Regional Studies, 2015
Vol. 49, No. 7, 1176–1192, http://dx.doi.org/10.1080/00343404.2013.812783

Explaining Interregional Migration Changes in China, 1985–2000, Using a Decomposition Approach

JIANFA SHEN
Department of Geography and Resource Management, and Research Centre for Urban and Regional Development, The Chinese University of Hong Kong, Shatin, NT, Hong Kong, China. Email: jianfa@cuhk.edu.hk

(Received June 2012: in revised form May 2013)

SHEN J. Explaining interregional migration changes in China, 1985–2000, using a decomposition approach, *Regional Studies*. This paper estimates the effects of changing parameters in migration models and the changing demographic, social, and economic attributes in origin and destinations on migration flows. A decomposition approach is developed based on multilevel Poisson migration models. Multilevel Poisson migration models are estimated for migration in China for the periods 1985–1990 and 1995–2000. Overall, the total inter-provincial migration was increased by 22 million in China from 1985–1990 to 1995–2000. The decomposition result shows that 62.28% of this increase was due to changes in the value of explanatory variables, while 37.72% was due to changes in the value of model parameters.

Migration change   Regional development   Migration decomposition   China

SHEN J. 运用分解的方法，解释中国 1985 年至 2000 年跨区人口迁移的转变，*区域研究*。本文评估迁移模型参数变化以及迁出地、迁入地人口、社会与经济指标的变化对迁移流的影响。本研究建立了一个基于多层级泊松迁移模型的分解方法，并分别对中国 1985 年—1990 年和 1995 年—2000 年间的人口迁移，建立多层级泊松迁移模型。总的来说，从 1985 年—1990 年至 1995 年—2000 年期间，中国的省际迁移增加了二千二百万人。分解结果表明，62.28% 的增加是因为解释变量数值的变化，37.72% 的增加则是因为模型参数值的变化。

迁移变化   区域发展   迁移分解   中国

SHEN J. Expliquer les modifications des flux migratoires interrégionaux en Chine, de 1985 à l'an 2000, employant une méthode de décomposition fonctionnelle, *Regional Studies*. Ce présent article estime les effets de la modification des paramètres dans les modèles de migration et l'évolution des caractéristiques démographiques, sociales et économiques des pays d'origine et des pays d'accueil sur les flux migratoires. On développe une méthode de décomposition basée sur les modèles de migration multiniveaux de Poisson. On estime des modèles de migration multiniveaux de Poisson pour la migration en Chine de 1985 à 1990 et de 1995 à l'an 2000. Tout considéré, la migration interrégionale globale a augmenté de 22 millions en Chine pour la période de 1985 à 1990 jusqu'à la période allant de 1995 à l'an 2000. Le résultat de la décomposition montre que 62,28% de cette augmentation s'expliquait par le changement des valeurs des variables explicatives, alors que 37,72% s'expliquait par les changements de la valeur des paramètres du modèle.

Flux migratoires   Aménagement du territoire   Décomposition de la migration   Chine

SHEN J. Erklärung der Veränderungen bei der Migration zwischen Regionen in China im Zeitraum von 1985 bis 2000 mit Hilfe eines Dekompositionsansatzes, *Regional Studies*. In diesem Beitrag werden die Auswirkungen der veränderten Parameter von Migrationsmodellen und der veränderten demografischen, sozialen und wirtschaftlichen Merkmale in den Herkunfts- und Zielgebieten auf die Migrationsströme geschätzt. Ausgehend von mehrschichtigen Poisson-Migrationsmodellen wird ein Dekompositionsansatz entwickelt. Die mehrschichtigen Poisson-Migrationsmodelle werden für die Migration in China im Zeitraum von 1985 bis 1990 sowie von 1995 bis 2000 geschätzt. Insgesamt stieg die gesamte Migration zwischen chinesischen Provinzen vom Zeitraum 1985–1990 bis zum Zeitraum 1995–2000 um 22 Millionen. Aus den Ergebnissen der Dekomposition geht hervor, dass 62,28% dieses Anstiegs durch Veränderungen bei den Werten der erklärenden Variablen verursacht wurde, während 37,72% auf Veränderungen bei den Werten der Modellparameter zurückzuführen sind.

Migrationsänderungen   Regionalentwicklung   Dekomposition von Migration   China

SHEN J. Explicación de los cambios de la migración interregional en China, entre 1985 y 2000, mediante un enfoque de descomposición, *Regional Studies*. En este artículo se calculan los efectos de los parámetros cambiantes en los modelos de migración y los atributos demográficos, sociales y económicos que varían en origen y destino en lo que afecta a los flujos de migración. Se desarrolla un enfoque de descomposición basado en los modelos de migración de Poisson de varios niveles. Se calculan estos modelos multiniveles de migración de Poisson para la migración en China durante los periodos entre 1985 y 1990 y entre 1995 y 2000. En general, la migración total interprovincial aumentó en 22 millones en China de 1985–1990 a 1995–2000. El resultado de descomposición muestra que un 62,28% de este aumento se debió a cambios en el valor de las variables explicativas, mientras que un 37,72% fue inducido por cambios en el valor de los parámetros de los modelos.

Cambio migratorio   Desarrollo regional   Descomposición de la migración   China

JEL classifications: R23

INTRODUCTION

Migration has become an increasingly important component of populations (VAN IMHOFF *et al.*, 1997; REES *et al.*, 2004; BELL *et al.*, 2002; HENRY *et al.*, 2003; ANDRIENKO and GURIEV, 2004). Micro- and macro-approaches are different approaches to migration analysis (STILLWELL and CONGDON, 1991). At the micro-level, the discrete-choice approach based on utility maximization is a common approach (STILLWELL, 2005). At the macro-level, aggregate data on migration are often analysed and used to model the macro-relationships between migration and related factors. A brief review of macro-studies is provided here, which is closely related to this paper. HUA and PORELL (1979) offered a detailed review of the development of the gravity model and STILLWELL (2005) published an excellent review of the recent development in interregional migration modelling. Various migration measures are often used for analysis and projection of migration. ROGERS (1992) argued that many migration rates are not true occurrence–exposure rates and they are ambiguous. Net migration rates, lifetime migration rates and return migration proportions are imperfect measures of migration propensities and they should be avoided in migration analysis.

Many efforts have been made to model and analyse migration with two major focuses. The first is on interregional migration flows. Previous studies have attempted to decompose migration into constituent components to identify their impacts. KRIESBERG and VINING (1978) probably made the first attempt to examine the components of changes of net migration, examining whether changes in net migration are caused primarily by changes in gross in-migration or out-migration. Using the Japanese inter-prefectural migration data in the 1970s, they concluded that changes in in-migration were more significant for urban prefectures, whereas changes in out-migration made the greater contribution for rural prefectures.

WILLEKENS (1983) demonstrated the relationship between the gravity model and the log–linear model and the meanings of balancing factors and first-order interaction effects in the log–linear model. MUESER (1989) made significant progress to decompose migration into sending effects, drawing effects and separation effect using a general spatial interaction model. Various forms of distance functions were used to improve the estimation of the separation effect. Generalized distance was also estimated. More recently, ROGERS *et al.* (2002) suggested using the log–linear specification of the spatial interaction model to describe the migration structure. The parameters of the log–linear model can fully describe the relative emissiveness and attractiveness of specific regions, as well as the level of interaction between pairs of regions. The method can be used to analyse migration and to capture the effect of historical migration patterns on contemporary migration (ROGERS, 1990).

PLANE (1987) made the first attempt to estimate a set of geographic components of temporal change in a system of interregional flows using a spatial adaptation of the shift–share technique. Dynamic change in net migration is conceptualized to consist of system growth, system mobility, geographic mix and competitive components.

More recently, SHEN (1999) considered the issue of identifying the effects of spatial structure and the origin and destination attributes on migration. A decomposition approach was developed making use of estimated migration models. A set of migration data in China for 1985–1990 was used to estimate various migration models to decompose the effects of spatial structure and the origin and destination attributes on migration in China.

MUESER (1989) and ROGERS *et al.* (2002) described the migration structure by the measures of origin, destination and interaction. They and PLANE (1987) focused on migration matrix or its changes, although MUESER (1989) used different distance functions to explain spatial interaction. Therefore, the characteristics of a region or spatial relations between regions are not used to explain migration in these studies, except for SHEN (1999). The origin and destination factors are considered in migration modelling.

The second focus of research is to model migration flows directly, explicitly using distance, populations, and other factors at origin and destination to explain migration. The estimated parameters are often used to describe how changes in an independent variable will affect the size of migration. Spatial interaction models have been formulated mathematically and statistically (WILSON, 1970; STILLWELL, 2005). The classical gravity model only considers three variables (HUA and PORELL, 1979). Extended demo-economic models often consider other socio-economic and environmental factors (ISSERMAN, 1985; STILLWELL and CONGDON, 1991; HENRY *et al.*, 2003). It is well known that different model specification and different estimation methods may produce different parameter estimates and goodness of fit (CONGDON, 1991). Thus, many studies on migration have focused on the issue of model specification as well as on distance measurement (BOYLE and FLOWERDEW, 1997). A lengthy debate in the 1970s (CLIFF *et al.*, 1974, 1976; CURRY *et al.*, 1975; JOHNSTON, 1973, 1975; SHEPPARD, 1979) focused primarily upon the problems of autocorrelation in the mass terms and the map pattern effect in the distance-decay parameters. It has been shown that origin-specific distance-decay terms are likely to be related to the map pattern (FOTHERINGHAM, 1981; JOHNSTON, 1973, 1975). FOTHERINGHAM (1984, 1991) argued that destinations may be related by forces of agglomeration or competition. Considering the local variations in parameters, geographically weighted regression (GWR) has also been used (FOTHERINGHAM *et al.*, 2002).

In a case study of US interstate migration, FIK and MULLIGAN (1998) found empirical evidence to support the argument that the use of highly restrictive log–linear specifications may be inappropriate and problematic. A Poisson model is considered as a more realistic description of migration process than a log–linear model (FLOWERDEW and AITKIN, 1982; BOYLE and HALFACREE, 1995). Multilevel models have also become popular to model random variation at different group or regional levels (JONES, 1991) and been used to explore the relationship between migration and the factors at individual and regional levels (BOYLE and SHEN, 1997; BOYLE *et al.*, 1998). A two-stage approach to migration modelling has been used by REES *et al.* (2004). Out-migration from each area is separately modelled at stage 1 and the distribution of migrants between destinations is then modelled at stage 2.

