportunity are the main motivation for mobility. We expect that the signs of real wages at the origin and destination are opposite, meaning that people will move to a region with higher income from a region with lower income. Compared with the constant railway distance, travel time can reflect improvement in transportation infrastructure. A negative decay effect of time distance is expected.

We use a log transformation of both dependent variable and explanatory variables to interpret the space-time effects as elasticities. To avoid the endogeneity problem, we lagged the dependent variable and explanatory variables by five years (LeSage and Pace 2008; Patuelli and Arbia 2016). Thus, the lagged migration flows at the initial period (1985−1990) were taken from the 1987 1% population sample survey (NBSC 1988), while data on regional population size, GDP, the proportion of labour force population, and real wages were collected from the China Statistical Yearbook on the first year of their corresponding periods (NBSC 1986, 1991, 1996, 2001, 2006, 2011). The minimum travel time by railway was collected from the China Railway Timetable at the same base year as the other exogenous explanatory variables (TBMR 1985, 1990, 1995, 2000, 2005, 2010). To check for the problem of potential multicollinearity, we calculate the variance inflation factor (VIF) for all variables (under 2.50), indicating no serious multicollinearity in our model. Table 1 describes these explanatory variables in detail.

We construct the spatial weight matrix W based on the first-order contiguity, reflecting whether a common boundary exists between two provinces, denoted by 0 or 1. Since Hainan province is an island that once belonged to Guangdong province, we designate Guangdong as its neighbour. The spatial weight matrix W is then row-normalized to produce three different network weight matrices W_o, W_d, and W_w. In practice, we remove the diagonal elements from W_o, W_d, and W_w to derive the $N' \times N'$ matrices $\widetilde{W}_o$, $\widetilde{W}_d$, and $\widetilde{W}_w$ ($N' = n \times (n-1)$) without considering intraregional flows. Different specifications of the spatial weight matrix have little impact on the subsequent

estimates and inferences, consistent with LeSage and Pace (2014), and hence we focus on the results from the first-order contiguity matrix.

Table 1:   Description of explanatory variables

Variable	Description (unit)
Popu	The size of total population in each province (ten thousands)
GDP	Gross domestic product at 1980 constant price in each province (billion yuan)
Age	The proportion of labour population aged 15–64
Wage	The ratio of average urban wage level to consumer price index (CPI) in each province (yuan)
Dis	Minimum railway travel time between each pair of provincial capitals (seconds)

Source: NBSC (1986, 1991, 1996, 2001, 2006, 2011); TBMR (1985, 1990, 1995, 2000, 2005, 2010).

4.1 Model parameter estimates

To further investigate the space-time dynamics in a migration system, we applied both nonspatial and spatial dynamic panel data models to our case study. The Bayesian MCMC sampling methods were implemented in the MATLAB based on the spatial econometrics toolbox (jplv7) (LeSage 2010). We estimated the nonspatial dynamic panel data model based on 500 draws, with the first 400 omitted for burn-in. Considering more parameters were involved in the spatial dynamic panel model, we then sampled 150,000 draws with the first 140,000 excluded for burn-in to ensure estimation accuracy. To check whether the MCMC samples converge, we initialized three chains at values that were over-dispersed relative to these parameter estimates. Table 2 reports the parameter estimates of the nonspatial and spatial dynamic panel data models.

Table 2: **Parameter estimates for nonspatial and spatial dynamic panel models**

Parameter	Nonspatial dynamic panel model				Spatial dynamic panel model			
	Total	500	Burn-in	400	Total	150000	Burn-in	140000
	Lower 0.05	*Mean*	*Upper 0.95*	*t-Statistic*	*Lower 0.05*	*Mean*	*Upper 0.95*	*t-Statistic*
α	−5.7341	−1.6694	1.9551	−0.7946	−6.7247	−3.3330	0.0827	−1.9194
$\gamma(Dis)$	−0.4963	−0.4552	−0.4132	−18.6179	−0.2020	−0.1575	−0.1126	−6.8940
$\beta_o(Popu)$	0.3892	0.4261	0.4639	21.8601	0.1659	0.2017	0.2375	11.0550
$\beta_o(GDP)$	−0.4281	−0.3509	−0.2610	−7.8175	−0.1860	−0.1235	−0.0614	−3.8753
$\beta_o(Age)$	0.7202	1.5053	2.2712	3.4021	0.3136	0.9088	1.5014	2.9905
$\beta_o(Wage)$	−0.2668	−0.1473	−0.0033	−2.1084	−0.1664	−0.0692	0.0293	−1.3864
$\beta_d(Popu)$	0.2778	0.3280	0.3652	15.9367	0.0958	0.1282	0.1612	7.6806
$\beta_d(GDP)$	0.1485	0.2300	0.3285	5.2076	0.0534	0.1182	0.1830	3.5679
$\beta_d(Age)$	−1.3312	−0.5256	0.3042	−1.1194	−0.6388	−0.0439	0.5549	−0.1443
$\beta_d(Wage)$	0.2009	0.3525	0.4737	4.9009	−0.0194	0.0770	0.1736	1.5615
ϕ	0.5318	0.5377	0.5442	141.1144	0.5787	0.6009	0.6220	54.0953
ρ_o	—	—	—	—	0.4251	0.4592	0.4828	31.3645
ρ_d	—	—	—	—	0.2954	0.3229	0.3501	22.5314
ρ_w	—	—	—	—	0.0468	0.0712	0.1132	3.7827
θ_o	—	—	—	—	−0.2691	−0.2374	−0.2007	−13.7004
θ_d	—	—	—	—	−0.1774	−0.1428	−0.1025	−7.4954
θ_w	—	—	—	—	−0.2575	−0.2121	−0.1727	−9.7941
Adjusted R^2			0.7435				0.7855	
BIC	−38697				−39743			

In general, the spatial dynamic panel data model is superior to the nonspatial one in characterizing the Chinese interprovincial migration process during 1985–2015, as seen by its smaller BIC (Bayesian Information Criterion) value. Furthermore, Chinese interprovincial migration is a path-dependent process, as demonstrated by the fact that temporal, spatial, and space-time diffusion dependencies are significantly different from zero. The significant and positive temporal dependence (ϕ) indicates that past migrants tend to produce future sustained outflows and inflows. In China, migration flows are usually dependent on social linkages, such as relying on family members, friends, or fellow-townspeople for job opportunities (Liu 2014). Even if an entire family moves, their migration process is often unsynchronized. It is the positive network externalities that make subsequent migrants follow the path blazed by the pioneers, giving rise to strong temporal dependence during migration processes (Epstein 2008). Surprisingly, temporal dependence has been strengthened by including space-time interactions in the spatial dynamic panel data model, rather than the contrary, reflecting network externalities' possible interaction with herd effects.

