Changing Patterns and Determinants of Interprovincial Migration in China 1985–2000

Jianfa Shen*
Department of Geography and Resource Management, Urban and Regional Development Programme, Hong Kong Institute of Asia-Pacific Studies, The Chinese University of Hong Kong, Shatin, N.T. Hong Kong, China

ABSTRACT

Most migration analyses focus on inter-regional migration in one particular period. Recently, efforts have also been made to compare migration intensity and patterns across countries. Instead of comparison over space, this paper is concerned with the temporal dimension. Can the same model, using the same set of variables, explain migration patterns in two different periods? Is there any change in the determinants of interprovincial migration in China? This paper attempts to answer these questions using migration data from the 1990 and 2000 censuses of China. It was found that coastal regions such as Guangdong, Zhejiang, Tianjin, and Fujian joined Shanghai and Beijing as the top destinations of migration over time. ‘Pull’ factors in the growing regions play an important role in the expanding migration in China. The same set of variables was used to explain the migration in 1985–1990 and 1995–2000. The same model can explain 51.6% of the variation of migration in 1995–2000, greater than 41.6% in 1985–1990. Nevertheless, absolute parameters in the model were bigger in 1995–2000 than in 1985–1990 to reflect the increasing scale of migration. A better model for 1995–2000 is also established using a new set of variables. Two income variables have a positive impact on migration, confirming the importance of income in internal migration in China. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: migration model, migration pattern, determinants, census data, China

INTRODUCTION

Migration is an important component of demographic change (Bell et al., 2002; Henry et al., 2003; Andrienko and Guriev, 2004; Dennett and Stillwell, 2010). With a declining natural population increase rate in China and many developed countries, migration becomes very important in population redistribution in a country and has important implications for socio-economic development. A key distinction in migration modelling is that between micro and macro approaches (Stillwell and Congdon, 1991). A common approach in macro migration analysis is to establish migration models using various variables to explain migration flows directly. A brief review of macro migration models is provided here. Hua and Porell (1979) offered a detailed review of the development of the gravity model (GM), and Stillwell (2005) made an excellent review of the recent development in inter-regional migration modelling. The GM was first developed by Zipf (1946) involving only three variables: distance and populations at the origin and at the destination. Wilson (1967, 1970) introduced balancing factors and formulated spatial interaction models based on entropy-maximising techniques. Alonso (1978) also formulated similar models. Their models ensure consistency in the observed and predicted sum of flows, that is, in-migration totals or out-migration totals. The model was extended by Fotheringham (1983, 1986, 1991) by the incorporation of a competing destinations variable to remove the effect of spatial structure.

* Correspondence to: Jianfa Shen, Department of Geography and Resource Management, Urban and Regional Development Programme, Hong Kong Institute of Asia-Pacific Studies, The Chinese University of Hong Kong, Shatin, N.T. Hong Kong, China. E-mail: jianfa@cuhk.edu.hk

Copyright © 2011 John Wiley & Sons, Ltd.
Spatial interaction models have also been formulated statistically. A typical log-linear GM with a random error term is equivalent to Wilson’s unconstrained spatial interaction model. Lowry (1966) extended the set of independent variables in addition to population and distance variables. Extended demo-economic models have also been developed (Isserman, 1985; Stillwell and Congdon, 1991; Henry et al., 2003).

Willekens (1983) demonstrated the relationship between the GM and the log-linear model and also the meanings of balancing factors and the first-order interaction effects in the log-linear model. Rogers et al. (2002) suggested the use of the log-linear specification of the spatial interaction model to describe the migration structure.

Many studies on migration have focused on the issue of model specification (Johnston, 1975; Kau and Sirmans, 1979; Fotheringham, 1981, 1991; Congdon, 1991) as well as distance measurement (Boyle and Flowerdew, 1997). In a case study on 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 the migration process than a log-linear model (Flowerdew and Aitkin, 1982; Boyle and Halfacree, 1995). Multilevel models have also become popular to model random variations at different groups or regional levels (Jones, 1991) and have been used to explore the relationship between migration and the factors at individual and regional levels (Boyle and Shen, 1997). The Poisson migration model will be used in this study.

Most previous models are one-step models modelling migration flows between regions simultaneously. Considering migration as a two-stage process where migrants first decide to leave and then choose a suitable destination, two-stage migration models have also been developed (Champion et al., 2003; Fotheringham et al., 2004; Rees et al., 2004).

Most migration analyses focus on inter-regional migration in one particular period with a few exceptions. Parameters can also be compared for models estimated for different periods with particular attention to the change distance decay effect on migration (Ledent, 1986; Li, 2004; Fan, 2005a). Ledent (1986) studied the temporal changes in the parameters when the migration models were separately estimated for each of a time series of Canadian migration matrices. This paper will compare the results of modelling migration in China in two different periods. Can the same models using the same set of variables explain migration patterns in two different periods, given the increased scale of migration? Are there better models and new determinants that can explain migration patterns in a new period?

Previous studies on interprovincial migration in China have focused on the spatial patterns of temporary and permanent migration (Chan et al., 1999; Yang, 2000; Zhu, 2003; Liang and Ma, 2004; Zhu and Chen, 2010), the regional concentration of migration flows (He and Pooler, 2002; Ding et al., 2005), the determinants of migration (Li and Li, 1995; Liang and White, 1997; Wei, 1997; Cai and Wang, 2003; Li, 2004; Fan, 2005a), the gender difference in migration (He and Gober, 2003), and the relationship between migration and regional development (Fan, 1996, 2005b). Various migration models were estimated to identify the effects of spatial structure on interprovincial migration using 1990 census data (Shen, 1999). Yang and Guo (1999) examined the determinants of temporary migration in Hubei province. They found that community-level factors were very important. Provincial-level factors that will be analysed in this research are expected to be even more important in migration, given vast differences among Chinese provinces (Fan, 1995; Wei, 1999). In summary, government policy, rural reform, household registration system reform, regional development, development of non-state-owned sectors, and foreign investments have been identified as major factors contributing to increasing migration in China.