Most migration models are used to analyse migration for a particular period with a few exceptions. Parameters can also be compared for models estimated for different times with particular attention being paid to the changing distance-decay effect on migration (LEDENT, 1986; LI, 2004; FAN, 2005a). LEDENT (1986) studied the temporal changes in the parameters when migration models were separately estimated for each of a time-series of Canadian migration matrices. Some attempts have considered temporal variability in the origin–destination migration intensities (STILLWELL, 2005). PLANE (1982) used a destination-population-weighted (DPW) model with a balancing term to ensure that the probabilities sum to 1. COURGEAU (1995) suggested that an origin–destination migration intensity based on the populations of both the origin and destination can be defined. This paper adopts a different approach. It attempts to estimate the contribution of changes in determinants and model parameters to the change in migration based on migration models for two periods. While previous studies have attempted to decompose migration into different components, this paper takes a further step to decompose the change of migration into different components making use of the multilevel Poisson migration model (PLANE, 1987; MUESER, 1989; ROGERS *et al.*, 2002; SHEN, 1999).

Previous studies on inter-provincial migration in China have focused on the spatial patterns of migration (ZHU and CHEN, 2010; LIANG and MA, 2004; ZHU, 2003; YANG, 2000; CHAN *et al.*, 1999; HE and GOBER, 2003; SHEN, 2013), the regional concentration of migration flows (HE and POOLER, 2002; DING *et al.*, 2005), the determinants of migration (WEI, 1997; LI and LI, 1995; LIANG and WHITE, 1997; LIANG *et al.*, 2002; CAI and WANG, 2003; LI, 2004; FAN, 2005a, 2011; LIU and SHEN, 2013a, 2013b), and the relationship between migration and regional development (FAN, 2005b). CROMLEY *et al.* (2010) estimated net provincial migration using census data from 1990 and 2000. Various migration models were estimated to identify the effects of spatial structure on inter-provincial migration using 1990 census data (SHEN, 1999). PONCET (2006) studied how migratory forces evolved between 1985 and 1995 among rural–urban migrants in China. She found that aggregate migration costs declined between two periods: 1985–1990 and 1990–1995, along with the relaxation of migration restrictions.

Assuming that a migration process can be described by a migration model including both origin and destination attributes and a distance variable, the change in migration in a regional system can be caused in two ways: changes in the value of attributes reflecting the demographic, social, economic, and environmental changes in origin and destinations, and changes in the value of model parameters indicating the changing relationship between the independent variable and migration process. In this paper, inter-provincial migration models will be estimated for 1985–1990, 1995–2000 and 2005–2010, respectively, using data from three censuses conducted in China. Total in-flows and out-flows will be decomposed into several components reflecting the effects of various factors. The decomposition focuses on the period between the 1990 and 2000 censuses as the same migration model for 2005–2010 can only explain a small proportion of variation in migration flows.

CHANGING INTER-PROVINCIAL MIGRATION IN CHINA, 1985–2010

The migration datasets used here refer to migration among twenty-nine provincial regions in mainland China except for Tibet in the census periods 1985–1990, 1995–2000 and 2005–2010 (POPULATION OFFICE OF THE STATE COUNCIL and DEPARTMENT OF POPULATION STATISTICS (DPS), 1993; POPULATION CENSUS OFFICE OF THE STATE COUNCIL and DEPARTMENT OF POPULATION, SOCIAL, SCIENCE AND TECHNOLOGY STATISTICS OF THE NATIONAL BUREAU OF STATISTICS (DPSSTS), 2002; POPULATION CENSUS OFFICE OF THE STATE COUNCIL and DEPARTMENT OF

Population and Employment Statistics of the National Bureau of Statistics (DPES), 2012).

Since the early 1980s, the volume of internal migration in China has been increasing due to more relaxed migration policies and unbalanced regional development (Fan, 2005b; Shen, 2013; Jiang and Shen, 2010). The direction of internal migration has also undergone a transition from east to west migration in the pre-reform period to west to east migration in the reform period (Shen, 1996). Inter-provincial migration in China reached 11.01 million people in 1985–1990. The scale of inter-provincial migration tripled, reaching 33.89 million in the period 1995–2000. It doubled again, reaching 67.10 million in the period 2005–2010. While many studies have been done on migration in 1985–1990 and 1995–2000, respectively (Fan, 2005a, 2005b; He and Pooler, 2002; Li, 2004; Ding *et al.*, 2005), there is a need to explore systematically how and why inter-provincial migration has changed in China between census periods.

Some important observations can be made about net migration and migration efficiency of the twenty-nine regions in three periods 1985–1990, 1995–2000 and 2005–2010 (Table 1). First, Guangdong, Shanghai and Beijing, the three most advanced provincial units, remained the top three regions in the first two periods and were in the top four in the final period in the scale of net migration. Sichuan remained the province with the largest loss of population through migration.

Second, there were also important changes in the ranking of net migration. Zhejiang was ranked bottom of twenty-seven with large net out-migration in 1985–1990. But it changed to number four in net in-migration in 1995–2000 and number three in net in-migration in 2005–2010. On the other hand, Hubei's case was the opposite. It was ranked number eight in net migration in 1985–2000, but it changed into a bottom province ranked number twenty-three in 1995–2000 and number twenty-five in 2005–2010, losing much of its population through migration.

Third, migration efficiency is the ratio of net migration to the total of in-migration and out-migration. Migration efficiency was increased significantly in most provincial regions from 1985–1990 to 1995–2000, indicating that migrants were moving into specific regions without much return migration flow. However, between the period 1995–2000 and 2005–2010, migration efficiency deceased in fourteen provinces, indicating a more diversified pattern of migration.

Table 1. Net migration of selected provinces for 1985–1990, 1995–2000 and 2005–2010 (millions)

Region	1985–1990	1995–2000	2005–2010	Rank
Guangdong	1.01	11.92	14.08	1
Zhejiang	–0.30	1.79	7.38	2
Shanghai	0.53	2.12	4.76	3
Beijing	0.54	1.80	3.77	4
Jiangsu	0.17	0.71	3.28	5
Tianjin	0.17	0.40	1.56	6
Fujian	0.01	0.80	1.53	7
Heilongjiang	–0.24	–0.65	–1.24	20
Hebei	–0.13	–0.13	–1.25	21
Guizhou	–0.12	–1.05	–2.28	22
Guangxi	–0.45	–1.67	–2.52	23
Jiangxi	–0.07	–2.59	–3.06	24
Hubei	0.08	–1.66	–3.27	25
Hunan	–0.26	–3.10	–4.38	26
Anhui	–0.20	–2.73	–5.01	27
Henan	–0.11	–1.96	–5.54	28
Sichuan	–0.87	–4.71	–5.65	29

Note: Only provinces with net migration over 1 million in 2005–2010 are listed.

METHOD FOR DECOMPOSING CHANGES IN MIGRATION FLOWS

It is common to use explanatory variables in migration models to describe the volume of migration, M_{ij}, between places i and j for a particular period. This section develops an approach to decompose changes in migration flow into various components caused by changes in the value of explanatory variables or the value of model parameters. This section only needs to focus on the expected migration flow; the specification of the random component will be discussed in the following section.

Assuming that a migration model has been properly specified and estimated, the expected migration flow, M_{ij}^E, between region i and j can be expressed by a migration model. Consider a general migration model that can be expressed as follows using a distance variable, d_{ij} (or x_{2N+1}); and N indicators at origin and destination x_{ik} and x_{jk} ($k = 1, 2, \ldots, N$) (or x_k, $k = 1, 2, \ldots, 2N$), respectively:

$$M_{ij}^E = e^{a_0} \prod_{k=1}^{N} x_{ik}^{a_k} \prod_{k=1}^{N} x_{jk}^{a_{N+k}} d_{ij}^{a_{2N+1}}$$

$$= e^{a_0} \prod_{k=1}^{N} x_k^{a_k} \prod_{k=N+1}^{2N} x_k^{a_k} x_{2N+1}^{a_{2N+1}} = e^{a_0} \prod_{k=1}^{2N+1} x_k^{a_k} \quad (1)$$

where a_0 is a constant; and a_k is an estimated parameter for variable x_k. The expected migration at periods T1 and T2 can be expressed using the following equations, assuming that there are different values in explanatory variables and model parameters in the two periods:

$$M_{ij}^{ET2} = e^{a_{0T2}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT2}} \quad (2)$$

$$M_{ij}^{ET1} = e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT1}^{a_{kT1}} \quad (3)$$

The difference in expected migration between two periods (M_{ij}^{ET1T2}) can be expressed by two major components: changes in the value of explanatory variables (M_{ij}^{VT1T2}) keeping the values of the model parameter at period T1; and changes in the value of model parameters

(M_{ij}^{PT1T2}) keeping the values of explanatory variables at period T2, in the following equation:

$$\begin{aligned} M_{ij}^{ET1T2} &= e^{a_{0T2}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT2}} - e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT1}} \\ &\quad + e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT1}} - e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT1}^{a_{kT1}} \\ &= M_{ij}^{PT1T2} + M_{ij}^{VT1T2} \end{aligned} \tag{4}$$

where $T1$ refers to values at period T1; and $T2$ refers to values at period T2. The following is assumed:

$$x_{0T1} = x_{0T2} = e \tag{5}$$

Then:

$$M_{ij}^{VT1T2} = e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT1}} - e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT1}^{a_{kT1}} \tag{6}$$

$$\begin{aligned} M_{ij}^{VT1T2} &= e^{a_{0T1}} \left[\left(\frac{x_{1T2}}{x_{1T1}} \right)^{a_{1T1}} - 1 \right] x_{1T1}^{a_{1T1}} \prod_{k=2}^{2N+1} x_{kT2}^{a_{kT1}} \\ &\quad + e^{a_{0T1}} x_{1T1}^{a_{1T1}} \prod_{k=2}^{2N+1} x_{kT2}^{a_{kT1}} - e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT1}^{a_{kT1}} \end{aligned} \tag{7}$$

$$\begin{aligned} M_{ij}^{VT1T2} &= e^{a_{0T1}} \left[\left(\frac{x_{1T2}}{x_{1T1}} \right)^{a_{1T1}} - 1 \right] x_{1T1}^{a_{1T1}} \prod_{k=2}^{2N+1} x_{kT1}^{a_{kT1}} \\ &\quad \times \prod_{k=2}^{2N+1} \left(\frac{x_{kT2}}{x_{kT1}} \right)^{a_{kT1}} + e^{a_{0T1}} x_{1T1}^{a_{1T1}} \prod_{k=2}^{2N+1} x_{kT2}^{a_{kT1}} \\ &\quad - e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT1}^{a_{kT1}} \end{aligned} \tag{8}$$

$$\begin{aligned} M_{ij}^{VT1T2} = &\left\{ \sum_{k=1}^{2N} \left[\left(\frac{x_{kT2}}{x_{kT1}} \right)^{a_{kT1}} - 1 \right] \prod_{l=k+1}^{2N+1} \left(\frac{x_{lT2}}{x_{lT1}} \right)^{a_{lT1}} \right. \\ &\left. + \left[\left(\frac{x_{2N+1T2}}{x_{2N+1T1}} \right)^{a_{2N+1T1}} - 1 \right] \right\} e^{a_{0T1}} \prod_{l=1}^{2N+1} x_{lT1}^{a_{lT1}} \end{aligned} \tag{9}$$

$$\begin{aligned} M_{ij}^{VT1T2} = &\left\{ \sum_{k=1}^{2N} \left[\left(\frac{x_{kT2}}{x_{kT1}} \right)^{a_{kT1}} - 1 \right] \right. \\ &\times \prod_{l=k+1}^{2N+1} \left(\frac{x_{lT2}}{x_{lT1}} \right)^{a_{lT1}} \\ &\left. + \left[\left(\frac{x_{2N+1T2}}{x_{2N+1T1}} \right)^{a_{2N+1T1}} - 1 \right] \right\} \\ &\times \prod_{l=0}^{2N+1} x_{lT1}^{a_{lT1}} \end{aligned} \tag{10}$$