The significant and positive spatial dependence parameters (ρ_o, ρ_d, and ρ_w) can corroborate the existence of herd effects during the development of migration flows (Epstein 2008). The origin-, destination-, and origin-to-destination-based spatial dependencies represent three different but complementary migration paths. Specifically, the origin-based path implies that migration flows from the origin to the destination may promote outflows from surrounding origins to the same destination. Similarly, migration flows from the origin (or neighbouring origins) to neighbouring destinations are also enhanced, as implied by the destination-based (or origin-to-destination-based) path. Comparatively, the origin-based path dominates the other two; the largest spatial dependence ($\rho_o = 0.4592$) is further supported by the fact that origins with strong outflows are mainly located in China's central and western regions. Like the time decay effect, the strength of origin-, destination-, and origin-to-destination-based spatial dependences of less than 1 further corroborates the distance-decay effects over space.

Mutual space-time interactions have negative diffusion effects on migration flows, as reported by the significant and negative parameters (θ_o, θ_d, and θ_w). The origin-, destination-, and origin-to-destination-based space-time paths are exerting opposite roles compared with those three migration paths across space. In other words, if the number of migration flows from an origin to a destination at one point in time increases, the flows departing from the neighbouring origins to the same destination during the next time period will decrease. Similarly, the migration flows from one origin (or neighbouring origin regions) to neighbouring destinations in the next period will also decrease. Among these space-time paths the flows along the origin-based space-time path will decrease sharply due to the coexistence of herd and network effects, evidenced by the smallest space-time parameter ($\theta_o = -0.2374$). This

phenomenon fully reflects the self-adjustment and adaptability of a complex migration system, which involves much more complicated processes than the nonspatial model suggests.

Distance, as measured by minimum railway travel time, has significant and negative effects, as expected. Notably, the distance-decay effect in the spatial dynamic panel data model is significantly reduced compared with that in the nonspatial model, since spatial dependence can play a role similar to that of distance (LeSage and Thomas-Agnan 2015). Although the rapid progress of transportation technology has greatly facilitated migration activity, longer distances and higher time costs still restrict people's aspirations to migrate, as does the relative scarcity of information about distant regions (Huff and Jenks 2015).

The coefficient estimates of explanatory variables in the spatial dynamic panel data model should not be directly interpreted as partial derivative impacts of regional factors on migration flows. To better understand the space-time dynamics in determining interprovincial migration processes in China, we perform a comprehensive quantitative analysis by calculating various space-time effects of explanatory variables and drawing valid inferences about their statistical significance.

4.2 Analysis of space-time effects

Tables 3 and 4 show various effect estimates of regional population size, GDP, age structure, and real wages in the nonspatial and spatial dynamic panel data models, respectively. The rows labelled '0' represent the immediate responses of migration flows to changes in regional characteristics in terms of contemporaneous effects. The following rows labelled between '1' and '5' capture the disaggregated period-by-period impacts on the follow-on migration flows at different time periods. The last rows labelled '…' reflect the cumulative long-term effects. The columns labelled 'Mean' record the means of the transitory period-by-period effects and the columns labelled 'Cumulative' the accumulated effects from the initial period to the current period.

Influenced by the temporal dependence and distance-decaying effects, various period-by-period effects are gradually reduced to zero over time, while the corresponding cumulative effects tend to approach their long-term effects in the future. This phenomenon conforms to the evolution of a migration system in reality. The introduction of space-time interactions in the spatial dynamic panel data model brings about profound changes in all kinds of effects. Overall, the role of regional factors is significantly extended and deepened in the spatial model, reflecting that space-time interactions not only promote closer connections between different interprovincial

migration flows but also stimulate a more lasting and sustained influence on the whole system.

Table 3:   Effect estimates for nonspatial dynamic panel model

Horizon T	Popu Mean	Popu Cumulative	GDP Mean	GDP Cumulative	Age Mean	Age Cumulative	Wage Mean	Wage Cumulative
Origin effects								
0	0.4261***	0.4261***	−0.3509***	−0.3509***	1.5053***	1.5053***	−0.1473**	−0.1473**
1	0.2291***	0.6552***	−0.1887***	−0.5396***	0.8095***	2.3147***	−0.0792*	−0.2265*
2	0.1232***	0.7784***	−0.1015**	−0.6411***	0.4353***	2.7500***	−0.0426*	−0.2691*
3	0.0662***	0.8447***	−0.0546**	−0.6957***	0.2341***	2.9841***	−0.0229*	−0.2921*
4	0.0356***	0.8803***	−0.0293*	−0.7250***	0.1259***	3.1101***	−0.0123*	−0.3044*
5	0.0192*	0.8995***	−0.0158**	−0.7408***	0.0677***	3.1778***	−0.0066*	−0.3110*
...	–	0.9218***	–	−0.7591***	–	3.2566***	–	−0.3188**
Destination effects								
0	0.3280***	0.3280***	0.2300***	0.2300***	−0.5256	−0.5256	0.3525***	0.3525***
1	0.1764***	0.5044***	0.1237***	0.3537***	−0.2827	−0.8083	0.1895***	0.5420***
2	0.0948***	0.5992***	0.0665***	0.4203***	−0.1521	−0.9604	0.1019***	0.6440***
3	0.0510***	0.6502***	0.0358***	0.4560***	−0.0818	−1.0422	0.0548***	0.6988***
4	0.0274***	0.6776***	0.0192***	0.4753***	−0.0440	−1.0862	0.0295***	0.7283***
5	0.0147***	0.6924***	0.0103***	0.4856***	−0.0237	−1.1099	0.0159***	0.7441***
...	–	0.7095***	–	0.4977***	–	−1.1374	–	0.7626***
Total effects								
0	0.7541***	0.7541***	−0.1209***	−0.1209***	0.9797*	0.9797*	0.2052***	0.2052***
1	0.4055***	1.1596***	−0.0650***	−0.1859***	0.5267*	1.5064*	0.1103***	0.3155***
2	0.2180***	1.3776***	−0.0349***	−0.2208***	0.2832*	1.7896*	0.0593***	0.3748***
3	0.1172***	1.4949***	−0.0188***	−0.2396***	0.1523*	1.9420*	0.0319***	0.4067***
4	0.0630***	1.5579***	−0.0101***	−0.2497***	0.0819*	2.0239*	0.0171***	0.4239***
5	0.0339***	1.5918***	−0.0054**	−0.2552***	0.0440*	2.0679*	0.0092*	0.4331***
...	–	1.7450***	–	−0.2615***	–	2.1192**	–	0.4438***

Notes: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Interval values of various effect estimates as well as mean and cumulative values can be obtained from the authors upon request.