With dramatic changes in the patterns of regional development, the scale of interprovincial migration increased from just 12 million in 1985–1990 to 33 million in 1995–2000 in China. Does such an increase just mean proportional expansion of migration in all directions? Is there any change in the determinants of interprovincial migration in China? Can we use the same set of explanatory variables to explain interprovincial migration in 1985–1990 and 1995–2000, given rapid social and economic changes in China?
This paper attempts to answer the aforementioned questions, which have important methodological implications, using migration data from the 1990 and 2000 censuses of China. The paper is organised as follows: the changing patterns of interprovincial migration in China is described in the next section. Then, migration models will be estimated to examine the main factors of migration in the two periods. Some conclusions are reached in the final section.

CHANGING PATTERNS OF INTERPROVINCIAL MIGRATION IN CHINA 1985–2000

Reliable migration data were collected in the 1990 and 2000 population censuses in China. The 1990 census considered the following people as migrants: those whose usual place of residence changed over the 5-year period 1 July 1985 to 1 July 1990. The 2000 census considered the following people as migrants: those whose usual place of residence changed in the 5-year period 1 November 1995 to 1 November 2000. Because of well-known inconsistencies between household registration and the residence of many individuals, the current place of residence was considered as the usual residence of the migrants if they had left their place of household registration for over a year in the 1990 census or over half a year in the 2000 census.

No migration data on migrants moving into Tibet were collected in the 1990 census. Based on the data from the 2000 census, the estimated numbers of migrants moving in and out of Tibet were 77,506 and 38,472, respectively, in the period 1995–2000, after making appropriate adjustments for undercounting in the census. For data consistency, migration between Tibet and the other regions in China will not be considered in this paper. The migration data sets used here refer to the migration flows among the 29 provincial regions in mainland China, except Tibet, in the census periods 1985–1990 and 1995–2000 (Population Census Office of State Council and DPS, 1993; Population Census Office of State Council and DPSSTS, 2002). Figure 1 shows the provincial regions in China. Here, migration is measured based on the number of migrants by using a transition approach instead of a movement approach.

Since the early 1980s, migration control has been relaxed because of the changes in the household registration system in China. In 1984, rural migrants who engaged in various kinds of small business enterprises in towns were allowed to register as a ‘household with self-supplied food grain’. In 1985, the registration of temporary population was introduced, allowing people to move to places away from the place of household registration and paving the way for large-scale migration in China. The majority of migrants in China today belong to the category of temporary population whose household registration is still in their hometown. The volume of internal migration in China has been increasing since then. 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). Interprovincial migration in China reached 11.01 million in 1985–1990. Nevertheless, 1985–1990 was an initial period of increasing mobility as there was still a keen debate among the public, scholars, and government officials as to whether such migration was desirable. The policy was tightened at some stage, such as in 1989, when society was perceived to be unstable. China went back to the reform track in 1992, and developing socialist market economy was officially adopted as the objective of reform in 1994 (Shen, 2007). Economic growth has sped up, and the society has become more receptive to migrants since then. They became the main driver of increasing mobility in the period 1995–2000 (Shen et al., 2002; Shen, 2010). Thus, the scale of interprovincial migration doubled, reaching 33.89 million in the period 1995–2000. Although many studies have been done on migration in 1985–1990 and 1995–2000 (He and Pooler, 2002; Li, 2004; Ding et al., 2005; Fan, 2005a, b), there is a need to do a systematic study on the changes in interprovincial migration in China in terms of spatial patterns and the driving factors between the two census periods. This paper will examine these two aspects in great detail.

This section will examine the total in-migration, out-migration, and net migration to and from the provincial regions in China in order to provide an overview of migration patterns. They refer to the marginals of the 1985–1990 and 1995–2000 migration flow matrices. Table 1 presents in-migration,
net migration, and migration efficiency of the 29 regions in the two periods 1985–1990 and 1995–2000, sorted by the net migration in 1995–2000. Some important observations can be made. First, Guangdong, Shanghai, and Beijing, the three most advanced provincial units, remained the top three in the scale of net migration in the two periods. 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 27 with a large net out-migration in 1985–1990, but it changed into the fourth top region with a large net in-migration. On the other hand, Hubei’s case was just the opposite. It was ranked eighth in net migration in 1985–2000, but it changed into a bottom province ranked 23rd in 1995–2000, losing much of its population through migration. Third, migration efficiency is the ratio of net migration to the total of in-migrations and out-migrations. Migration efficiency was increased significantly in most provincial regions, indicating that the migrants were moving into specific regions without much return migration flow. In the period 1985–1990, the largest migration efficiency was around +60% to +70% or −60% in Guangdong, Shanghai, Beijing, and Guangxi. In the period 1995–2000, the largest migration efficiency increased to around +80% to +90% or −80% in Guangdong, Shanghai, Beijing, Jiangxi, and Anhui. Fourth, the scale of in-migration and net migration in the top regions was increased much more significantly than the national average of about two times from 1985–1990 to 1995–2000. The in-migration to Guangdong, Zhejiang, and Fujian was increased by nearly 9, 8, and 5 times, respectively. The net migration to Fujian, Yunnan, Xinjiang, and Guangdong was increased by 63, 15, 13, and 11 times, respectively. The net migration from Jiangxi, Hubei, Henan, Anhui, and Hunan was increased by 36, 21, 16, 13, and 11 times, respectively. Clearly, migration in China became more concentrated in some provinces as destinations and origins.