The change in migration flow (M_{kij}^{VT1T2}), caused by a change in the value of variable x_k, is expressed as follows:

$$\begin{aligned} M_{kij}^{VT1T2} &= \left[\left(\frac{x_{kT2}}{x_{kT1}} \right)^{a_{kT1}} - 1 \right] \\ &\quad \times \prod_{l=k+1}^{2N+1} \left(\frac{x_{lT2}}{x_{lT1}} \right)^{a_{lT1}} \prod_{l=0}^{2N+1} x_{lT1}^{a_{lT1}} \end{aligned} \tag{11}$$

$(k = 1, 2, 3, ..., 2N)$

$$M_{kij}^{VT1T2} = \left[\left(\frac{x_{kT2}}{x_{kT1}} \right)^{a_{kT1}} - 1 \right] \prod_{l=0}^{2N+1} x_{lT1}^{a_{lT1}} \tag{12}$$

$(k = 2N + 1)$

If c_i refers to the change in the value of model parameter of variable i, then:

$$c_i = a_{iT2} - a_{iT1} \tag{13}$$

$$M_{ij}^{PT1T2} = e^{a_{0T2}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT2}} - e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT1}} \tag{14}$$

$$\begin{aligned} M_{ij}^{PT1T2} &= (e^{c_0} - 1) e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT2}} \\ &\quad + e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT2}} - e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT1}} \end{aligned} \tag{15}$$

$$\begin{aligned} M_{ij}^{PT1T2} &= (e^{c_0} - 1) e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT2}^{c_k} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT1}} \\ &\quad + e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT2}} - e^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT1}} \end{aligned} \tag{16}$$

$$\begin{aligned} M_{ij}^{PT1T2} &= (x_{0T2}^{c_0} - 1) \prod_{k=1}^{2N+1} x_{kT2}^{c_k} \prod_{k=0}^{2N+1} x_{kT2}^{a_{kT1}} \\ &\quad + x_{0T2}^{a_{0T1}} \prod_{k=1}^{2N+1} x_{kT2}^{a_{kT2}} - \prod_{k=0}^{2N+1} x_{kT2}^{a_{kT1}} \end{aligned} \tag{17}$$

$$\begin{aligned} M_{ij}^{PT1T2} = &\left\{ \sum_{k=0}^{2N} \left[\left(x_{kT2}^{c_k} - 1 \right) \prod_{l=k+1}^{2N+1} x_{lT2}^{c_l} \right] \right. \\ &\left. + x_{2N+1T2}^{c_{2N+1}} - 1 \right\} \prod_{l=0}^{2N+1} x_{lT2}^{a_{lT1}} \end{aligned} \tag{18}$$

The change in migration flow (M_{kij}^{PT1T2}), caused by a change in the model parameter c_k, is expressed as follows:

$$M_{kij}^{PT1T2} = (x_{kT2}^{c_k} - 1) \prod_{l=K+1}^{2N+1} x_{lT2}^{c_l} \prod_{l=0}^{2N+1} x_{lT2}^{a_{lT1}} \quad (k = 0, 1, 2, ..., 2N) \tag{19}$$

$$M_{kij}^{PT1T2} = (x_{kT2}^{c_k} - 1) \prod_{l=0}^{2N+1} x_{lT2}^{a_{lT1}} \quad (k = 2N + 1) \tag{20}$$

The expected migration flow from region i to region j has now been decomposed into two major components and detailed components caused by changes in various explanatory variables and model parameters in equations (10)–(12) and (18)–(20).

INTERREGIONAL MIGRATION MODELS

Migration models need to be estimated first before the above decomposition approach can be applied to identify the effects of changing attributes and migration model parameters on migration changes from 1985–1990 to 1995–2000 and 2005–2010. Three models are considered.

The first is the classical gravity model of migration, which can be estimated using least squares estimation (LSE) in log–linear form. The model is specified as follows:

$$\ln M_{ij} = a_0 + a_1 \ln p_i + a_2 \ln p_j + a_3 \ln d_{ij} + e_{ij} \tag{21}$$

where e_{ij} is assumed to be a normal distributed random variable.

It is noted that the above migration model (21) is estimated to minimize the residual squares of logged number of migration. According to previous studies, there are two problems in the LSE estimation (SHEN, 1999). First, the LSE estimation of the log–linear migration model aims to minimize the sum of the squares of logged ratios of the real migration size to the expected migration size. This kind of estimation criteria often results in poor fitting of large migration flows (CONGDON, 1991). Another problem of the LSE estimation is that the total numbers of the actual and expected migrants will be different due to log-transformation.

A Poisson model is a more realistic description of the migration process than a log–linear model (FLOWERDEW and AITKIN, 1982; SHEN, 1999). The second migration model is a Poisson gravity model with three variables. The total of actual migrants will be equal to the total expected migrants in the model:

$$M_{ij} = \exp(a_0 + a_1 \ln p_i + a_2 \ln p_j + a_3 \ln d_{ij}) + u_{ij} \tag{22}$$

where migration flow is assumed to be a Poisson-distributed variable; and u_{ij} is the random residual.

In recent years, multilevel models have been developed to model the random variation at different group or regional levels (JONES, 1991). BOYLE and SHEN (1997) used a multilevel modelling approach to explore the relationship between migration and individual level and regional-level factors. In terms of spatial migration, some origin or destination-specific processes might be in operation that will affect origin and destination-specific migrations systematically. Thus, a second level based on origin or destination can be specified for multilevel migration modelling. Indeed, migration between two areas is also affected by other environmental, social and economic factors (HENRY *et al.*, 2003). Thus, a multilevel Poisson migration model that also includes several socio-economic and demographic variables will be used to model inter-provincial migration in China in the periods 1985–1990 and 1995–2000. The model is specified as follows:

$$M_{ij} = \exp(a_0 + \sum_k a_k \ln x_{ik} + \sum_k a_{N+k} \ln x_{jk} + a_{2N+1} \ln d_{ij} + e_{0i}) + u_{ij} \tag{23}$$

where migration flow is assumed to be a Poisson-distributed variable at level one; and u_{ij} is the random residual at level one. It is noted that in a multilevel migration model, each sample at level two should contain a group of samples at level one so that various parameters at level two can be estimated. Either origins or destinations can be specified at level two, but not at both. Level two is defined on the basis of origin regions here. This is based on many empirical findings that migration from origins is more stable depending mainly on demographic factors than migration to destinations depending on both economic and other factors (SHEN, 1996). It is assumed that there is random variation at level two which is represented by a normally distributed random variable e_{0i}. It is noted that all other variables are specified at level one, including variables based on origins. This means that only the random variation is assumed to be different among various origins, while the coefficient of each origin or destination variable is assumed to be the same among various origins or destinations, respectively. Such a model is called a variance components model, which can be estimated using MLwiN software (RASBASH *et al.*, 2009).

In a recent review of macro-migration modelling, STILLWELL (2005) identified the following groups of determinants: gravity variables, economic variables, labour market variables, housing market variables and environment variables. Classical migration studies indicate the importance of job growth rate (LEWIS, 1955), the unemployment rate (TODARO, 1969), and migration stock (FAN, 2005b). For the analysis of migration in the periods 1985–1990, 1995–2000 and 2005–2010, eight variables are used in model (23) in addition to three variables in a gravity model (those for 1995–2000 and 2000–2005 are in parentheses below). These variables are:

GDPIi annual gross national product (GNP) growth rate over the period 1981–1989 at the origin (%)
(Annual gross domestic product (GDP) growth rate over the period 1995–2000 and 2005–2010 at the origin (%))
GDPIj annual GNP growth rate over the period 1981–1989 at the destination (%)
(Annual GDP growth rate over the period 1995–2000 and 2005–2010 at the destination (%))
ILLIj percentage of the illiterate and semi-illiterate population aged fifteen years and above in 1990 (2000 and 2010) at the destination (%)
AGRILi percentage of agricultural employment in total rural employment in 1990 (2000 and 2010) at the origin (%)
AGRILj percentage of agricultural employment in total rural employment in 1990 (2000 and 2010) at the destination (%)
POPGi percentage of population increase between 1982 and 1990 (1992–2000 and 2002–2010) at the origin. (The population decreased in 2005–2010 for a few provinces due to out-migration. Each value had 5% added so that all the values of this indicator would be positive and a log-transformation could be performed) (%)
DENSITYi population density in 1990 (2000 and 2010) at the origin (persons/km)
DENSITYj population density in 1990 (2000 and 2010) at the destination (persons/km)

These variables describe important demographic, socio-economic situations in various areas that may affect the migration process. They have been selected by stepwise regression to formulate the best model for the migration data of 1985–1990. The same set of variables is used for 1995–2000 and 2005–2010 for the decomposing analysis in this paper. *GDPIi* and *GDPIj* describe the impact of economic growth rate, thus job growth rate on migration (LEWIS, 1955). *ILLIj* describes the education level, as the education level of the population is a good indicator of development (TODARO, 1997). *AGRILi* and *AGRILj* describe the extent of rural industrialization (a smaller value means more rural industrialization) as urban and rural areas were the main destination and origin of inter-provincial migration (SUN and FAN, 2011, pp. 98–102). *POPGi*, *DENSITYi* and *DENSITYj* measure population pressure. People in areas with less land are more likely to migrate (ROZELLE *et al.*, 1999). A region with a high population density would attract fewer migrants but send out more migrants. The unemployment rate (TODARO, 1969) is not used in this study due to unreliable unemployment statistics that do not cover temporary migrants. Migration stock (FAN, 2005b) is not used in this study as the scale of migration increased dramatically in China after the 1980s.