Table 4: Space-time effect estimates for spatial dynamic panel model

Horizon T	Popu Mean	Popu Cumulative	GDP Mean	GDP Cumulative	Age Mean	Age Cumulative	Wage Mean	Wage Cumulative
Origin effects								
0	0.4181***	0.4181***	-0.2257***	-0.2257***	1.8629***	1.8629***	-0.1118***	-0.1118
1	0.2061***	0.6242***	-0.1307***	-0.3563***	0.9847***	2.8476***	-0.0667	-0.1785
2	0.1230***	0.7472***	-0.0795***	-0.4358***	0.5924***	3.4400***	-0.0409	-0.2193
3	0.0769***	0.8241***	-0.0501***	-0.4858***	0.3717***	3.8116***	-0.0259	-0.2452
4	0.0495***	0.8736***	-0.0323***	-0.5181***	0.2395***	4.0512***	-0.0167	-0.2619
5	0.0324***	0.9060***	-0.0212***	-0.5393***	0.1570***	4.2081***	-0.0110	-0.2729
⋮	—	0.9724***	—	-0.5826***	—	4.5291***	—	-0.2954
Destination effects								
0	0.2145***	0.2145***	0.1436***	0.1436***	0.1015	0.1015	0.0955	0.0955
1	0.0825***	0.2970***	0.0763***	0.2199***	-0.0287	0.0728	0.0497	0.1452
2	0.0424***	0.3394***	0.0417***	0.2616***	-0.0230	0.0499	0.0270	0.1722
3	0.0225***	0.3619***	0.0230***	0.2846***	-0.0151	0.0347	0.0149	0.1871
4	0.0123***	0.3742***	0.0128***	0.2974***	-0.0092	0.0255	0.0083	0.1955
5	0.0067***	0.3809***	0.0072***	0.3046***	-0.0056	0.0199	0.0047	0.2002
⋮	—	0.3889***	—	0.3141***	—	0.0103	—	0.2066
Spillover effects								
0	1.7500***	1.7500***	0.0467	0.0467	4.5353*	4.5353*	0.1155	0.1155
1	-0.0608	1.6892***	0.0146	0.0613	-0.1724	4.3629*	0.0118	0.1273
2	-0.1218*	1.5674***	0.0164	0.0778	-0.3727*	3.9902*	0.0046	0.1319
3	-0.1002***	1.4672***	0.0135	0.0912	-0.3111**	3.6791*	0.0032	0.1352
4	-0.0680***	1.3992***	0.0105	0.1017	-0.2167**	3.4624*	0.0029	0.1380
5	-0.0450***	1.3542***	0.0079	0.1096	-0.1468**	3.3156*	0.0024	0.1404
⋮	—	1.2637***	—	0.1288	—	3.0094*	—	0.1476
Total effects								
0	2.3827***	2.3827***	-0.0354	-0.0354	6.4997**	6.4997**	0.0992	0.0992
1	0.2277	2.6104***	-0.0397	-0.0751	0.7837	7.2833**	-0.0052	0.0940
2	0.0436	2.6540***	-0.0214*	-0.0965	0.1967	7.4800**	-0.0092	0.0848
3	-0.0008	2.6532***	-0.0136***	-0.1101	0.0454	7.5254**	-0.0077	0.0771
4	-0.0062	2.6470***	-0.0089***	-0.1190	0.0136	7.5391**	-0.0055	0.0716
5	-0.0059***	2.6411***	-0.0061***	-0.1251	0.0046	7.5436**	-0.0039	0.0677
⋮	—	2.6250***	—	-0.1397	—	7.5488*	—	0.0588

Notes: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Interval values of various effect estimates as well as mean and cumulative values can be obtained from the authors upon request.

4.2.1 Population size

Population size is often chosen as the basic explanatory variable to characterize regional mass in traditional gravity models (Zipf 1946; Stillwell 2008). Migration flows between origins and destinations are positively proportional to their population sizes if other variables are held constant. Tables 3 and 4 show that, as hypothesized, the origin and destination effects of population size are significantly positive, exerting strong push and pull forces on interprovincial migration in China from 1985 to 2015. In addition, the cumulative origin and destination effects of population size continue to increase over time, indicating that changes in regional population have profound impacts on migration flows. The origin effects of population size are larger than the corresponding destination effects, reflecting the fact that the pressure from large populations is much stronger than their agglomeration effects during the migration process. In a typical region, a 1% increase in population size produces 0.42% more outflow during the same period. As time passes, transitory responses remain significantly positive, eventually leading to 0.97% additional outflow over the long term. On the other hand, a 1% increase in population size in a typical region produces 0.21% more inflow during the same period and 0.39% additional inflow in the long run, less than half of its long-term origin effect.

More importantly, changes in regional population size also impact the inflows and outflows to and from other regions through the migration network over time. It is these complex spillover effects that demonstrate the close link between migration flows. The spatial dynamic panel data model allows us to measure these invisible effects; the magnitude of such effects is beyond our expectations. Specifically, the cumulative spillover effects at different time periods are greater than 1.0, indicating a more rapid increase of migration flows arising from the same population growth across the network. On average, if a given population size in a typical region increases by 1%, the migration flows between other provinces will increase by 1.75% over the same period, well beyond the increase of outflows and inflows from and to the given region. This kind of spillover effect occurring at the same time (herd effect) may cause large flows even before networks are created, providing a complementary explanation for the formation of migration flows (Epstein 2008). As time passes, negative feedback effects in the following periods gradually weaken the cumulative network effects and eventually lead to a long-term effect of 1.2637, over two-thirds of its contemporaneous effect. These negative network externalities are also consistent with the reality that the number of immigrants may gradually reduce as the increase in new arrivals results in more competition for jobs. Declining network effects may help prevent the migration system from becoming out of control.

Total effects, represented by the sum of origin, destination, and spillover effects, reflect broader societal responses to changes in regional population size. The estimates

in Table 4 show that population size has the second largest impact on Chinese interprovincial migration flows. On average, a 1% increase in regional population in a typical region increases migration flows by 2.38% during the same period and by 2.63% over the long term. Although the spillover effects of population may decrease over time, they are still greater than its origin and destination effects. In other words, we cannot ignore the role of spillover effects in addition to traditional push and pull forces on the formation and development of migration flows.