In 1995–2000, Guangdong received the largest net gain of 11.92 million migrants because of its rapidly expanding economy and heavy inflow of capital from Hong Kong, Macao, and foreign countries (Shen, 2006, 2008a; Lu and Wei, 2007). Beijing, the capital of China, and Shanghai, the
leading economic centre of China (Yeung and Shen, 2008), received a net gain of 2.12 and 1.80 million migrants, respectively. Zhejiang, a booming coastal province near Shanghai (Wei et al., 2007), also received a net gain of 1.79 million. The most populous province, Sichuan, is the largest loser of migrants during the period (4.67 million). Hunan, Anhui, and Jiangxi also lost 3.10 million, 2.73 million, and 2.59 million people through migration in the period 1995–2000 (Shen, 2008b).

The overall picture of migration gains in China is as follows: those regions with rapidly expanding economies are usually the top destinations of migrants. The main source of migrants is not necessarily the poorest regions in the country but the medium and less developed regions. Main origin and destination regions have high migration efficiency.

The above analysis is based on the total number of migrants moving between provinces. It is affected by the population size of provinces, as a large province would certainly send out or receive more migrants. To measure the intensity or attractiveness to migrants, in-migration rates and out-migration rates are calculated. The eastern, central, and western regions are considered separately because of their different


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In-migration</td>
<td>Net migration</td>
</tr>
<tr>
<td>Guangdong</td>
<td>12,397,081</td>
<td>11,920,728</td>
</tr>
<tr>
<td>Shanghai</td>
<td>2,288,563</td>
<td>2,115,346</td>
</tr>
<tr>
<td>Beijing</td>
<td>1,982,498</td>
<td>1,797,478</td>
</tr>
<tr>
<td>Zhejiang</td>
<td>2,828,915</td>
<td>1,794,503</td>
</tr>
<tr>
<td>Xinjiang</td>
<td>1,118,717</td>
<td>799,445</td>
</tr>
</tbody>
</table>

Based on migration among the 29 regions in China, excluding to and from Tibet.
demographic–economic conditions (Yeung and Shen, 2004; Jiang and Shen, 2010).

According to Figure 2, most provinces in eastern China had in-migration rates above the national average in 1985–1990 and 1995–2000, but the in-migration rates were below the national average in Shandong and Hebei in both periods. The in-migration rates were also below the national average in Zhejiang and Fujian in 1985–1990 and Liaoning in 1995–2000. Shanghai and Beijing had in-migration rates over 4.9% in 1985–1990, well above the other provinces. Their in-migration rates then jumped to over 13% in 1995–2000 along with Guangdong. The net migration rates to Beijing, Guangdong, and Shanghai increased dramatically in the same period (Fig. 3). Zhejiang, Tianjin, and Fujian had a similar but less significant jump in in-migration rates from below 2.8% in 1985–1995 to over 4.2% in 1995–2000. Over time, Guangdong, Zhejiang, Tianjin, and Fujian joined Shanghai and Beijing as the top destinations of migration, but this belt of high in-migration rate spatially focused on the south part of the eastern coastal region from Shanghai to Guangdong, plus Beijing and Tianjin. The in-migration rates were modest in Hainan, Jiangsu, Liaoning, Shandong, and Hebei, slightly above or even below the national average in both periods. Interprovincial labour migration was not strong to these provinces. There is clear difference among the provincial units in the eastern region.

In the central part of China, almost all provinces had in-migration rates below the national average, except that Shanxi and Heilongjiang had slightly higher in-migration rates than the national average in 1985–1990. All in-migration rates were below 1.3%. In the western part of China, almost all provinces had an in-migration rate below the national average except Xinjiang in both periods and Qinghai, Ningxia, and Inner Mongolia in
1985–1990. Xinjiang, Qinghai, Ningxia, and Inner Mongolia were traditional destinations of migration in the pre-reform period. The in-migration rates in these regions had only a modest increase or even decline from 1985–1990 to 1995–2000, confirming the reversal of migration direction in the reform period (Shen, 1996). Xinjiang was exceptional as its in-migration rate was increased significantly and was above the national average in the period 1985–2000 as a result of significant agricultural and resource development there (Pannell and Ma, 1997; Becquelin, 2004). Clearly, the central and western regions were no longer the main destinations of migration in China in the reform period. However, some provincial units there had become the main origins of migration.

In the central region, the out-migration rates increased dramatically in Jiangxi, Hunan, Anhui, and Hubei. They were below the national average in 1985–1990 but well over the national average in 1995–2000 (Fig. 4). There was a significant increase in the net migration rates from these four provinces (Fig. 5). The out-migration rates of Henan and Shanxi also increased but still below the national average in both periods. The out-migration rates in Heilongjiang and Jilin only increased slightly, changing from above the national average in 1985–1990 to below the national average in 1995–2000. Heilongjiang and Jilin were also traditional destinations of migration in the pre-reform period. Geographically, Heilongjiang, Jilin, and Shanxi were far away from the major economic growth centres, reducing their out-migration in the reform period.

The out-migration rate in the provinces of the western region was above the national average in 1985–1990, except Yunnan (Fig. 6), but the out-migration rate changed differently in the period 1985–2000. The out-migration rate in Sichuan, Guangxi, and Guizhou jumped from below 1.4% in 1985–1990 to over 3.6% in 1995–2000. The net migration rate from there also increased significantly (Fig. 7). They became important sources of migrants such as Anhui, Jiangxi, Hunan, and Hubei in the central region. The out-migration rates in the other provinces were only increased slightly, changing from above the national average in 1985–1990 to below the national average in 1995–2000. The out-migration rate was below the national average in Yunnan in both periods. Ningxia, Xinjiang, and Yunnan had positive net migration rates in 1995–2000. The net migration rate to Xinjiang increased significantly as a result of the increase in in-migration rates and the decrease in out-migration rates in the same period (Fig. 7).

Generally, most provinces in the eastern region had out-migration rates below the national average in the period 1985–2000 although it increased slightly during the period. All out-migration rates were below 2.3%. During the period 1985–1990, when the national out-migration rate was low, the out-migration rates in Zhejiang, Beijing, Hainan, Hebei, and Shanghai were above the national average. As mentioned before, Beijing and Shanghai had high in-migration rates even in the period 1985–1990, so they may have had high out-migration rates too. Hebei, Hainan, and Zhejiang had also been sources of migrants to nearby areas.
such as Beijing, Guangdong, and Shanghai, but by 1995–2000, their out-migrant rates were overtaken by some provinces in the central and western regions.