These data for provincial regions of China are available from the DEPARTMENT OF POPULATION STATISTICS (DPS), NATIONAL BUREAU OF STATISTICS (NBS) (1991) and NATIONAL BUREAU OF STATISTICS (1990, 1991, 1996, 2001, 2011). Table 2 presents the estimation results of migration models (22) and (23) for the three periods 1985–1990, 1995–2000 and 2005–2010.

In the Poisson gravity models for 1985–1990, 1995–2000 and 2005–2010, all three variables are highly significant. The models only explain 23.72%, 6.76% and 5.06% of the variation of the number of migrants for all flows in three periods, respectively.

Table 2. Estimated parameters of Poisson gravity models and multilevel Poisson models for 1985–1990, 1995–2000 and 2005–2010

	1985–1990		1995–2000		2005–2010	
Variables	Poisson gravity model	Multilevel Poisson model	Poisson gravity model	Multilevel Poisson model	Poisson gravity model	Multilevel Poisson model
Constant	−1.0049	3.2546	−16.4362*	19.8394*	−15.7336*	−9.9599
Distance	−0.8673*	−1.1109*	−0.7316*	−1.0142*	−0.5528*	−0.6690*
Origin Population	0.6944*	0.8318*	1.1352*	1.2466*	1.3222*	1.0953*
Destination Population	0.2714*	0.6211*	0.7074*	1.7217*	0.4377*	1.0826*
GDPIi		0.2396		−1.1361		0.0726
GDPIj		2.2713*		0.2921		−0.2428
ILLIj		−0.7766*		−2.1231*		−1.2091*
AGRILi		0.0913		−0.9578		−0.0272
AGRILj		−1.9665*		−7.4981*		−1.6490*
POPGi		−0.8084		−1.2067*		−0.7011*
DENSITYi		−0.3878*		−0.4188		−0.1581
DENSITYj		−0.3565*		−1.3527		−0.1836
Level 1 variance	24101	9575	253660	103007	1002311	254034
Level 2 variance		0.0937		0.1194		0.0895
R^2 (logged M)	0.6400	0.5329	0.5417	0.5098	0.6037	0.4679
R^2 (unlogged M)	0.2372	0.4147	0.0676	0.5975	0.0506	0.1815

Note: *Significant parameter at the 0.05 level.
Source: Calculated by the author.

In the multilevel Poisson model for 1985–1990, the variations at levels one and two are both significant. All variables except *GDPIi* (the annual GNP growth rate over the period 1981–1989 at the origin), *AGRILi* (the percentage of agricultural employment in total rural employment in 1990 at the origin) and *POPGi* (the percentage of population increase between 1982 and 1990) are significant at the 0.05 level. The multilevel Poisson migration model explained over 41.47% of the variation in the number of migrants in all flows in the period 1985–1990. The key significant factors are as follows.

GDPIj (annual GNP growth rate over the period 1981–1989 at the destination) has a positive value of 2.2713 indicating a clear pulling effect of rapidly growing regions to migrants. On the other hand, *ILLIj* (the percentage of the illiterate and semi-illiterate population aged fifteen years and above in 1990 at the destination), *AGRILj* (the percentage of agricultural employment in total rural employment in 1990 at the destination), and *DENSITYj* (the population density in 1990 at the destination) have negative parameters of –0.7766, –1.9665 and –0.3565, respectively. This indicates that, if everything were equal, regions with poor education (a high percentage of the illiterate and semi-illiterate population), less rural industrialization (a high percentage of the rural population engaged in agricultural employment), and high population density were not attractive to migrants in China in 1985–1990. It is clear that inter-provincial migration in China was stimulated by economic growth and industrialization in that period.

One origin variable *DENSITYi* (the population density in 1990 at the origin) also has a negative parameter of –0.3878. Thus, if everything were equal, regions with high population density would send out fewer migrants. This seems contradictory to the common understanding that areas with high 'population pressure' (high population density) would send out more migrants. This can be explained in the dynamic spatial context of China. Areas with high population density are the most advanced areas in China and the population in such regions is less likely to move away. For example, Shandong in the eastern region had a low out-migration rate in the period 1985–1990. On the other hand, people living in areas with a high 'population pressure' may not have the willingness to move to other areas. Thus, the so called 'PUSHING' mechanism was not operating effectively in interregional migration in 1985–1990.

The result shows that the same multilevel Poisson model works to explain a significant proportion of variation for 1995–2000, 59.76%, which is even greater than the 41.47% in 1985–1990. But there are significant differences in the model parameters between multilevel Poisson models for 1985–1990 and 1995–2000. For instance, the origin and destination population parameters become bigger and are over 1. The distance parameter becomes smaller in absolute value, while the constant parameter becomes bigger. The change in the distance effect confirms the findings of Fan (2005a) but differs from the result of Li (2004) on earlier changes from 1985–1990 to 1990–1995.

For eight other socio-economic variables, five are significant at the 0.05 level in the model for 1995–2000. *ILLIj*, *AGRILj*, *DENSITYi* and *DENSITYj* are all significant and have the same negative impact on migration in the two periods. Their absolute parameters become bigger reflecting the increasing scale of migration. *GDPIi* and *AGRILi* are not significant in the two periods.

POPGi (the percentage of population increase between 1992 and 2000) has a negative impact on migration and becomes significant in 1995–2000. Thus, if everything were equal, regions with a high population growth rate would send out fewer migrants. *GDPIj* has a positive impact on migration, but it becomes insignificant in 1995–2000. However, the same multilevel Poisson model can only explain 18.15% of variation in migration flows in 2005–2010 and level two variance also becomes insignificant, indicating that a new migration model is probably needed for the migration pattern in this period. There are also significant differences in model parameters between multilevel Poisson models for 1995–2000 and 2005–2010. All parameters become smaller. The smaller distance parameter and parameter for *AGRILj* in absolute value reflect the increasing scale of migration. The distance parameter and the origin and destination population parameters remain significant in 2005–2010.

ILLIj, *AGRILj* and *POPGi* are all significant and have the same negative impact on migration in 1995–2000 and 2005–2010. *GDPIi*, *GDPIj* and *AGRILi* are not significant in the two periods. *DENSITYi* and *DENSITYj* have a negative impact on migration but become insignificant in 2005–2010. The constant parameter also becomes insignificant.

These differences indicate significant changes in migration processes. How have changes in model parameters contributed to changes in migration versus changes in the regional attributes in the period between the 1990 and 2000 censuses? The decomposition approach can be used to answer this question. The following decomposition will focus on the period between the 1990 and 2000 censuses as the same migration model for 2005–2010 can only explain a small proportion of variation in migration flows. As the multilevel Poisson model is better than the other models, they will be used in the decomposition of migration flows.

DECOMPOSING CHANGING MIGRATION FLOWS IN CHINA

The multilevel Poisson models estimated for the migration data in two periods 1985–1990 and

1995–2000 can now be used to decompose regional outflows and inflows in China. The interregional flows are calculated using the estimated migration models. The total outflows and inflows are the sum of interregional flows from or to specific provinces. Migrations rates that are problematic especially for in-migration rates are not used here (ROGERS, 1992, 1990). Among twenty-nine provinces in China, in-migration increased in twenty-six and decreased in three from 1985–1990 to 1995–2000. According to the estimated changes in migration for two periods, nineteen out of twenty-six provinces were correctly predicted to have an increase in in-migration. But seven out of twenty-six provinces were incorrectly predicted to have a decrease in in-migration, including Shandong, Hainan, Guizhou, Yunnan, Shaanxi, Gansu and Ningxia. Two out of three provinces, Anhui and Qinghai, were correctly predicted to have a decrease in in-migration. But one province, Heilongjiang, was incorrectly predicted to have an increase in in-migration. Except Shandong and Yunnan, the provinces that were incorrectly predicted involved small in-migrations at less than 0.15 million.

According to the estimated changes in out-migration for two periods, twenty-six out of twenty-eight provinces were correctly predicted to have an increase in out-migration. But two out of twenty-eight provinces, Qinghai and Ningxia, were incorrectly predicted to have a decrease in out-migration. Xinjiang was correctly predicted to have an increase in out-migration. Two migration models perform much better in the prediction of out-migration. Only three provinces were incorrectly predicted in terms of the direction of change, all involving small out-migration at less than 50000 people. Table 3 presents the changes in the size of estimated migration and the contributions by changes in the value of explanatory variables and model parameters from 1985–1990 to 1995–2000 based on the decomposing result.