4.2.2 GDP

Unbalanced regional economic development in China has led to continuing internal migration flows since the late 1970s (Cai and Wang 2003; Zhu 2003; Fan 2005). Tables 3 and 4 show that regional GDP plays a significant role in causing people to move between different areas, consistent with the neoclassical migration theory (Sjaastad 1962). The negative origin effects' inhibiting role is strong during the initial period. On average, a 1% increase in regional GDP produces 0.23% less outflow in the same period ($T = 0$). As time passes the period-by-period effects decrease but remain significant. Over the long run the origin effect of GDP reaches -0.5826, an absolute value much larger than that of GDP during the contemporaneous period. In other words, regional economic development tends to reduce migration outflows over time.

As expected, the development of a regional economy attracts more immigrants, as evidenced by the significant and positive destination effects of GDP. On average, a 1% increase in GDP in a typical region attracts 0.14% more inflows during the same period and 0.08% more inflows during the following period ($T = 1$). Although the period-by-period effects appear to decay over time, the cumulative effect still reaches 0.3141 over the long term, twice that of the contemporaneous effect. If we do not consider the space-time interactions among migration flows (Table 3), most of the destination effects of GDP tend to be overestimated, thus ignoring the self-adjustment of migration processes.

Regional GDP's spillover effects are not as significant as its origin and destination effects. In other words, the role of regional GDP in generating and attracting migration flows is mainly limited to the local inflows and outflows associated with the changing regions, exerting less externalities on the other surroundings. However, the positive period-by-period responses of GDP may accumulate to a larger long-term effect. Eventually, a 1% increase in regional GDP in a typical region produces 0.13% more flows to other provinces throughout the migration system, which is very close to its contemporaneous destination effect.

Overall, the total effects of regional GDP can help explain the inverted U-shape relationship between development and migration. After sustained and rapid growth of interprovincial migration, China has experienced a declining trend since 2010 (NBSC 1993, 1997, 2002, 2007, 2012, 2016). Furthermore, the period-by-period effects of regional GDP for the contemporaneous time period out to one time horizon ($T = 1$) are not significantly different from zero, indicating that development may not necessarily decrease migration flows. In reality, regional economic development, at least in the initial time periods, may enhance migration because development enables and inspires people to move (Zelinsky 1971, De Haas 2007, 2010b). However, the transitory period-by-period effects of regional GDP are significant from the two time horizons ($T = 2$) on, reflecting that development does have a lagged decaying effect on migration.

4.2.3 Age structure

The labour force population is the most influential driver of migration flows. According to population censuses and national survey data (NBSC 1993, 1997, 2002, 2007, 2012, 2016), migrant workers dominate in the total migration population. From Table 4 we see that the origin effects of the proportion of working age population are the largest of all the explanatory variables, consistent with the nonspatial model estimates. The impact of a 1% increase in working age population at origin would result in a 1.86% increase in outflows during the same period and eventually a 4.53% outflow increase, more than four times the origin effects of population size. The enhancement of origin effects in the spatial model is the result of feedback effects arising from changes in network flows associated with regions neighbouring each OD dyad. The small magnitude of destination effects suggests that an increase in the proportion of local labour population does not significantly attract more foreigners to move in, indicating that its agglomeration effects are rather limited.

The spillover effects arising from the increase in the proportion of the labour-force population are considerable, making a great difference to the evolution of a migration system. Specifically, a 1% increase in the proportion of labour force in a typical region leads to a 4.54% contemporaneous increase in other flows throughout the network, surpassing its cumulative origin effect in the long run. In practice this increase would be mainly concentrated in areas neighbouring the region where the proportion of labour force has increased. More strikingly, such period-by-period spillover effects tend to be more significant over time. It can be reasonably speculated that a change in the regional labour force market will have a lasting impact on other regions because of the significant space-time interactions. In the long run, the spillover effect for a 1% increase in the proportion of labour force in a typical region will produce 3.01% more

flows to other provinces throughout the migration network, making it a remarkable driving force.

Overall, interprovincial migration flows in China are extremely sensitive to the age structure of the population. Specifically, if the proportion of working age population in a typical region increases (or decreases) by 1%, there will be an increase (or decrease) of migration flows by 6.50% during the same period and by 7.55% over the long-term. That is, a slight change in the age structure of a regional population will have a great influence on current and future flows, especially in the contemporaneous period. This suggests that herd effects may dominate in migration processes, in line with the employment-oriented long-distance migration trends (Thomas, Gillespie, and Lomax 2019).

4.2.4 Real wages

A large number of migration studies indicate that people migrate to seek better job opportunities and to increase their quality of life (Chi and Voss 2005; Yue et al. 2010; Chen and Wang 2018). Without the consideration of space-time interactions, real wages do play a dominant role in the Chinese interprovincial migration process, as evidenced by various effect estimates shown in Table 3. However, this is not the case in the spatial model. We can see from Table 4 that the effect estimates of real wages are no longer significant, reflecting that real wages are not as important as expected in migration processes. These empirical results confirm that migrants are willing to give up higher wages initially, but prefer to move to more prosperous regions over time (Epstein 2008).

In practice, real wages express both the income and consumption of urban residents. On the one hand, developed regions with higher wage levels also have higher consumption standards and more survival pressure, reducing these regions' attractiveness to some potential migrants and affecting their migration decisions. On the other hand, remote and border regions with higher wage levels supported by national policies, such as the Tibet Autonomous Region, are not major destinations in China, going against the general expectation that regions with higher wages attract more migrants (Liu and Shen 2014). Overall, an increase in real wages reduces local residents' motivation to migrate and encourages more people to move in. A balance between wage income and living costs is essential for potential migrants' decision-making. If the quality of life in an area meets their expectations, most residents tend to stay where they live, reflecting the traditional Chinese attitude towards life: 'happiness lies in contentment' (*Zhi zu chang le*).

5. Conclusions

Migration is arguably playing an increasingly important role in regional growth and development than ever before. To understand the space-time dynamics of complex migration processes, this study constructs a spatial dynamic panel data model by incorporating temporal, spatial, and space-time diffusion dependencies into traditional gravity models. Moreover, we decompose origin, destination, and spillover effects into their contemporaneous, short-, and long-term components, thereby providing the tools to quantify the transitory and cumulative responses of migration flows to changes in regional characteristics. We then apply this model to interprovincial migration flows between 31 mainland provinces in China from 1985 to 2015, choosing population size, GDP, age structure, real wages at origin and destination, and railway travel time between provincial capitals as explanatory variables. To evaluate the importance of space-time interactions, we use a general (nonspatial) dynamic panel data model as a benchmark specification. Finally, we employ Bayesian MCMC procedures to estimate various effects and draw valid inferences regarding the mechanisms of Chinese interprovincial migration processes.