Thus, the spatial pattern of migration in China has changed significantly from 1985–1990 to 1995–2000. Over time, Guangdong, Zhejiang, Tianjin, and Fujian joined Shanghai and Beijing as
the top destinations of migration. Some provinces with high in-migration rates in 1985–1990 such as Hainan and Qinghai were overtaken by the aforementioned top destinations. Only Xinjiang in the western region kept a high in-migration rate in both periods. Only some provinces in the central and western regions, including Jiangxi, Hunan, Anhui, Hubei, Sichuan, Guangxi, and Guizhou had made use of a growing economy in the eastern coastal region by sending large numbers of migrants in 1995–2000. The top origins of out-migration, in order of out-migration rate, were Jiangxi, Hunan, Sichuan, Anhui, Guangxi, Hubei, and Guizhou in 1995–2000. Contrary to common understanding, Jiangxi and Hunan had higher out-migration rates than Sichuan. Some provinces, including Henan, Heilongjiang, Jilin, Shanxi, Shaanxi, Gansu, Inner Mongolia, Qinghai, Yunnan, Ningxia, and Xinjiang, had not established a strong migration link with the coastal region. Clearly, migration is a selective spatial process affected by distance and socio-economic factors. Migration models will be estimated to examine the main factors of migration in the two periods in the next section.

IDENTIFY THE MAIN FACTORS OF MIGRATION USING INTER-REGIONAL MIGRATION MODELS

Specification of Inter-Regional Migration Models

In this paper, migration models will be estimated for China for the periods 1985–1990 and 1995–2000. One classic migration model is the GM, which includes a distance variable, $d_{ij}$, and populations at origin and destination $p_i$ and $p_j$. The model can be estimated using least squares estimation (LSE) in a log-linear form as follows:

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

(1)

Here, $e_{ij}$ is assumed to be a normal distributed random variable.

It is noted that the above migration model is estimated in order to minimise the residual squares of logged number of migration. According to previous studies, there are two problems in the LSE (Shen, 1999). First, the LSE of the log-linear migration model aims to minimise the sum of the squares of logged ratios of the real migration size to expected migration size. This kind of estimation criteria often results in poor fitting of large migration flows (Congdon, 1991). Second, the total numbers of the actual and expected migrants will be different because of the 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). A Poisson GM with three variables can be described as follows:

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

(2)

Here, the 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). The multilevel modelling approach has also been used to examine the relationship between migration and individual level and regional level factors (Boyle and Shen, 1997). In terms of spatial migration, some origin-specific or destination-specific processes might be in operation, which will affect origin-specific and destination-specific migrations systematically. Thus, a second level based on origin or destination can be specified for multilevel migration modelling. This level 2 is defined on the basis of origin regions. 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). This is also confirmed by a larger parameter for origin population than destination population in this study. Indeed, migration between two areas is also affected by other environmental, social, and economic factors (Henry et al., 2003). Thus, a multilevel extended Poisson model (EPM), which also includes several socio-economic and demographic variables, will be used to model interprovincial migration in China in the periods 1985–1990 and 1995–2000. The model is specified as follows:

$$M_{ij} = \exp\left(a_0 + \sum_k a_{1k} \ln x_{ik} + \sum_k a_{2k} \ln x_{jk} + b \ln d_{ij} + e_{ij}\right) + u_{ij}$$  

(3)

This is a simple two-level model with random variations at level 1 (all interprovincial flows)
and level 2 (origin region) (Jones, 1991). Here, migration flow is assumed to be a Poisson-distributed variable at level 1, and $u_{ij}$ is the random residual at level 1. There is random variation at level 2, represented by a normally distributed random variable $v_{ij}$. The model can be estimated using MLwiN software (Centre for Multilevel Modelling, University of Bristol, Bristol) (Rasbash et al., 2009). Extra Poisson variation at level 1 is assumed, as the model fails to converge if using a fixed Poisson variation of 1.

This paper will use the multilevel extended EPM with gravity and socio-economic variables to describe the interprovincial migration in China. The Poisson GM including only GM variables is also estimated for comparison.


For the analysis of migration in the period 1985–1990, eight variables were used in multilevel EPM in addition to the three variables in a GM. These variables are as follows:

- **GDPI$_i$**: Annual gross national product (GNP) growth rate over the period 1981–1989 at origin, %
- **GDPI$_j$**: Annual GNP growth rate over the period 1981–1989 at destination, %
- **ILLI$_j$**: Percentage of illiterate and semi-illiterate population aged 15+ in 1990 at destination, %
- **AGRIL$_i$**: Percentage of agricultural employment in total rural employment in 1990 at origin, %
- **AGRIL$_j$**: Percentage of agricultural employment in total rural employment in 1990 at destination, %
- **POPG$_i$**: Percentage of population increase between 1982 and 1990 at origin, %
- **DENSITY$_i$**: Population density in 1990 at origin, persons/km$^2$
- **DENSITY$_j$**: Population density in 1990 at destination, persons/km$^2$

These variables describe important demographic and socio-economic situations in various areas, which may affect the migration process. They have been selected by stepwise regression for the migration data of 1985–1990. Annual GNP growth rate over the period 1981–1989 was used to reflect the accumulated impact on migration in the 1980s. For most other indicators, data at the end of the period were used, as reliable data are only available from the census. GDPI$_i$ and GDPI$_j$ describe the impact of economic growth rate on migration. A region with a high economic growth rate would attract more migrants but send out fewer migrants. Thus, it is expected that GDPI$_i$ and GDPI$_j$ should have negative and positive impacts on migration at the origin and the destination, respectively. ILLI$_j$ describes the education level. A region with a high proportion of low-educated population would not be attractive to migrants. Thus, it is expected that ILLI$_j$ should have negative impact on migration at the destination. AGRIL$_i$ and AGRIL$_j$ describe the extent of rural industrialisation (low value means more rural industrialisation). A region with a high level of rural industrialisation would attract more migrants but send out fewer migrants. Thus, AGRIL$_i$ and AGRIL$_j$ should have positive and negative impacts on migration at the origin and the destination, respectively. POPG$_i$, DENSITY$_i$, and DENSITY$_j$ measure population pressure. A region with a high population growth rate or high population density would attract fewer migrants but send out more migrants. Thus, POPG$_i$ and DENSITY$_j$ should have positive impacts on migration at the origin, and DENSITY$_i$ should have negative impacts on migration at the destination.