According to Table 2 and Fig. 1, Guangdong and Zhejiang had the largest increase in the size of in-migration, over 2.9 million each, from 1985–1990 to 1995–2000. These are followed by Beijing, Hebei, Shanxi, Shanghai, Jiangsu, Jiangxi and Sichuan with increases in in-migration of 1–2 million each. On the other hand, in-migration decreased in some provinces

Table 3. Changes in the size of migration and the percentage contributions by changes in the value of explanatory variables and model parameters from 1985–1990 to 1995–2000

Region	In-migration			Out-migration		
	Total change	Share of variable changes (%)	Share of parameter changes (%)	Total change	Share of variable changes (%)	Share of parameter changes (%)
Beijing	1 605 896	34.52	65.48	18 447	413.23	–313.23
Tianjin	59 582	1189.89	–1089.89	37 814	295.66	–195.66
Hebei	1 553 941	92.80	7.20	1 245 593	75.55	24.45
Shanxi	1 070 964	51.00	49.00	438 881	95.36	4.64
Inner Mongolia	160 893	74.99	25.01	355 619	155.83	–55.83
Liaoning	330 596	22.14	77.86	754 695	51.69	48.31
Jilin	9549	3497.32	–3397.32	40 300	311.43	–211.43
Heilongjiang	10 331	–733.81	833.81	684 435	76.17	23.83
Shanghai	1 759 423	54.21	45.79	90 172	65.11	34.89
Jiangsu	1 706 461	75.69	24.31	1 101 971	63.19	36.81
Zhejiang	2 933 808	29.36	70.64	679 881	47.62	52.38
Anhui	–68 384	–853.90	953.90	1 874 571	77.23	22.77
Fujian	124 600	452.99	–352.99	256 856	96.43	3.57
Jiangxi	1 231 239	73.61	26.39	1 014 889	64.53	35.47
Shandong	–119 032	–425.48	525.48	1 920 878	60.22	39.78
Henan	–54 649	–1002.92	1102.92	1 827 414	63.24	36.76
Hubei	660 198	154.07	–54.07	951 067	73.08	26.92
Hunan	832 703	97.77	2.23	2 079 832	50.76	49.24
Guangdong	6 014 043	10.52	89.48	160 847	65.60	34.40
Guangxi	891 301	48.77	51.23	1 530 070	36.56	63.44
Hainan	–85 275	44.18	55.82	32 799	95.58	4.42
Sichuan	1 338 052	26.81	73.19	2 609 571	25.27	74.73
Guizhou	–123 642	–25.02	125.02	759 724	49.65	50.35
Yunnan	–146 069	–5.02	105.02	439 923	54.83	45.17
Shaanxi	–45 893	–453.14	553.14	625 321	77.87	22.13
Gansu	–66 066	–151.89	251.89	217 573	108.09	–8.09
Qinghai	–48 346	–82.60	182.60	–29 439	–350.88	450.88
Ningxia	–64 172	–84.21	184.21	–18 998	–237.71	337.71
Xinjiang	250 607	–20.11	120.11	21 956	242.66	–142.66
Total	21 722 659	62.28	37.72	21 722 662	62.28	37.72

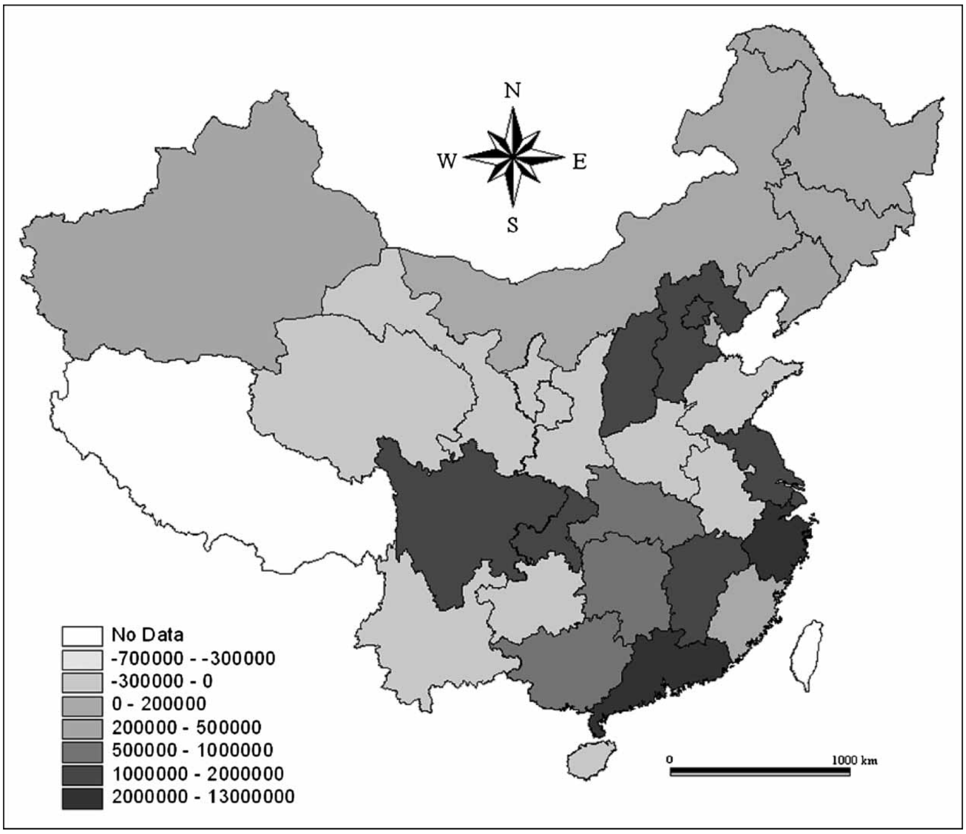

Fig. 1. Estimated total change in the in-migration to various regions from 1985–1990 to 1995–2000

such as Anhui and Qinghai in the period. Overall, the total inter-provincial migration was increased by 22 million in the above period. The decomposition result shows that 62.28% of this increase was due to changes in the value of explanatory variables, while 37.72% was due to changes in the value of model parameters. Thus, the demographic, social and economic changes among various provinces play a greater role than the changes in the parameters of migration models in the dramatic rise of the inter-provincial migration in China.

It should be noted that the changes in the value of explanatory variables and changes in the value of model parameters may have the opposite impact on the size of in-migration in some provinces. For example, the changes in the value of explanatory variables had a negative impact on the following provinces: Heilongjiang, Anhui, Shandong, Henan, Guizhou, Yunnan, Shaanxi, Gansu, Qinghai, Ningxia and Xinjiang. The changes in the value of model parameters had a negative impact in the following provinces: Tianjin, Jilin, Fujian and Hubei. In these cases the other component would contribute over 100% to the overall change in the size of in-migration.

According to Fig. 2, the in-migration to Hebei, Jiangsu and Hubei increased by 1–2 million people each due to changes in the value of explanatory variables. But the in-migration to Heilongjiang, Hainan and Xinjiang was reduced due to the same factor. Fig. 2 shows a clear pattern that the changes in the value of explanatory variables contributed a greater increase in in-migration in the eastern and central regions of China following the overall increase in in-migration in these regions.

Table 4 presents the impact of the change in the value of selected explanatory variables on the size of in-migration in various provinces. For all provinces as a whole, *Origin Population* (7.67%), *POPGi* (24.48%), *AGRILj* (16.38%), and *ILLIj* (48.36%) had a positive impact on in-migration and *ILLIj* had a large impact with 48.36% of the total change in in-migration. This indicates that the improvement in education level increased the inflow of migrants due to the negative effect of *ILLIj* in the migration model. *GDPIj* (–24.15%) and *DENSITYj* (–4.37%) had a negative impact on in-migration. As the GDP growth rate was smaller in the period 1995–2000 than in the previous period, the inflow of migrants was affected negatively. There were differences among provinces for these effects. For example, the inflow of migrants to Jiangsu increased by 57.46% and the inflow to Guangdong increased by 12.45% due to improvement in the education level.

According to Fig. 3, the in-migration to Zhejiang and Guangdong increased by 2–6 million each and to Beijing, Shanxi, Liaoning, Shanghai, Jiangsu, Jiangxi, Guangxi, Sichuan and Xinjiang by 0.2–1.1 million each due to

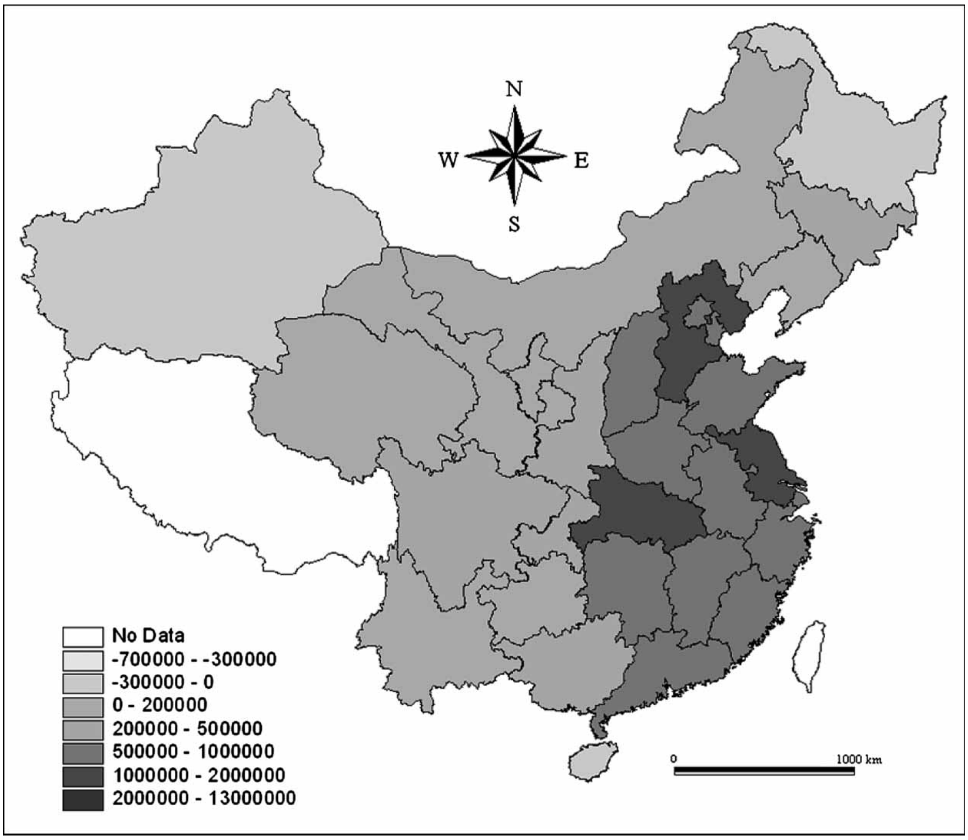

Fig. 2. Estimated effect of the change in the value of explanatory variables on the change in in-migration to various regions from 1985–1990 to 1995–2000

the changes in the value of model parameters. Clearly, changes in model parameters contributed significantly to the increase of in-migration to two top destinations, Zhejiang and Guangdong. On the other hand, in-migration to Tianjin, Jilin, Anhui, Fujian, Shandong, Henan and Hubei was reduced by 0.3–0.7 million each due to changes in model parameters.

Table 5 presents the impact of the change in the value of model parameters on the size of in-migration in various provinces. For all provinces as a whole, *Origin Population* (181.13%) and *Distance* (1512.52%) had a positive impact on in-migration. *Distance* had a very large impact, 1512.52%, of the total change in in-migration. From the migration model for 1985–1990 to the migration model for 1995–2000, the absolute *Distance* parameter was decreased by 8.70%, from –1.1109 to –1.0142. But this effect was multiplied to over a 1500% increase in the inflow of migrants. This case indicates the important role of the reduction of spatial friction in migration.