Our empirical results show that the spatial dynamic panel data model is more suitable for characterizing the path-dependent migration process. Past migrants tend to produce sustained outflows and inflows, as evidenced by more strong temporal dependence. Moreover, the origin-, destination-, and origin-to-destination-based paths complement one another during the generation and development of migration flows, as confirmed by the significant and positive spatial dependence. Comparatively, the origin-based path dominates the other two. However, the origin-, destination-, and origin-to-destination-based space-time paths provide different channels with which to adjust migration flows, as reflected by significant and negative space-time diffusion interactions. The origin-based space-time path is likely to be the most effective way to adjust the migration system toward an equilibrium state. Migration flows are perpetuated by the inertia of positive temporal and spatial dependence as well as by negative space-time diffusion dependence.

The disaggregation of origin, destination, and spillover effects over time further allows us to investigate the complex space-time dynamics underlying path-dependent migration processes. First, the development of migration flows in China is very sensitive to the age structure of regional populations, as a reflection of the workforce composition. A slight change in the proportion of the workforce will not only have a strong impact on outflows through transitory and cumulative origin effects, but will also influence other flows through diffusion effects, especially contemporaneous herd effects. Second, regional population size also plays a prominent role in migration processes. Specifically, the significant and positive spillover effects are much larger

than the sum of origin and destination effects, indicating that herd and network externality effects dominate in the migration system. Third, the inverted U-shape relationship between development and migration can be partly explained by the total effects of regional GDP. Finally, the role of real wages in driving migration flows is rather weak, suggesting that migrants are willing to give up higher wages initially but prefer prosperous regions in the long run.

The gravity model remains popular within the field of migration studies (O'Kelly 2004). Adding space-time components to this model will not only help promote theoretical, methodological, and empirical aspects of migration modelling, but also advance its application in other spatial interaction modelling areas such as transportation and international and interregional trade. More importantly, temporal, spatial, and space-time dependencies represent inertia in a system: Monitoring and observing their presence will help us understand the vulnerability and resilience of a regional system (Griffith and Chun 2015). In a rapidly changing society it is of critical importance to be able to predict what will happen to a regional migration system under shock, and how to respond. The space-time framework can be applied to forecast future migration flows as well as to explain past and present processes (Baltagi, Fingleton, and Pirotte 2019).

The spatial dynamic panel approach proposed in this study is a powerful means to understand how migration flows respond to regional socioeconomic changes and how space and time influence migration processes. However, the various coefficient and effect estimates are computationally expensive because more parameters are involved in the spatial model: Practical parallel strategies need to be tested. In addition, we only considered global spillover effects with a dependent variable lagged in both space and time: Possible local spillover effects can be investigated in a spatial dynamic Durbin panel model. Furthermore, we ignored model uncertainty in the selection of explanatory variables: The Bayesian model averaging method can be further explored with more regional social, economic, and environmental factors. Overall, to track the complex evolution of a regional migration system in future research, other types of spatial-temporal models for dyadic flow data should be specified, interpreted, and calibrated.

6. Acknowledgements

References

Anselin, L., Le Gallo, J., and Jayet, H. (2008). Spatial panel econometrics. In: Mátyás, L. and Sevestre, P. (eds.). *The econometrics of panel data: Fundamentals and recent developments in theory and practice*. Berlin: Springer: 625–660. doi:10.1007/978-3-540-75892-1_19.

Baláž, V. and Williams, A.M. (2007). Path-dependency and path-creation perspectives on migration trajectories: The economic experiences of Vietnamese migrants in Slovakia. *International Migration* 45(2): 37–66. doi:10.1111/j.1468-2435.2007.00403.x.

Baltagi, B.H., Fingleton, B., and Pirotte, A. (2019). A time-space dynamic panel data model with spatial moving average errors. *Regional Science and Urban Economics* 76: 13–31. doi:10.1016/j.regsciurbeco.2018.04.013.

Bauer, T., Epstein, G.S., and Gang, I.N. (2007). The influence of stocks and flows on migrants' location choices. *Research in Labor Economics* 26: 199–229. doi:10.1016/S0147-9121(06)26006-0.

Bavaud, F., Kordi, M., and Kaiser, C. (2018). Flow autocorrelation: A dyadic approach. *Annals of Regional Science* 61(1): 95–111. doi:10.1007/s00168-018-0860-y.

Beenstock, M. and Felsenstein, D. (2019). *The econometric analysis of non-stationary spatial panel data.* Switzerland: Springer. doi:10.1007/978-3-030-03614-0.

Bellos, S. and Subasat, T. (2012). Governance and foreign direct investment: A panel gravity model approach. *Bulletin of Economic Research* 64(4): 565–574. doi:10.1111/j.1467-8586.2010.00370.x.

Bernard, A. (2017). Cohort measures of internal migration: Understanding long-term trends. *Demography* 54(6): 2201–2221. doi:10.1007/s13524-017-0626-7.

Bhargava, A. and Sargan, J.D. (1983). Estimating dynamic random effects models from panel data covering short time periods. *Econometrica* 51(6): 1635–1659. doi:10.2307/1912110.

Biagi, B., Faggian, A., Rajbhandari, I., and Venhorst, V.A. (2018). *New frontiers in interregional migration research*. Cham: Springer International Publishing. doi:10.1007/978-3-319-75886-2.

Bijak, J. (2011). *Forecasting international migration in Europe: A Bayesian view*. Berlin: Springer. doi:10.1007/978-90-481-8897-0.

Black, W. (1992). Network autocorrelation in transportation network and flow systems. *Geographical Analysis* 24(3): 207−222. doi:10.1111/j.1538-4632.1992.tb00 262.x.

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

Chen, J. and Wang, W. (2018). Economic incentives and settlement intentions of rural migrants: Evidence from China. *Journal of Urban Affairs* 41(3): 372–389. doi:10.1080/07352166.2018.1439339.

Chi, G. and Voss, P.R. (2005). Migration decision-making: A hierarchical regression approach. *Journal of Regional Analysis and Policy* 35(2): 11–22.

Chib, S. and Carlin, B.P. (1999). On MCMC sampling in hierarchical longitudinal models. *Statistics and Computing* 9(1): 17−26. doi:10.1023/A:1008853808677.

Chort, I. and de la Ruppelle, M. (2016). Determinants of Mexico–U.S. outward and return migration flows: A state-level panel data analysis. *Demography* 53(5): 1453−1476. doi:10.1007/s13524-016-0503-9.