To test whether the same set of variables is still good enough to explain the new migration pattern in 1995–2000, a set of similar variables was used for the analysis of migration in the period 1995–2000.

- **GDPI$_i$**: Annual gross domestic product (GDP) growth rate over the period 1995–2000 at origin, %
- **GDPI$_j$**: Annual GDP growth rate over the period 1995–2000 at destination, %
- **ILLI$_j$**: Percentage of illiterate and semi-illiterate population aged 15+ in 2000 at destination, %
- **AGRIL$_i$**: Percentage of agricultural employment in total rural employment in 2000 at origin, %
- **AGRIL$_j$**: Percentage of agricultural employment in total rural employment in 2000 at destination, %
- **POPG$_i$**: Percentage of population increase between the 1992 and 2000 census
The significance of the variables in the multilevel EPM for the period 1985–1990 is further illustrated by the estimated parameters. All variables except GDPI$_{i}$ (annual GNP growth rate over the period 1981–1989 at origin) and AGRIL$_{i}$ (percentage of agricultural employment in the total rural employment in 1990 at origin) are significant at 0.05 level. The multilevel extended Poisson migration model explained over 41.6% of the variation of the number of migrants in all flows in the period 1985–1990. This is a significant improvement over the Poisson GM. Compared with similar estimations of migration models, the model for 1985–1990 can be considered satisfactory in explaining internal migration in China in that period. The key significant factors are as follows.

The variable GDPI$_{j}$ (annual GNP growth rate over the period 1981–1989 at destination) has a positive value of 2.2706 indicating the clear pulling effect of rapidly growing regions to migrants. On the other hand, ILLI$_{j}$ (percentage of illiterate and semi-illiterate population aged 15+ in 1990 at destination), AGRIL$_{j}$ (percentage of agricultural employment in the total rural employment in 1990 at destination), and DENSITY$_{j}$ (population density in 1990 at destination) have negative parameters of −0.7749, −1.9565, and −0.3537, respectively. These four variables have the expected impacts on migration mentioned previously. This indicates that, if everything was equal, regions with poor education (a high percentage of illiterate and semi-illiterate population), less rural industrialisation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>−1.0346</td>
<td>2.4785</td>
</tr>
<tr>
<td>Distance</td>
<td>−0.8641*</td>
<td>−1.1061*</td>
</tr>
<tr>
<td>POP$_{i}$</td>
<td>0.6948*</td>
<td>0.8289*</td>
</tr>
<tr>
<td>POP$_{j}$</td>
<td>0.2715*</td>
<td>0.6175*</td>
</tr>
<tr>
<td>GDPI$_{i}$</td>
<td>0.1932</td>
<td>0.1932</td>
</tr>
<tr>
<td>GDPI$_{j}$</td>
<td>2.2706*</td>
<td>2.2706*</td>
</tr>
<tr>
<td>ILLI$_{i}$</td>
<td>−0.7749*</td>
<td>−0.7749*</td>
</tr>
<tr>
<td>AGRIL$_{i}$</td>
<td>0.1007</td>
<td>0.1007</td>
</tr>
<tr>
<td>AGRIL$_{j}$</td>
<td>−1.9565*</td>
<td>−1.9565*</td>
</tr>
<tr>
<td>POPG$_{i}$</td>
<td>−0.5855*</td>
<td>−0.5855*</td>
</tr>
<tr>
<td>DENSITY$_{i}$</td>
<td>−0.3848*</td>
<td>−0.3848*</td>
</tr>
<tr>
<td>DENSITY$_{j}$</td>
<td>−0.3537*</td>
<td>−0.3537*</td>
</tr>
<tr>
<td>Level 1 variance</td>
<td>24,130*</td>
<td>9,696*</td>
</tr>
<tr>
<td>Level 2 variance</td>
<td>0.0911*</td>
<td>0.1191*</td>
</tr>
<tr>
<td>$R^{2}$ (logged M)</td>
<td>0.533</td>
<td>0.640</td>
</tr>
<tr>
<td>$R^{2}$ (unlogged M)</td>
<td>0.237</td>
<td>0.416</td>
</tr>
</tbody>
</table>

Source: Calculated by the author.
GM, gravity model; EPM, extended Poisson model.
*Significant parameter at 0.05 level.
(a high percentage of rural population engaged in agricultural employment), and high population density were not attractive to migrants in China in 1985–1990. It is clear that interprovincial migration in China was stimulated by economic growth and industrialisation in that period.

Two origin variables POPG, (percentage of population increase between 1982 and 1990 census periods at origin) and DENSITY, (population density in 1990 at origin) also have negative parameters of −0.5855 and −0.3848, respectively. Thus, if everything was equal, then regions with rapid population growth and high population density would send out fewer migrants. This is contradictory to the expected impact and the common understanding that areas with high ‘population pressure’ (high population density and rapid population growth) 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 people living in such regions are 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 the areas with high ‘population pressure’ (such as rapid population growth) may not have the willingness or ability to move to other areas. As shown in Figure 4, Yunnan in the western region had the lowest out-migration rate in the period 1985–1990. Thus, the so-called ‘pushing’ mechanism was not operating effectively in the inter-regional migration in 1985–1990. For interprovincial migration, the origin population seems to be more important in determining the out-migration flow, whereas in-migration is much more selective towards rapidly growing regions. Indeed, the parameter of origin population (0.8289) is much greater than that of destination population parameter (0.6175), indicating that population size is more important in outflows than inflows.