The changes in the parameters for *GDPIi* (–1.86%), *POPGi* (–3.46%), and *DENSITYi* (–28.94%) had a

Table 4. Estimated effect of the change in the value of explanatory variables on the change in in-migration to various regions from 1985–1990 to 1995–2000 (%)

Region	*Origin Population*	*POPGi*	*GDPIj*	*AGRILj*	*DENSITYj*	*ILLIj*
Beijing	5.33	18.22	–20.76	6.67	–5.58	30.72
Hebei	8.85	27.94	–3.18	18.38	–3.73	54.39
Shanxi	6.48	23.09	–19.12	5.74	–3.77	43.43
Shanghai	6.49	10.20	14.61	–7.63	–7.07	41.27
Jiangsu	7.82	34.83	–29.61	14.81	–3.69	57.46
Zhejiang	3.56	10.60	–20.98	18.68	–1.76	22.34
Jiangxi	6.59	20.96	–11.64	17.45	–2.64	50.03
Guangdong	1.57	6.28	–16.07	7.90	–2.37	12.45
Sichuan	3.89	10.73	–16.71	14.60	–1.05	20.05
Total	7.67	24.48	–24.15	16.38	–4.37	48.63

Note: Figures are percentage change in the in-migration to various regions from 1985–1990 to 1995–2000. Only regions with estimated changes over 1 million are listed.

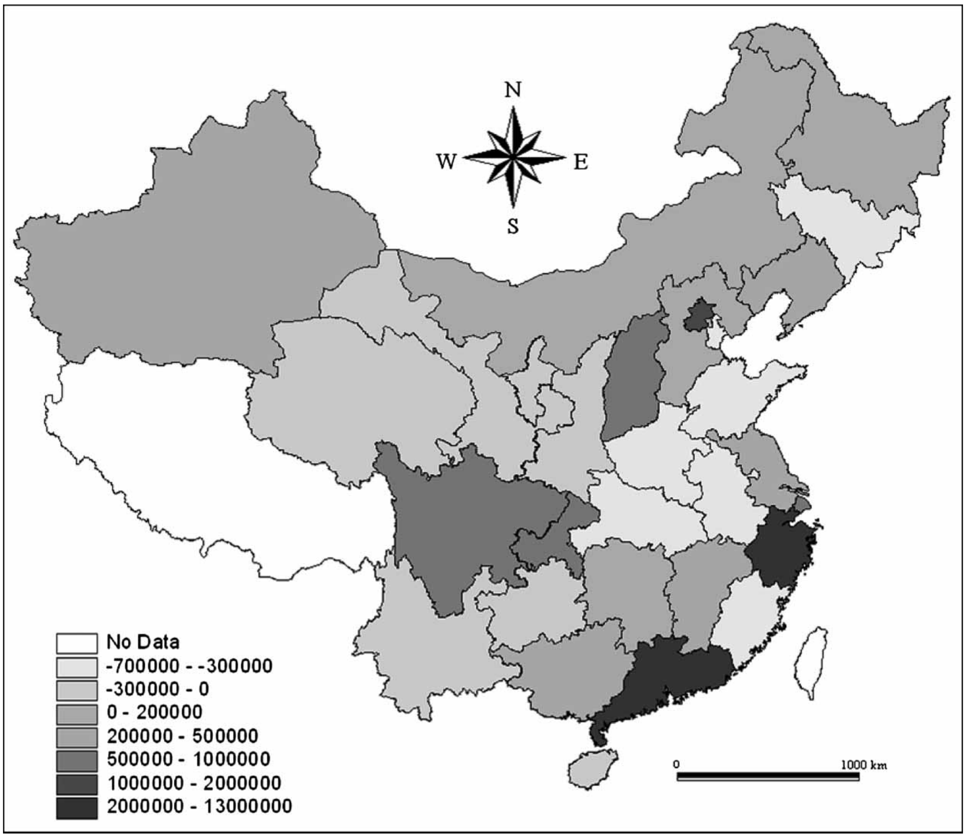

Fig. 3. Estimated effect of the change in the value of parameters on the change in in-migration to various regions from 1985–1990 to 1995–2000

negative impact on in-migration. But *GDPIi* and *POPGi* had very little impact while *DENSITYi* had a significant impact on the change in the inflow of migrants. There were differences among provinces for these effects. For example, the inflow of migrants to Zhejiang was increased by 144.16% and the inflow to Jiangsu was increased by 188.53% due to the increase in the *Origin Population* parameter.

According to Table 3 and Fig. 4, Hunan and Sichuan had the largest increase in the size of out-migration, over 2 million each, from 1985–1990 to 1995–2000. This was followed by Hebei, Jiangsu, Anhui, Jiangxi, Shandong, Henan and Guangxi with an increase in out-migration of 1–2 million each. Overall, the total inter-provincial out-migration was increased by 22 million in the above period. The decomposition result shows that 62.28% of this increase was due to changes in the value of explanatory variables, while 37.72% was due to changes in the value of model parameters. These shares are the same for both in-migration and out-migration for the country as a whole. This is because total in-migration and total out-migration are

Table 5. Estimated effect of the change in the value of parameters on the change in in-migration to various regions from 1985–1990 to 1995–2000 (%)

Region	*Origin Population*	*GDPIi*	*POPGi*	*DENSITYi*	*Distance*
Beijing	171.40	−1.84	−3.44	−27.41	2019.66
Hebei	161.03	−1.62	−3.05	−25.83	791.84
Shanxi	173.66	−1.78	−3.30	−27.54	506.07
Shanghai	165.55	−1.78	−3.37	−27.48	6233.88
Jiangsu	188.53	−1.98	−3.55	−30.93	1825.86
Zhejiang	144.16	−1.52	−2.88	−23.45	984.34
Jiangxi	148.58	−1.51	−2.79	−23.98	515.12
Guangdong	136.45	−1.36	−2.48	−21.50	948.05
Sichuan	147.92	−1.57	−3.02	−22.72	473.88
Total	181.13	−1.86	−3.46	−28.94	1512.52

Note: Figures are percentage change in the in-migration to various regions from 1985–1990 to 1995–2000. Only regions with estimated changes over 1 million are listed.

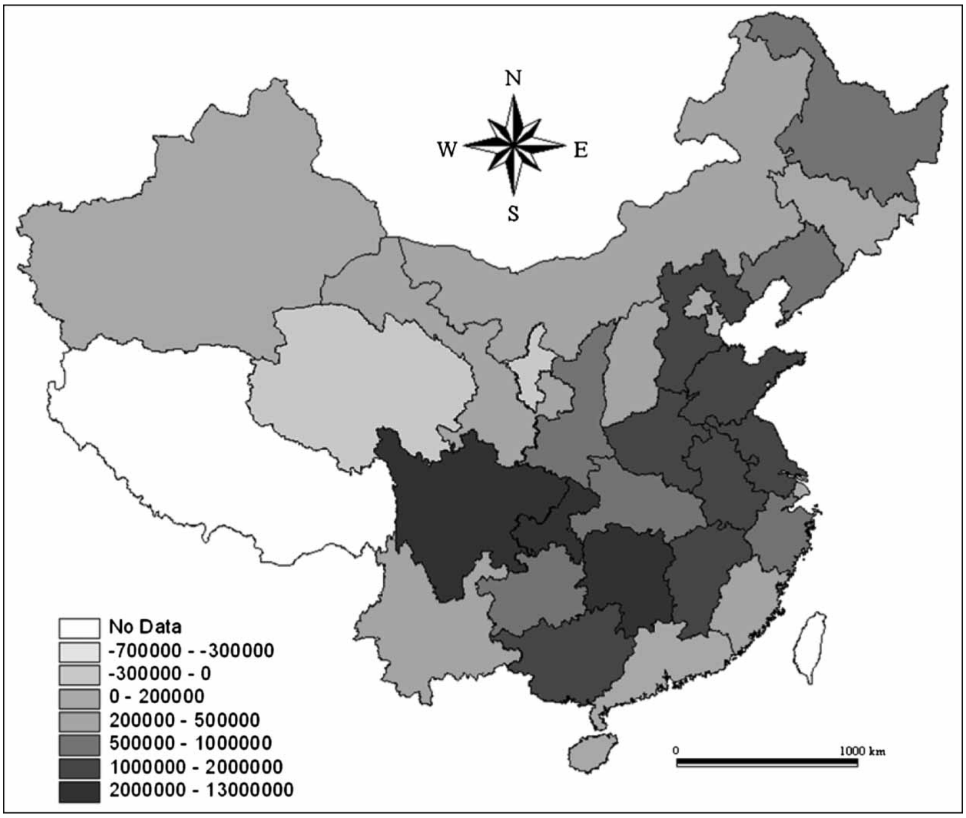

Fig. 4. Estimated total change in the out-migration from various regions from 1985–1990 to 1995–2000

the sum of inter-provincial migration flows between origins and destinations. They subject to the same migration process and underlying factors.

It is noted that the changes in the value of explanatory variables had a positive impact on out-migration from twenty-seven out of twenty-nine provinces (Table 3). But the changes in the value of model parameters had a positive impact on out-migration from twenty-three provinces and a negative impact on out-migration from six provinces, including Beijing, Tianjin, Inner Mongolia, Jilin, Gansu and Xinjiang. In these six provinces, the changes in the value of explanatory variables contributed to over 100% of the total change in out-migration. Furthermore, the changes in the value of explanatory variables contributed to 70–100% of the total change in out-migration from Hebei, Shanxi, Heilongjiang, Anhui, Fujian, Hubei, Hainan and Shaanxi. The changes in the values of model parameters contributed to 50–70% of the total change in the out-migration from Liaoning, Shanghai, Jiangsu, Jiangxi, Shandong, Henan and Guangdong.

For all provinces as a whole, the impact of the change in the value of selected explanatory variables on the size of out-migration is the same as the impact on the size of in-migration. But for individual provinces, such impacts were different on the size of out-migration and in-migration. For example, the outflow of migrants from Beijing was increased by 161.97% and the outflow from Sichuan increased by 9.55% due to different improvements in rural industrialization (*AGRILj*).

According to Fig. 5, the out-migrations from Anhui, Shandong, Henan and Hunan increased by 1–2 million each due to the changes in the value of explanatory variables. Fig. 5 shows a clear pattern that the changes in the value of explanatory variables contributed a greater increase to out-migration in the central region of China following an overall increase in out-migration in the region.

According to Fig. 6, the out-migration from Sichuan and Hunan increased by over 1 million and from Shandong, Henan and Guangxi increased by 0.5–1 million each due to the changes in the values of model parameters. Clearly, changes in model parameters contributed significantly to the increase of out-migration from the top origins, Sichuan and Hunan. On the other hand, out-migration from Beijing, Tianjin, Inner Mongolia, Jilin, Gansu, Qinghai, Ningxia and Xinjiang was reduced due to changes in the model parameters. Figs 3 and 6 show that the changes in the value of model parameters were more likely to produce both positive and negative impacts on in-migration and out-migration from various provinces than the changes in the value of explanatory variables.