Chun, Y. (2008). Modeling network autocorrelation within migration flows by eigenvector spatial filtering. *Journal of Geographical Systems* 10(4): 317−344. doi:10.1007/s10109-008-0068-2.

Chun, Y. and Griffith, D. (2011). Modeling network autocorrelation in space–time migration flow data: An eigenvector spatial filtering approach. *Annals of the Association of American Geographers* 101(3): 523−536. doi:10.1080/00045 608.2011.561070.

Church, J. and King, I. (1993). Bilingualism and network externalities. *The Canadian Journal of Economics* 26(2): 337−345. doi:10.2307/135911.

Crown, D., Jaquet, T., and Faggian, A. (2018). Interregional migration and implications for regional resilience. In: Biagi, B., Faggian, A., Rajbhandari, I., and Venhorst, V.A. (eds.). *New frontiers in interregional migration*. Cham: Springer International Publishing: 231−252. doi:10.1007/978-3-319-75886-2_11.

Cushing, B. and Poot, J. (2003). Crossing boundaries and borders: Regional science advances in migration modelling. *Papers in Regional Science* 83(1): 317−338. doi:10.1007/s10110-003-0188-5.

Davidescu, A.A.M., Strat, V.A., Grosu, R.M., and Zgură, I.D. (2017). Determinants of Romanians' migration within the European Union: Static and dynamic panel gravity approaches. *Amfiteatru Economic* 19(46): 621−639.

De Haas, H. (2007). Turning the tide? Why development will not stop migration. *Development and Change* 38(5): 819−841. doi:10.1111/j.1467-7660.2007.004 35.x.

De Haas, H. (2010a). The internal dynamics of migration processes: A theoretical inquiry. *Journal of Ethnic and Migration Studies* 36(10): 1587−1617. doi:10.1080/1369183X.2010.489361.

De Haas, H. (2010b). Migration and development: A theoretical perspective. *International Migration Review* 44(1): 227−264. doi:10.1111/j.1747-7379.2009. 00804.x.

Debarsy, N., Ertur, C., and LeSage, J. P. (2012). Interpreting dynamic space–time panel data models. *Statistical Methodology* 9(1): 158−171. doi:10.1016/j.stamet.2011. 02.002.

Elhorst, J.P. (2010). Spatial panel data models. In: Fischer, M.M. and Getis, A. (eds.). *Handbook of applied spatial analysis: Software tools, methods and applications*. Berlin: Springer: 377−407. doi:10.1007/978-3-642-03647-7_19.

Elhorst, J.P. (2012). Dynamic spatial panels: Models, methods and inferences. *Journal of Geographical Systems* 14(1): 5−28. doi:10.1007/s10109-011-0158-4.

Elhorst, J.P. (2014). *Spatial econometrics: From cross-sectional data to spatial panels*. London: Springer. doi:10.1007/978-3-642-40340-8.

Epstein, G.S. (2008). Herd and network effects in migration decision-making. *Journal of Ethnic and Migration Studies* 34(4): 567−583. doi:10.1080/13691830801 961597.

Etzo, I. (2011). The determinants of the recent interregional migration flows in Italy: A panel data analysis. *Journal of Regional Science* 51(5): 948–966. doi:10.1111/j.1467-9787.2011.00730.x.

Fan, C.C. (2005). Modeling interprovincial migration in China, 1985−2000. *Eurasian Geography and Economics* 46(3): 165−184. doi:10.2747/1538-7216.46.3.165.

Fotheringham, A.S. (1991). Migration and spatial structure: The development of the competing destinations model. In: Stillwell, J. and Congdon, P. (eds.). *Migration models: Macro and micro approaches*. London: Belhaven Press: 57−72.

Golgher, A.B. and Voss, P.R. (2016). How to interpret the coefficients of spatial models: Spillovers, direct and indirect effects? *Spatial Demography* 4(3): 175−205. doi:10.1007/s40980-015-0016-y.

Greenwood, M.J. (1969). An analysis of the determinants of geographic labor mobility in the United States. *Review of Economics and Statistics* 51(2): 189–194. doi:10.2307/1926728.

Greenwood, M.J. (1970). Lagged response in the decision to migrate. *Journal of Regional Science* 10(3): 375–384. doi:10.1111/j.1467-9787.1970.tb00059.x.

Griffith, D.A. (2009). Spatial autocorrelation in spatial interaction: Complexity-to-simplicity in journey-to-work flows. In: Reggiani, A. and Nijkanmp, P. (eds.). *Complexity and spatial networks*. Berlin: Springer: 221–237. doi:10.1007/978-3-642-01554-0_16.

Griffith, D.A. and Chun, Y. (2015). Spatial autocorrelation in spatial interactions models: Geographic scale and resolution implication for network resilience and vulnerability. *Networks and Spatial Economics* 15(2): 337–365. doi:10.1007/s11067-014-9256-4.

Griffith, D.A. and Fischer, M.M. (2013). Constrained variants of the gravity model and spatial dependence: Model specification and estimation issues. *Journal of Geographical System*s 15(3): 291–317. doi:10.1007/s10109-013-0182-7.

Griffith, D.A., Fischer, M.M., and LeSage, J.P. (2017). The spatial autocorrelation problem in spatial interaction modeling: A comparison of two common solutions. *Letters in Spatial and Resource Sciences* 10(1): 75–86. doi:10.1007/s12076-016-0172-8.

Griffith, D.A. and Jones, K. (1980). Explorations into the relationships between spatial structure and spatial interaction. *Environment and Planning A* 12(2): 187–201. doi:10.1068/a120187.

Huff, D.L. and Jenks, G.F. (2015). A graphic interpretation of the friction of distance in gravity models. *Annals of the Association of American Geographers* 58(4): 814–824. doi:10.1111/j.1467-8306.1968.tb01670.x.

Kau, J.B. and Sirmans, C.F. (1979). A recursive model of the spatial allocation of migrants. *Journal of Regional Science* 19(1): 47–56. doi:10.1111/j.1467-9787.1979.tb00570.x.

Laber, G. (1972). Lagged response in the decision to migrate: A comment. *Journal of Regional Science* 12(2): 307–310. doi:10.1111/j.1467-9787.1972.tb00352.x.

Lamonica, G.R. and Zagaglia, B. (2013). The determinants of internal mobility in Italy, 1995–2006: A comparison of Italians and resident foreigners. *Demographic Research* 29(16): 407–440. doi:10.4054/DemRes.2013.29.16.

Lee, E.S. (1966). A theory of migration. *Demography* 3(1): 47−57. doi:10.2307/20 60063.