As discussed in the previous section, the scale and spatial patterns of migration changed significantly from 1985–1990 to 1995–2000. It is interesting to find out if the same set of variables used in 1985–1990 can explain the migration in 1995–2000 effectively. The estimated parameters for the migration models of 1995–2000 are shown in Table 2. Regarding changes in the Poisson GM, the parameters for origin and destination populations increase to 1.1352 and 0.7074, respectively, in the model for 1995–2000. The distance decay effect becomes less steep as the absolute distance parameter is slightly reduced. The constant parameter becomes much bigger in absolute value. All model parameters are significant at 0.05 level for both periods. It is clear that the demographic factors such as origin and destination populations become more important in determining the size of migration flow as the distance decay effect is reduced in the model. But overall, the origin and destination populations and the distance become less powerful in explaining the migration in the period 1995–2000. The model for 1995–2000 can only explain 6.7% of the variation of migration, down from 23.7% for 1985–1990. Thus, the Poisson GM cannot explain any significant variation of migration in 1995–2000 because of the increasing scale and complexity of migration in China.

Other social and economic factors played more important roles in migration in 1995–2000. The result shows that the same multilevel EPM works to explain a significant proportion of variations for 1995–2000. The model for 1995–2000 can explain 54.1% of the variation of migration, even greater than 41.6% in 1985–1990.

Nevertheless, there are also significant differences in model parameters between the multilevel EPMs 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, whereas the constant parameter becomes bigger. The change in the distance effect confirms the findings of Fan (2005a) but differs from the result in the study by Li (2004) on earlier changes from 1985–1990 to 1990–1995.

For eight other socio-economic variables, six and five variables are significant at 0.05 level in 1985–1990 and 1995–2000, respectively. ILLI, (percentage of illiterate and semi-illiterate population aged 15+ at destination), AGRI, (percentage of agricultural employment in total rural employment in destination), POPG, (percentage of population increase in 1982–1990 or 1992–2000 at origin), and DENSITY, and DENSITY, (population density at origin and destination, respectively) both are significant and have the same negative impact on migration in the two periods. Their absolute
parameters become bigger to reflect the increasing scale of migration. GDPI_i (annual GNP growth rate in 1981–1989 or annual GDP growth rate in 1995–2000 in origin) and AGRIL_i (percentage of agricultural employment in the total rural employment in origin) are not significant in the two periods.

The variable GDPI_j (annual GNP growth rate in 1981–1989 or annual GDP growth rate in 1995–2000 in destination) has positive impact on migration, but it becomes insignificant in 1995–2000. These differences indicate significant changes in the migration processes, and new variables need to be considered to explain the migration in 1995–2000.

Considering the dramatic changes in the social and economic conditions of various regions from the 1980s to the 1990s, we made an attempt to establish a better model for 1995–2000 by using a new set of variables. The following factors at both origin (i) and destination (j) are considered in addition to the variables used in the model for 1985–1990:

- HHSIZE: Average household size in 2000, persons
- E2S: Share of secondary sector employment in 2000, %
- E3S: Share of tertiary sector employment in 2000, %
- POPUS: Share of urban population in the total population in 2000, %
- GDPA: GDP per unit of area in 2000, Renminbi (RMB) 1000/km²
- EGDP: Export as percentage of GDP in 2000, %
- FDIGDP: Foreign direct investment (FDI) as percentage of GDP in 2000, %
- GDPPC: GDP per capita in 2000, RMB per person
- GDPPCI: Annual growth rate of GDP per capita in 1995–2000, %
- RNONEUD: Share of rural (county) population aged 6+ with education below primary education in 2000, %
- RPSEDU: Share of rural (county) population aged 6+ with primary and secondary education in 2000, %
- UNIV: Share of population aged 6+ with university-level education in 2000, %
- RURALWS: Share of wage in the rural income per capita in 2000, %

Many households in China adopt a split-household strategy by sending some members to engage in labour migration (Fan, 2008). HHSIZE is considered important as a large household may send out more migrants. E2S and E3S measure economic structure. POPUS measures level of urbanisation. GDPA measures economic agglomeration. EGDP measures export intensity. FDIGDP measures inflow of foreign investment. GDPPC measures level of development. GDPPCI measures economic growth rate. RNONEUD, RPSEDU, and UNIV measure level of education. RURALWS and RURALW measure rural waged income. RURALI and URBANI measure rural and urban income per capita, respectively. These variables are important indicators of regional household size, urbanisation, economic development, education level, and income that are expected to affect migration in China (Fan, 1996, 2005a; Shen, 1999; Cai and Wang, 2003).

Stepwise regression of the log-linear equation using ordinary least squares obtained the best model with 17 of 45 variables (32 new variables plus 13 old variables for both origin and destination). Among the variables used in the analysis for the 1985–1990 migration data, POPG_i, DENSITY_j, AGRIL_i, AGRIL_j, GDPI_i, and ILL_j are not included in the best model for 1995–2000 data set. Furthermore, POPG_j and DENSITY_i have a high correlation with other variables in the 1995–2000 data set. Thus, except POP_i, POP_j, and distance, a completely new set of variables needs to be used for the best model of the 1995–2000 data set.

Starting from the full set of 45 variables introduced earlier, 24 variables highly correlated with the other variables are first removed from the data set to avoid the problem of multicollinearity (Table 3). The only exception is that both RURALI_i and URBANI_i, are retained although their correlation coefficient is also high (0.864). They are key indicators of urban and rural income, which play a significant role in regional migration in China (Cai and Wang, 2003). Their parameters in the final model also show expected signs not affected by multicollinearity. Based on...
the stepwise regression of the remaining variables, the best model includes 13 variables (Table 4). A multilevel EPM was then estimated based on these 13 variables. It can explain 65.9% of the variation of migration flows in 1995–2000. This is better than the model using variables used in 1985–1990. Both the level 1 and level 2 variances are significant for the new multilevel EPM for 1995–2000.