For all provinces as a whole, the impact of the change in the value of an individual model parameter on the size of in-migration is the same as the impact on the size of out-migration. But for an individual province, such

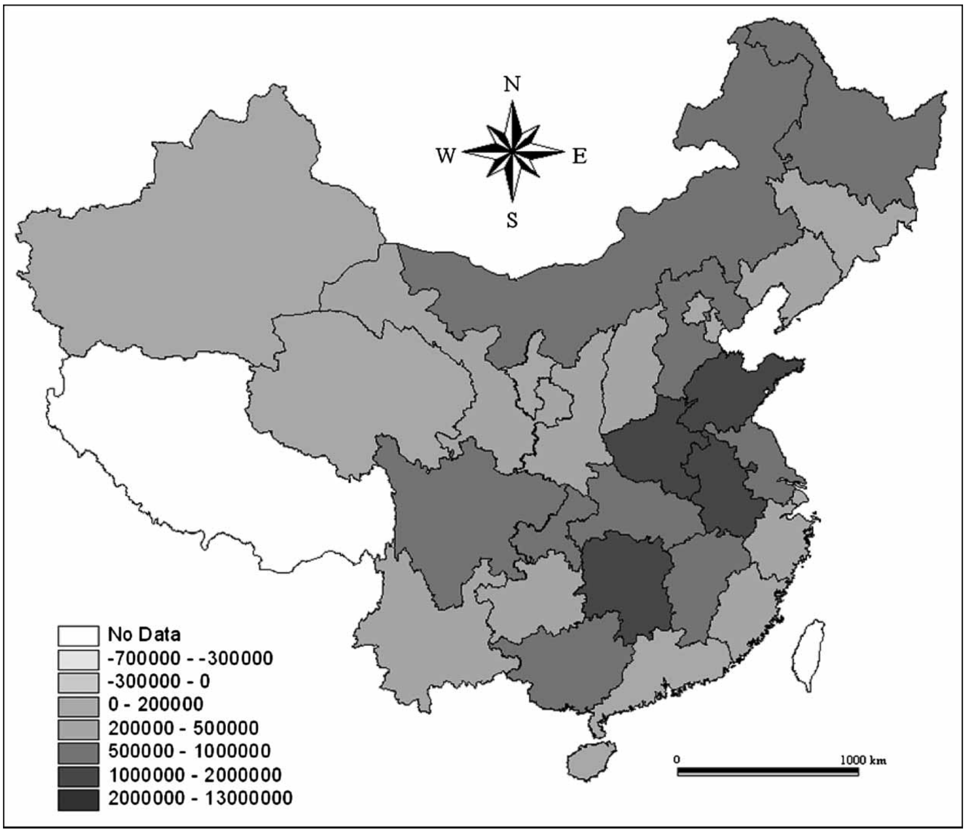

Fig. 5. Estimated effect of the change in the value of explanatory variables on the change in out-migration from various regions from 1985–1990 to 1995–2000

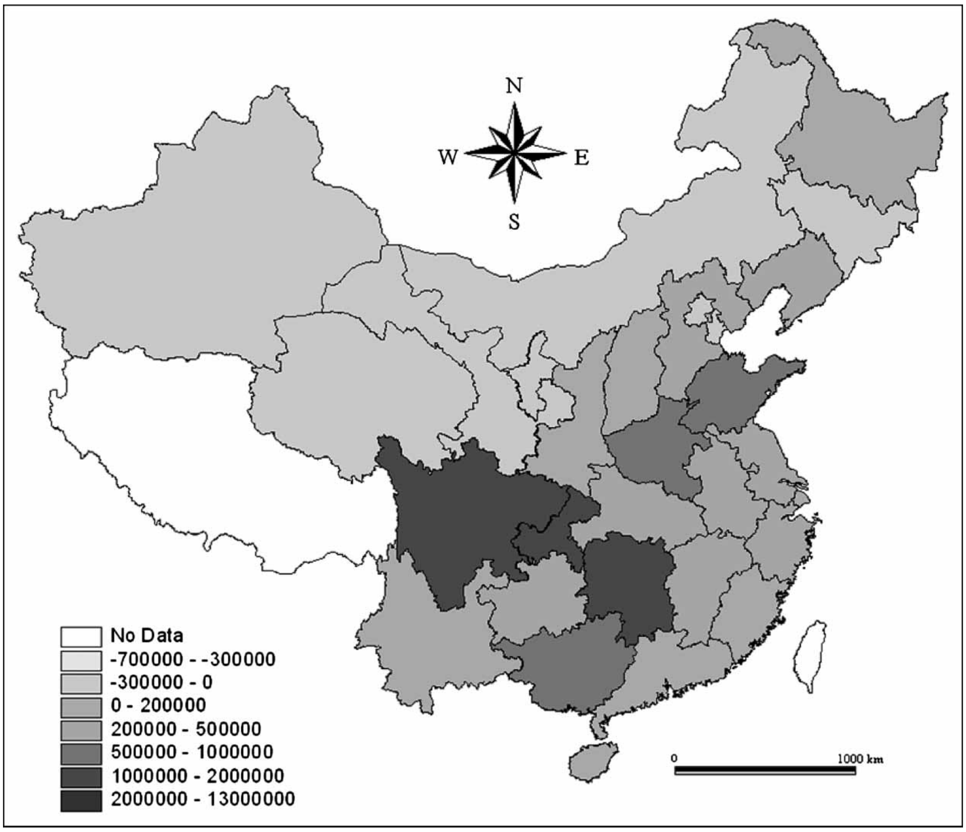

Fig. 6. Estimated effect of the change in the value of parameters on the change in out-migration from various regions from 1985–1990 to 1995–2000

impacts on in-migration and out-migration were different. For example, the outflow of migrants from Hunan increased by 150.81% and the outflow from Guangdong increased by 322.28% due to an increase in the parameter *Origin Population*.

CONCLUSIONS

Previous studies have attempted to decompose migration into sending effects, drawing effects and the separation effect (MUESER, 1989). The log–linear specification of the spatial interaction model is also used to describe the migration structure that can fully describe the relative emissiveness and attractiveness of specific regions, as well as the level of interaction between pairs of regions (ROGERS *et al.*, 2002). Making use of estimated migration models, SHEN (1999) proposed a decomposition approach to estimate the effects of spatial structure and the origin and destination attributes on migration.

This paper takes a further step to decompose the change of migration into different components. It is concerned with identifying the effects of changing parameters in migration models and the changing attributes in origin and destinations on migration flows. A decomposition approach is developed based on multilevel Poisson migration models. Inter-provincial migration models are estimated for 1985–1990, 1995–2000 and 2005–2010, respectively, using data from three censuses of China. The total in-flows and outflows are decomposed into several components reflecting the effects of various factors. The decomposition focuses on the period between the 1990 and 2000 censuses as the same migration model for 2005–2010 can only explain a small proportion of variation in migration flows. A different migration model for the period 2005–2010 with new explanatory variables can be estimated in further research. But such a different model cannot be used for decomposing migration changes between 1985–1990 and 2005–2010 as the analysis requires the same model.

Overall, total inter-provincial migration increased by 22 million in China from 1985–1990 to 1995–2000. The decomposition result shows that 62.28% of this increase was due to changes in the value of explanatory variables, while 37.72% was due to changes in the value of model parameters. Previous studies have revealed a dramatic increase in the scale of migration and changes in the spatial pattern of migration in China. Rapid economic growth in the Eastern coastal region of China has been considered a key factor of such mobility change (CAI and WANG, 2003; LIANG and MA, 2004; LI, 2004; FAN, 2005a; SUN and FAN, 2011; SHEN, 2013). However, it is unclear whether such a change is caused mainly by quantitative changes in regional development or by qualitative changes in the internal migration system such as the *hukou* system or people's propensity to migrate given the same economic stimuli. The decomposition approach used in this paper provides a clear and reliable answer that the demographic, social and economic changes among various provinces play a greater role than the changes in the parameters of migration models in the dramatic rise of inter-provincial migration in China.

This paper reveals that the same model works quite well for the periods 1985–1990 and 1995–2000, but not for the period 2005–2010. Further research is needed to reveal how migration patterns and processes have changed from 1995–2000 to 2005–2010. Preliminary analysis based on Table 1 indicates that the growth of net migration to South China (Guangdong) was slow from 1995–2000 to 2005–2010. But there was dramatic growth in net migration to East China (Zhejiang, Jiangsu and Shanghai) and North China (Beijing) in the same period, indicating a shift of the economic centre from South China to East China and North China.

Acknowledgements – The work described in this paper was supported by a research grant from the Research Grants Council of Hong Kong SAR (RGC Project Number 450107) and a Direct Research Grant from the Chinese University of Hong Kong (Project Code Number 2021070). Thanks for constructive comments and suggestions are due to two referees and the editors of the journal.

REFERENCES

ANDRIENKO Y. and GURIEV S. (2004) Determinants of interregional mobility in Russia: evidence from panel data, *Economics of Transition* **12(1)**, 1–27.

BELL M., BLAKE M., BOYLE P., DUKE-WILLIAMS O., REES P., STILLWELL J. and HUGO G. (2002) Cross-national comparison of internal migration: issues and measures, *Journal of the Royal Statistical Society A* **165(3)**, 435–464.

BOYLE P. J. and FLOWERDEW R. (1997) Improving distance estimates between areal units in migration models, *Geographical Analysis* **29(2)**, 93–107.

BOYLE P. J., FLOWERDEW R. and SHEN J. (1998) Analysing local-area migration using data from the 1991 Census: the importance of housing growth and tenure, *Regional Studies* **32(2)**, 113–132.

BOYLE P. J. and HALFACREE K. H. (1995) Service class migration in England and Wales, 1980–81: identifying gender-specific mobility patterns, *Regional Studies* **29(1)**, 43–57.

BOYLE P. J. and SHEN J. (1997) Public housing and migration: a multi-level modelling approach, *International Journal of Population Geography* **3**, 227–242.

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.

CHAN K. W., LIU T. and YANG Y. (1999) Hukou and non-Hukou migrations: comparisons and contrasts, *International Journal of Population Geography* **5(6)**, 425–448.

CLIFF A. D., MARTIN R. L. and ORD J. K. (1974) Evaluating the friction of distance parameter in gravity models, *Regional Studies* **8**, 281–286.