LeSage, J.P. (2010). Econometrics Toolbox. San Marcos: Texas State McCoy College of Business. http://spatial-econometrics.com.

LeSage, J.P. (2014). Spatial econometric panel data model specification: A Bayesian approach. *Spatial Statistics* 9: 122−145. doi:10.1016/j.spasta.2014.02.002.

LeSage, J.P. and Fischer, M.M. (2016). Spatial regression-based model specifications for exogenous and endogenous spatial interaction. In: Patuelli, R. and Arbia, G. (eds.). *Spatial econometric interaction modelling*. Cham: Springer International Publishing: 15−36. doi:10.1007/978-3-319-30196-9_2

LeSage, J.P. and Llano, C. (2013). A spatial interaction model with spatial structure origin and destination effects. *Journal of Regional Science* 15(3): 265−289. doi:10.1007/s10109-013-0181-8.

LeSage, J.P. and Pace, R.K. (2008). Spatial econometric modeling of origin-destination flows. *Journal of Regional Science* 48(5): 941−967. doi:10.1111/j.1467-9787. 2008.00573.x.

LeSage, J.P. and Pace, R.K. (2009). *Introduction to spatial econometrics*. Boca Raton: CRC press. doi:10.1201/9781420064254.

LeSage, J.P. and Pace, R.K. (2014). The biggest myth in spatial econometrics. *Econometrics* 2(4): 217−249. doi:10.3390/econometrics2040217.

LeSage, J.P. and Thomas-Agnan, C. (2015). Interpreting spatial econometric origin-destination flow models. *Journal of Regional Science* 55(2): 188−208. doi:10.1111/jors.12114.

Li, S. (2004). Population migration and urbanization in China: A comparative analysis of the 1990 population census and the 1995 national one percent sample population survey. *International Migration Review* 38(2): 655−685. doi:10.1111/ j.1747-7379.2004.tb00212.x.

Liang, Z. and Song, Q. (2016). Migration in China. In: White, M.J. (ed.). *International handbook of migration and population distribution*. Dordrecht: Springer Science + Business Media: 285−309. doi:10.1007/978-94-017-7282-2_14.

Liu, S. (2014). Factors analysis on interprovincial migration in China. *The Journal of Quantitative and Technical Economics* 31(4): 83−98 (in Chinese).

Liu, Y. and Shen, J. (2014). Jobs or amenities? Location choices of interprovincial skilled migrants in China, 2000−2005. *Population, Space and Place* 20(7): 592−605. doi:10.1002/psp.1803.

Mátyás, L. (1997). Proper econometric specification of the gravity model. *The World Economy* 20(3): 363−368. doi:10.1111/1467-9701.00074.

Mátyás, L. (2017). *The econometrics of multi-dimensional panels: Theory and applications*. Cham: Springer International Publishing. doi:10.1007/978-3-319-60783-2.

Mavroudi, E. and Nagel, C. (2016). *Global migration: Patterns, processes, and politics*. London: Routledge. doi:10.4324/9781315623399.

Mitze, T. (2012). *Empirical modelling in regional science: Towards a global time-space-structural analysis*. Berlin: Springer. doi:10.1007/978-3-642-22901-5.

Mitze, T. (2016). On the mutual dynamics of interregional gross migration flows in space and time. In: Patuelli, R. and Arbia, G. (eds.). *Spatial econometric interaction modelling*. Berlin: Springer: 415−440. doi:10.1007/978-3-319-30196-9_16.

National Bureau of Statistics of China (NBSC). (1986). *China statistics yearbook*. Beijing: China Statistics Press.

National Bureau of Statistics of China (NBSC). (1988). *Tabulations on the 1987 1% population sampling survey of the People's Republic of China*. Beijing: China Statistics Press.

National Bureau of Statistics of China (NBSC). (1991). *China statistics yearbook*. Beijing: China Statistics Press.

National Bureau of Statistics of China (NBSC). (1993). *Tabulation on the 1990 population census of the People's Republic of China*. Beijing: China Statistics Press.

National Bureau of Statistics of China (NBSC). (1996). *China statistics yearbook*. Beijing: China Statistics Press.

National Bureau of Statistics of China (NBSC). (1997). *Tabulation on the 1995 1% population sampling survey of the People's Republic of China*. Beijing: China Statistics Press.

National Bureau of Statistics of China (NBSC). (2001). *China statistics yearbook*. Beijing: China Statistics Press.

National Bureau of Statistics of China (NBSC). (2002). *Tabulation on the 2000 population census of the People's Republic of China*. Beijing: China Statistics Press.

National Bureau of Statistics of China (NBSC). (2006). *China statistics yearbook*. Beijing: China Statistics Press.

National Bureau of Statistics of China (NBSC). (2007). *Tabulation on the 2005 1% population sampling survey of the People's Republic of China*. Beijing: China Statistics Press.

National Bureau of Statistics of China (NBSC). (2011). *China Statistics Yearbook*. Beijing: China Statistics Press.

National Bureau of Statistics of China (NBSC). (2012). *Tabulation on the 2010 population census of the People's Republic of China*. Beijing: China Statistics Press.

National Bureau of Statistics of China (NBSC). (2016). *Tabulation on the 2015 1% population sampling survey of the People's Republic of China*. Beijing: China Statistics Press.

Odland, J. and Shumway, J.M. (1993). Interdependencies in the timing of migration and mobility events. *Papers in Regional Science* 72(3): 221−237. doi:10.1007/BF01434274.

O'Kelly, M.E. (2004). Isard's contributions to spatial interaction modelling. *Journal of Geographical Systems* 6(1): 43−54. doi:10.1007/s10109-004-0125-4.

Parent, O. and LeSage, J.P. (2010). A spatial dynamic panel model with random effects applied to commuting times. *Transportation Research Part B* 44(5): 633−645. doi:10.1016/j.trb.2010.01.004.

Parent, O. and LeSage, J.P. (2012) Spatial dynamic panel data models with random effects. *Regional Science and Urban Economics* 42(4): 727−738. doi:10.1016/j.regsciurbeco.2012.04.008.

Patuelli, R. and Arbia, G. (2016). *Spatial econometric interaction modelling*. Cham: Springer International Publishing. doi:10.1007/978-3-319-30196-9.

Petersen, W. (1958). A general typology of migration. *American Sociological Association* 23(3): 256−266. doi:10.2307/2089239.

Plane, D.A. and Rogerson, P.A. (1986). Dynamic flow modeling with interregional dependency effects: An application to structural change in the U.S. migration system. *Demography* 23(1): 91−104. doi:10.2307/2061411.