Among the 10 socio-economic variables, six new variables become significant factors of interprovincial migration in China. In destinations, RURALI, (rural income per capita in 2000) and URBANI, (urban income per capita in 2000) have a positive impact on migration. GDPPCI, (annual growth rate of GDP per capita in 1995–2000), and RURALWS, (share of wage in the rural income per capita in 2000) have a negative impact on migration. This means that regions with higher urban and rural income per capita would attract more migrants, indicating the importance of income in internal migration in China. But after controlling the income level, those regions with a higher growth rate of GDP per capita and a high share of wage in the rural income per capita would attract fewer migrants. In these regions, many local people may be employed in various industries, producing a high share of wage in their income and higher growth rate of GDP per capita. These regions have less dependence on migrants from other provinces.

In the origins, FDIGDP, (FDI as percentage of GDP in 2000) and RURALWS, (share of wage in the rural income per capita in 2000) have a negative impact on migration. This means that regions with a higher share of wage in the rural income per capita and higher FDI as a percentage of GDP would have less out-migration. This is reasonable as regions with more non-agricultural jobs and higher income for rural people and regions with more foreign direct investment would attract more people to stay rather than move away.

Among the 10 socio-economic variables, four variables become insignificant. As the multilevel EPM attempts to fit well with large migration flows, this means that GDPPCI, (annual growth rate of GDP per capita)

Table 3. Correlation coefficients between variables in the model and variables removed because of high correlation coefficients.

<table>
<thead>
<tr>
<th>Variables</th>
<th>FDIGDP</th>
<th>GDPPCI</th>
<th>RNONEDU</th>
<th>UNIV</th>
<th>RURALWS</th>
<th>RURALI</th>
<th>URBANI</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHSIZE</td>
<td></td>
<td>−0.712**</td>
<td></td>
<td></td>
<td>−0.734**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POPG</td>
<td>0.786**</td>
<td></td>
<td>0.702**</td>
<td>0.703**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DENSITY</td>
<td></td>
<td></td>
<td>0.670**</td>
<td>0.703**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E3S</td>
<td>0.837**</td>
<td></td>
<td>0.822**</td>
<td>0.704**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2S</td>
<td>0.697**</td>
<td>0.700**</td>
<td>0.822**</td>
<td>0.704**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGRIL</td>
<td>0.790**</td>
<td>0.877**</td>
<td>0.822**</td>
<td>0.704**</td>
<td>0.671**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POPUS</td>
<td>0.871**</td>
<td></td>
<td>0.792**</td>
<td>0.712**</td>
<td>0.711**</td>
<td>0.784**</td>
<td></td>
</tr>
<tr>
<td>GDPA</td>
<td>0.776**</td>
<td>0.877**</td>
<td>0.850**</td>
<td>0.711**</td>
<td>0.784**</td>
<td></td>
<td>0.810**</td>
</tr>
<tr>
<td>GDPI</td>
<td></td>
<td></td>
<td>0.792**</td>
<td>0.850**</td>
<td>0.711**</td>
<td>0.784**</td>
<td></td>
</tr>
<tr>
<td>ILL</td>
<td>−0.647**</td>
<td>0.927**</td>
<td>0.931**</td>
<td>0.810**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPSEDU</td>
<td>0.746**</td>
<td>−0.943**</td>
<td></td>
<td></td>
<td>0.938**</td>
<td>0.860**</td>
<td>0.769**</td>
</tr>
<tr>
<td>RURALW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: calculated by the author.

AGRIL, percentage of agricultural employment in the total rural employment; DENSITY, population density; E3S, share of tertiary sector employment; E2S, share of secondary sector employment; E3S, share of tertiary sector employment; EGPD, export as percentage of gross domestic product (GDP); FDIGDP, Foreign direct investment as percentage of GDP; GDPA, GDP per unit of area; GDPI, annual GDP growth rate; GDPPC, GDP per capita; GDPPCI, annual growth rate of GDP per capita; ILL, percentage of illiterate and semi-illiterate population; POPG, percentage of population increase; POPUS, share of urban population in the total population; RNONEDU, share of rural (county) population aged 6+ with education below primary education; RPSEDU, share of rural (county) population aged 6+ with primary and secondary education; RURALI, rural income per capita; RURALWS, share of wage in the rural income per capita; RURALI, rural income per capita; RURALWS, share of wage in the rural income per capita; UNIV, share of population aged 6+ with university-level education; URBANI, urban income per capita.

**Correlation is significant at the 0.01 level (two-tailed).
Government policy, rural reform, household registration system reforms, regional development, development of non-state-owned sector, and foreign investment have been identified as major factors of increasing migration (Li and Li, 1995; Liang and White, 1997; Wei, 1997; Cai and Wang, 2003; Li, 2004; Fan, 2005a; Li et al., 2010; Luo, 2010).

This paper is concerned with the temporal dimension of migration. Can the same model, using the same set of variables, explain the migration patterns in the two different periods, given the increased scale of migration? Does increased migration size just mean proportional expansion of migration in all directions? Is there any change in the determinants of interprovincial migration in China? This paper made an attempt to answer these questions using migration data from the 1990 and 2000 censuses of China.

The changes in the spatial patterns of interprovincial migration in China were examined first based on the regional in-migration rates and out-migration rates in the periods 1985–1990 and 1995–2000. It was found that Guangdong, Zhejiang, Tianjin, and Fujian joined Shanghai and Beijing as the top destinations of migration over time. The concentration of the top destinations in the coastal region is the most outstanding change in the migration pattern. The ‘pull’ factors in the growing regions play an important role in the expanding migration in China. Some provinces with high in-migration rate in 1985–1990 such as Hainan and Qinghai were overtaken by the aforementioned top destinations. Only Xinjiang in the western region kept a high in-migration rate in both periods. Only some provinces in the central and western regions, including Jiangxi, Hunan, Anhui, Hubei, Sichuan, Guangxi, and Guizhou, sent many migrants to the eastern coastal region in 1995–2000. Some provinces have not established a strong migration link with the coastal region. The ‘push’ factors have not been as strong as the ‘pull’ factors. These are further confirmed by the result of modelling the interprovincial migration flows using the Poisson GM and the multilevel EPM.