CLIFF A. D., MARTIN R. L. and ORD J. K. (1976) A reply to the final comment, *Regional Studies* **10**, 341–342.

CONGDON P. (1991) An application of general linear modelling to migration in London and South East England, in STILLWELL J. C. H. and CONGDON P. (Eds) *Migration Models: Macro and Micro Approaches*, pp. 113–137. Belhaven, London.

COURGEAU D. (1995) Migration theories and behavioural models, *International Journal of Population Geography* **1**, 19–27.

CROMLEY R. G., HANINK D. M. and EBENSTEIN A. Y. (2010) Estimating and mapping the intercensal internal net migration of China, 1990–2000, *Cartography and Geographic Information Science* **37(3)**, 229–238.

CURRY L., GRIFFITH D. A. and SHEPPARD E. S. (1975) Those gravity parameters again, *Regional Studies* **9**, 289–296.

DEPARTMENT OF POPULATION STATISTICS (DPS), NATIONAL BUREAU OF STATISTICS (NBS) (1991) *China Population Statistics Yearbook 1990*. China Statistics Press, Beijing.

DING J., LIU Z., CHENG D., LIU J. and ZOU J. (2005) Areal differentiation of inter-provincial migration in China and characteristics of flow field, *Acta Geographica Sinica* **60(1)**, 106–114.

FAN C. C. (2005a) Modeling interprovincial migration in China, 1985–2000, *Eurasian Geography and Economics* **46(3)**, 165–184.

FAN C. C. (2005b) Interprovincial migration, population redistribution, and regional development in China: 1990 and 2000 Census comparisons, *Professional Geographer* **57(2)**, 295–311.

FAN C. C. (2011) Settlement intention and split households: findings from a migrant survey in Beijing, China, *China Review* **11(2)**, 11–42.

FIK T. J. and MULLIGAN G. F. (1998) Functional form and spatial interaction models, *Environment and Planning A* **30**, 1497–1507.

FLOWERDEW R. and AITKIN M. (1982) A method of fitting the gravity model based on the Poisson distribution, *Journal of Regional Science* **22**, 191–202.

FOTHERINGHAM A. S. (1981) Spatial structure and distance-decay parameters, *Annals of the Association of American Geographers* **71**, 425–436.

FOTHERINGHAM A. S. (1984) Spatial flows and spatial patterns, *Environment and Planning A* **16**, 529–543.

FOTHERINGHAM A. S. (1991) Migration and spatial structure: the development of the competing destinations model, in STILLWELL J. C. H. and CONGDON P. (Eds) *Migration Models: Macro and Micro Approaches*, pp. 57–72. Belhaven, London.

FOTHERINGHAM A. S., BRUNSDON C. and CHARLTON M. (2002) *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Wiley, Chichester.

HE C. and GOBER P. (2003) Gendering interprovincial migration in China, *International Migration Review* **37(4)**, 1220–1251.

HE J. and POOLER J. (2002) The regional concentration of China's interprovincial migration flows, 1982–90, *Population and Environment* **24(2)**, 149–182.

HENRY S., BOYLE P. and LAMBIN E. F. (2003) Modelling inter-provincial migration in Burkina Faso, West Africa: the role of socio-demographic and environmental factors, *Applied Geography* **23**, 115–136.

HUA C. and PORELL F. (1979) A critical review of the development of the gravity model, *International Regional Science Review* **4(2)**, 97–126.

ISSERMAN A. M. (1985) Economic–demographic modelling with endogenously determined birth and migration rates: theory and prospects, *Environment and Planning A* **17**, 25–45.

JIANG Y. and SHEN J. (2010) Measuring the urban competitiveness of Chinese cities in 2000, *Cities* **27(5)**, 307–314.

JOHNSTON R. J. (1973) On frictions of distance and regression coefficients, *Area* **5**, 187–191.

JOHNSTON R. J. (1975) Map pattern and friction of distance parameters: a comment, *Regional Studies* **9**, 281–283.

JONES K. (1991) Specifying and estimating multi-level models for geographical research, *Transactions of Institute of British Geographers* n.s. **16**, 148–160.

KRIESBERG E. M. and VINING, JR D. R. (1978) On the contribution of out-migration to changes in net migration: a time series confirmation of Beale's cross-sectional results, *Annals of Regional Science* **12**, 1–11.

LEDENT J. (1986) Forecasting interregional migration: an economic–demographic approach, in ISSERMAN A. M. (Ed.) *Population Change and the Economy: Social Science Theories and Models*, pp. 53–77. Kluwer-Nijfoff, Boston, MA.

LEWIS W. A. (1955) *The Theory of Economic Growth*. George Allen & Unwin, London.

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.

LI W. L. and LI Y. H. (1995) Special characteristics of China's interprovincial migration, *Geographical Analysis* **27(2)**, 137–151.

LIANG Z., CHEN Y. P. and GU Y. M. (2002) Rural industrialisation and internal migration in China, *Urban Studies* **39(12)**, 2175–2187.

LIANG Z. and MA Z. (2004) China's floating population: new evidence from the 2000 Census, *Population and Development Review* **30 (3)**, 467–488.

LIANG Z. and WHITE M. J. (1997) Market transition, government policies, and interprovincial migration in China: 1983–1988, *Economic Development and Cultural Change* **45**, 321–336.

LIU Y. and SHEN J. (Forthcoming 2013a) Spatial patterns and determinants of skilled internal migration in China, 2000–2005, *Papers in Regional Sciences*.

Liu Y. and Shen J. (Forthcoming 2013b) Jobs or amenities? Location choices of interprovincial skilled migrants in China, 2000–2005, *Population, Space and Place*.

Mueser P. (1989) The spatial structure of migration: an analysis of flows between states in the USA over three decades, *Regional Studies* **23(3)**, 185–200.

National Bureau of Statistics (NBS) (1990) *Historical Statistics of Provinces in China 1949–1989*. China Statistics Press, Beijing.

National Bureau of Statistics (NBS) (1991) *China Statistics Yearbook 1991*. China Statistics Press, Beijing.

National Bureau of Statistics (NBS) (1996) *China Statistics Yearbook 1996*. China Statistics Press, Beijing.

National Bureau of Statistics (NBS) (2001) *China Statistics Yearbook 2001*. China Statistics Press, Beijing.

National Bureau of Statistics (NBS) (2011) *China Statistics Yearbook 2011*. China Statistics Press, Beijing.

Plane D. A. (1982) An information theoretic approach to the estimation of migration flows, *Journal of Regional Science* **22**, 441–456.

Plane D. A. (1987) The geographic components of change in a migration system, *Geographical Analysis* **19(4)**, 283–299.

Poncet S. (2006) Provincial migration dynamics in China: borders, costs and economic motivations, *Regional Science and Urban Economics* **36(3)**, 385–398.

Population Census Office of the State Council and Department of Population and Employment Statistics of the National Bureau of Statistics (DPES) (2012) *Tabulations on the 2010 Population Census of the People's Republic of China*, Vol. 1–3. China Statistics Press, Beijing.

Population Census Office of the State Council and Department of Population, Social, Science and Technology Statistics of the National Bureau of Statistics (DPSSTS) (2002) *Tabulations on the 2000 Population Census of the People's Republic of China*, Vol. 3. China Statistics Press, Beijing.

Population Office of the State Council and Department of Population Statistics (DPS) (1993) *Tabulations on the 1990 Population Census of the People's Republic of China*. China Statistics Press, Beijing.

Rasbash J., Charlton C., Browne W. J., Healy M. and Cameron B. (2009) *MLwiN Version 2.1*. Centre for Multilevel Modelling, University of Bristol, Bristol.

Rees P. H., Fotheringham A. S. and Champion A. G. (2004) Modelling migration for policy analysis, in Stillwell J. and Clarke G. (Eds) *Applied GIS and Spatial Analysis*, pp. 259–296. Wiley, Chichester.

Rogers A. (1990) Requiem for the net migrant, *Geographical Analysis* **22(4)**, 283–300.

Rogers A. (1992) Heterogeneity, spatial population dynamics, and the migration rate, *Environment and Planning A* **24(6)**, 775–791.

Rogers A., Willekens F., Little J. and Raymer J. (2002) Describing migration spatial structure, *Papers in Regional Science* **81**, 29–48.

Rozelle S., Guo L., Shen M., Hughart A. and Giles J. (1999) Leaving China's farms: survey results of new paths and remaining hurdles to rural migration, *China Quarterly* **158**, 367–393.

Shen J. (1996) Internal migration and regional population dynamics in China, *Progress in Planning* **45(3)**, 123–188.

Shen J. (1999) Modelling regional migration in China: estimation and decomposition, *Environment and Planning A* **31**, 1223–1238.

Shen J. (2013) Increasing internal migration in China from 1985 to 2005: institutional versus economic drivers, *Habitat International* **39(1)**, 1–7.

Sheppard E. S. (1979) Gravity parameter estimation, *Geographical Analysis* **11**, 120–132.

Stillwell J. (2005) Inter-regional migration modelling: a review and assessment. Paper presented at the 45th Congress of the European Regional Science Association (ERSA), the Netherlands, 23–27 August 2005.

Stillwell J. and Congdon P. (Eds) (1991) *Migration Models: Macro and Micro Approaches*. Belhaven, London.

Sun M. and Fan C. C. (2011) China's permanent and temporary migrants: differentials and changes, 1990–2000, *Professional Geographer* **63(1)**, 92–112.

Todaro M. P. (1969) A model of labor migration and urban unemployment in less developed countries, *American Economic Review* **59**, 138–148.

Todaro M. P. (1997) *Economic Development*. Longman, London.

Van Imhoff E. V., van der Gaag N., van Wissen L. and Rees P. H. (1997) The selection of internal migration models for European regions, *International Journal of Population Geography* **3**, 137–159.

Wei Y. (1997) Interegional migration in socialist countries: the case of China, *GeoJournal* **41**, 205–214.

Willekens F. (1983) Log–linear modelling of spatial interaction, *Papers of the Regional Science Association* **52**, 187–205.

Wilson A. G. (1970) *Entropy in Urban and Regional Modelling*. Pion, London.

Yang H. (2000) A comparative analysis of China's permanent and temporary migration during the Reform Period, *International Journal of Social Economics* **27(3)**, 173–193.

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.

Zhu Y. and Chen W. (2010) The settlement intention of China's floating population in the cities: recent changes and multifaceted individual-level determinants, *Population, Space and Place* **16(4)**, 253–267.