Rainer, H. and Siedler, T. (2009). The role of social networks in determining migration and labour market outcomes: Evidence from German Reunification. *Economics of Transition* 17(4): 739−767. doi:10.1111/j.1468-0351.2009.00365.x.

Rogers, A. (1990). Requiem for the net migrant. *Geographical Analysis* 22(4): 283−300. doi:10.1111/j.1538-4632.1990.tb00212.x.

Rogers, A., Willekens, F., Little, J., and Raymer, J. (2002). Describing migration spatial structure. *Papers in Regional Science* 81(1): 29–48. doi:10.1007/s101100 100090.

Sjaastad, L.A. (1962). The costs and returns of human migration. *Journal of Political Economy* 70(5): 80−93. doi:10.1086/258726.

Sen, A. and Smith, T.E. (1995). *Gravity models of spatial interaction behavior*. Berlin: Springer. doi:10.1007/978-3-642-79880-1.

Shen, J. (2015). Explaining interregional migration changes in china, 1985–2000, using a decomposition approach. *Regional Studies* 49(7): 1176−1192. doi:10.1080/ 00343404.2013.812783.

Shen, J. (2016). Error analysis of regional migration modeling. *Annals of the American Association of Geographers* 106(6): 1−15. doi:10.1080/24694452.2016.11977 67.

Shen, J. (2017). Modelling interregional migration in China in 2005–2010: The roles of regional attributes and spatial interaction effects in modelling error. *Population, Space and Place* 23(3): e2014. doi:10.1002/psp.2014.

Stillwell, J. (2008). Inter-regional migration modelling: A review. In: Poot, J., Waldorf, B., and van Wissen, L. (eds.). *Migration and human capital*. Cheltenham: Edward Elgar: 29−48.

Thomas, M., Gillespie, B., and Lomax, N. (2019). Variations in migration motives over distance. *Demographic Research* 40(38): 1097−1110. doi:10.4054/DemRes. 2019.40.38.

Transport Bureau of the Ministry of Railway (TBMR). (1985). *China railway timetable 1985*. Beijing: China Railway Publishing House.

Transport Bureau of the Ministry of Railway (TBMR). (1990). *China railway timetable 1990*. Beijing: China Railway Publishing House.

Transport Bureau of the Ministry of Railway (TBMR). (1995). *China railway timetable 1995*. Beijing: China Railway Publishing House.

Transport Bureau of the Ministry of Railway (TBMR). (2000). *China railway timetable 2000*. Beijing: China Railway Publishing House.

Transport Bureau of the Ministry of Railway (TBMR). (2005). *China railway timetable 2005*. Beijing: China Railway Publishing House.

Transport Bureau of the Ministry of Railway (TBMR). (2010). *China railway timetable 2010*. Beijing: China Railway Publishing House.

United Nations (2012). World urbanization prospects. The 2011 revision. *Highlight* (ESA/P/WP/224). New York: Department of Economic and Social Affairs, Population Division. doi:10.18356/08b807d4-en.

White, M.J. (2016). *International handbook of migration and population distribution*. Berlin: Springer.

Williamson, J. (1988). Migrant selectivity, urbanization, and industrial revolutions. *Population and Development Review* 14(2): 287−314. doi:10.2307/1973573.

Yue, Z., Li, S., Feldman, M.W., and Du, H. (2010). Floating choices: A generational perspective on intentions of rural–urban migrants in China. *Environment and Planning A* 42(3): 545−562. doi:10.1068/a42161.

Zelinsky, W. (1971). The hypothesis of the mobility transition. *Geographical Review* 61(2): 219−249. doi:10.2307/213996.

Zhu, Y. (2003). The floating population's household strategy and the migration's role in China's regional development and integration. *International Journal of Population Geography* 9(6): 485−502. doi:10.1002/ijpg.308.

Zimmer, M. (2008). *Network effects on domestic migration flows across Germany: A spatial autoregressive perspective with spatially structured origin and destination effects and heteroskedastic innovations*. Paper presented at the seventh workshop on spatial econometrics and statistics, Paris, France, June 9−10, 2008.

Zipf, G.K. (1946). The P_1P_2/D hypothesis: On the intercity movement of persons. *American Sociological Review* 11(6): 677−686. doi:10.2307/2087063.

Appendix

The intercept term α and the parameter $\beta = (\beta_o{}', \beta_d{}', \gamma)$, as well as the precision parameters σ_ε^{-2} and σ_μ^{-2}, are under the assumption of a nonhierarchical prior of the independent Normal-Gamma variety. This implies that the posterior distributions of α and β can be reducible to Normal distribution while the posterior distributions of σ_ε^{-2} and σ_μ^{-2} can be reducible to Gamma distribution. Thus,

$$\alpha \sim N(\alpha_0, M_\alpha^{-1}), \beta \sim N(\beta_0, M_\beta^{-1}),$$
$$\sigma_\varepsilon^{-2} \sim G\left(\frac{v_0}{2}, \frac{S_0}{2}\right), \sigma_\mu^{-2} \sim G\left(\frac{v_1}{2}, \frac{S_1}{2}\right). \tag{9}$$

There is a great deal of prior uncertainty regarding α and β, and hence we adopt uninformative priors for them to resolve the interference of prior information and increase the reliability of the estimation results (LeSage and Pace 2009). Specifically, we set $\alpha_0 = 0$, $\beta_0 = 0$, $M_\alpha^{-1} = 10^{12}$, and $M_\beta^{-1} = 10^{12} I_{2k+1}$; similarly, v_0, S_0, v_1, and S_1 are all set to 0.001.

In this case where the conditional distributions are known, Gibbs sampling is used to directly draw a large number of samples of α, β, σ_ε^{-2}, and σ_μ^{-2} from their posterior distributions under the Bayesian updating schemes.

To determine the space-time parameters $\rho = (\rho_o, \rho_d, \rho_w)$, ϕ, and $\theta = (\theta_o, \theta_d, \theta_w)$, whose posterior distributions are not reducible to standard distributions, this study adopts a more generally applicable step labelled the Metropolis-Hastings (M-H) sampler, which relies on a random-walk proposal with normally distributed increments as the proposal density (Parent and LeSage 2010, 2012; Debarsy, Ertur, and LeSage 2012). Uniform priors for ρ, ϕ, θ are adopted over the interval $(-1, 1)$. During the sampling and updating process the acceptance probabilities of these parameters evaluated at the old and new candidate draws are required to move between 0.40 and 0.60, resulting in an effective acceptance rate close to 0.50 after a burn-in period.

By using Bayesian MCMC procedures, the generated sequential samples allow us to proceed with inferences regarding all the parameters.