In the Poisson GM for 1985–1990, all three variables are highly significant, but the model only explains 23.7% variation of the number of migrants for all flows. In the multilevel EPM of 1985–1990, 11 variables were used. They

### DISCUSSION AND CONCLUSION

With dramatic changes in the patterns of regional development, showing significant growth in coastal regions, the scale of interprovincial migration has increased significantly in China.
described important demographic, socio-economic situations in various areas and were selected by stepwise regression. The model can explain over 41.6% of the variation of the number of migrants in all flows in the period 1985–1990. The result indicates that, if everything were equal, regions with poor education (a high percentage of illiterate and semi-illiterate population), less rural industrialisation (a high percentage of rural population engaged in agricultural employment), and high population density were not attractive to migrants in China in 1985–1990. It is clear that interprovincial migration in China was stimulated by economic growth and the industrialisation process in that period.

The same set of variables used in 1985–1990 was used to explain the migration in 1995–2000. It is found that the Poisson GM cannot explain any significant variation of migration in 1995–2000 because of the increasing scale and complexity of migration in China. It can only explain 6.7% of the variation of migration for 1995–2000.

But the same multilevel EPM works to explain a significant proportion of variation for 1995–2000. The model for 1995–2000 can explain 54.1% of the variation of migration, even greater than 41.6% in 1985–1990. The findings have important implications in migration projection (Wilson and Rees, 2005). It seems possible to use a set of known explanatory variables to estimate migration flows in the future period, even in cases when the migration scale is increased dramatically. Nevertheless, there are also significant differences in model parameters between the multilevel EPMs for 1985–1990 and 1995–2000. Five variables are significant and have the same negative impact on migration in the two periods, but their absolute parameters become bigger to reflect the increasing scale of migration. Thus, it is possible to adjust the model parameters to reflect the changing trend of migration scale to project inter-regional migration flows. Such migration estimates may then be used in multi-regional population projections (Shen, 1994a, b).

Finally, a better model for 1995–2000 is established by using a new set of variables. The new model can explain 65.9% of the variation of the migration flows in 1995–2000. Based on the initial stepwise regression of the 1995–2000 data set, no social and economic variables used in the model for 1985–1990 are retained in the new model.

In addition to origin population, destination population, and distance variables, six new variables become significant factors of interprovincial migration in China. Most interestingly, two income variables, RURALI and URBANI, have a positive impact on migration. This means that regions with higher urban and rural income per capita would attract more migrants, confirming the importance of income in internal migration in China (Cai and Wang, 2003). Furthermore, FDIGDP, and RURALWS, have a negative impact on migration. This means that regions with a higher share of wage in the rural income per capita and higher FDI as percentage of GDP would have less outmigration.

Clearly, compared with 1985–1990, income, foreign investment, and economic opportunities had become the most important determinants of interprovincial migration in China in the period 1995–2000. A few important theoretical and practical implications can be drawn based on the modelling results. First, labour migration from poor rural areas in the Central and Western regions to more developed urban areas in the Eastern region dominated the ongoing migration in China.

Second, a national labour market was emerging as the income level was driving the migration flows in 1995–2000. Although a GM based on population size and distance could explain a significant proportion of the variation of migration flows (23.7%) in the period 1985–1990, it had little explanatory power on the migration in the period 1995–2000, with much higher mobility under increasing impact of economic factors. Seemingly, China underwent a transition from a demographic–economic-driven migration to an economic-driven migration in the period 1985–2000.

Third, it is the high income level rather than the high growth rate of the income that attracts migration. In the case of China in 1995–2000, after controlling the income level, those regions with a higher growth rate of GDP per capita would attract fewer migrants. Some poor regions in China achieved a high economic growth rate, but they were not very attractive to migrants because of low income levels.

Fourth, foreign investment had an important impact on migration in China. As a result of economic opportunities created by such investment, regions receiving more investment would send out fewer migrants.
Fifth rural industrialisation had a special impact on migration. Regions with high rural industrialisation had less dependence on migrants from other provinces. Many local people may be employed in various rural industries, producing a high share of wage in the rural income. Such regions would attract fewer migrants from outside but would also send out fewer migrants. Thus, economic growth led by exogenous factors (such as FDI) and endogenous factors (such as rural industrialisation) has different consequences for migration. FDI-driven export-oriented economies would stimulate large-scale migration, whereas endogenous rural industrialisation would not increase migration greatly.

It is clear that rapid economic growth, unbalanced regional development, emerging market economy, increasing economic openness, relaxed migration control, and improved communication and transportation systems are creating a new economy–migration nexus in China (Shen, 2006; Fan, 2008; Zhu and Chen, 2010). This nexus has likely been enhanced since 2000 with further social-economic changes in China. The new migration data to be released from the 2010 population census would be useful to verify such expectations. Nevertheless, the new emphasis on rural development, balanced regional development, transition from an export-oriented economy to a domestic consumption-driven economy, declining fertility, and labour shortage in the coastal region would have different impacts on origins and destinations of migration. The scale of migration may be further increased but perhaps not as rapidly as in the period 1985–2000.

ACKNOWLEDGEMENTS

The research on which this paper is based was supported by the Hong Kong Institute of Asia-Pacific Studies, The Chinese University of Hong Kong and a research grant from Research Grants Council of Hong Kong Special Administrative Region (RGC project no. 450107). Thanks are due to two anonymous referees for helpful comments.

REFERENCES


