Evolution Characteristics of Spatial Concentration Patterns of Interprovincial Population Migration in China from 1985 to 2015

Ziliang Liu[1] · Hengyu Gu[2]

Abstract

People migrate between places constantly, promoting population redistribution and socioeconomic element flows. Since China's interprovincial population migration flows are highly concentrated in particular places and there are comparatively few studies using a spatial concentration perspective in this field, there is room for a spatial analysis. The Gini Index (GI) is commonly used in analysing the concentration of topological data, and it works well with migration flow data. Also, GI is scale-invariant due to its normalized nature, and hence it is readily interpretable. However, due to missing data, the traditional GI specification is often an underestimate. Thus, the authors propose the GIc for better standardisation, and demonstrate its utility in the characterization of interprovincial migration flows in China from 1985 to 2015, quantifying the spatial concentration of migration. The results indicate that (a) the GIc measures the concentration degree of population migration better; (b) collectively, the concentration is consistently high, and it shows a trend in an inverse U-shape over six periods, and there are strong linkages between regions with high origin GIcs and regions with high destination GIcs; (c) since 1995, the concentration at migration origins is high in middle China and some western regions; migration destinations with high concentration are the Yangtze River Delta and the Pearl River Delta, which show a smooth temporal change; (d) origins with high concentration are often connected to destinations with high concentration by strong migration flows that move large populations.

Keywords Population migration · Spatial concentration · Coefficient correction · Gini index · Evolution characteristics · China

✉ Hengyu Gu
  henry.gu@pku.edu.cn

  Ziliang Liu
  20160781028@m.scnu.edu.cn

Extended author information available on the last page of the article

🙂 Springer

Introduction

Population migration is known to influence the political order (Ammendola 2005), economic development and cultural ecology among regions (Smith et al. 2016), playing an increasingly important role in globalization. Current society is in an "Age of Migration" (Castles et al. 2013), hallmarked by an enhanced element flow, and regional imbalance caused by population migration, which also causes conflicts and poses challenges to the stability and development of demographic regions (Castles and Miller 2009). On a macroscopic level, both intercontinental and domestic population migration promote resource redistribution and enhance regional construction. However, microscopically, mobility offered by migration is crucial for individuals to achieve personal goals (Bernard et al. 2014). According to United Nations documents, with proper policy support, migration can promote sustainable economic growth, enhance allocation of labour sources and lower old-age dependency ratios (United Nations et al. 2017). Nevertheless, population migration inevitably intensifies development imbalances among regions, and it can lead to various problems, such as ecological deterioration, social instability and a larger gap between the rich and the poor (Adey 2010).

China is a typical and ideal study area for analysing the spatial distribution of population migration and its spatial concentration patterns. In China, interprovincial migration is a long-standing socioeconomic phenomenon that fuels the spatial redistribution of population (Liu and Shen 2016; Wu 1994). Data from previous population censuses and population sample surveys suggest that, since the relaxation of migration restrictions in the 1980s, population migration has been constantly increasing, making it an even more important macroscopic index (Li et al. 2015). Meanwhile, migration flows have changed over the years, with an increasing spatial concentration. The migrant population tends to migrate to destinations with more employment opportunities, higher income, better social security systems and more developed transportation networks (Fan 2005a). Collectively, as research into multiple periods has suggested, provinces in middle and western China such as Sichuan, Hunan, Hubei and Anhui are the major origins of population migration, and those flows go towards more developed eastern coastal areas such as Beijing, Shanghai and Guangdong, which hold a great number of immigrants (Ding et al. 2005; Wang and Pan 2016; Yang et al. 2015; Zang et al. 2016; Zeng et al. 2013). Most of the migration flows are concentrated between these major origins and destinations, which can be quantitatively addressed (Bell et al. 2002). Insights on the spatial concentration pattern of population migration in China can not only influence policymaking in China, but also provide enlightenment and social implications to other countries or regions, as there are similarities among countries. In addition, it can benefit cross-national comparative research in the field of population migration.

The pattern of spatial concentration (also known as spatial focusing) of population migration flows has been a long-studied subject, since it contains information about the non-uniform spatial distribution of migration (White 1986). Building a well-off society in an all-round way by 2020 is a top priority, in which population migration will play an important role. Population migration produces element flows of human resources and social capital (Liu and Xu 2017), and it impacts economic coordination, people's lives and cultural building profoundly. Recently, in light of the increasingly imbalanced population migration and hence population distribution, scholars have thoroughly discussed how spatial redistribution of population relates to socioeconomic issues,

including: (a) How does population migration affect urbanization in different regions (Yang and Ning 2015)? (b) Is migration reducing or creating regional imbalance (Duan and Liu 2012)? (c) Should the rural population be urbanized locally or directed to large cities to exploit economies of scale (Lu et al. 2012, 2014)? Nevertheless, such discussions cannot be made quantitative unless migration flows can be correctly inferred from collected data, which are often incomplete. Therefore, a reliable inference framework should further promote such discussion as well as providing references for policymaking and execution.

To measure the spatial concentration of population migration, there are five popular methods: Coefficient of Variation (CV), Index of Migration Dissimilarity (IMD), Gini Index (GI), Theil's Entropy Index (TEI) (Theil 1972) and Atkinson's Index (Allison 1978; Atkinson 1970; Bell et al. 2002; Duncan and Duncan 1955; Griffith and Amrhein 1991; Rogers and Raymer 1998; White 1986). Plane and Mulligan (1997) commented that GI is the natural choice for measuring regional imbalance (including spatial concentration of population migration), as the other indexes are often artificial (even random), hard to interpret, and/or less suitable. In contrast, GI has the following advantages (Duncan and Duncan 1955; Plane and Mulligan 1997): (a) all migration flows are standardized to the same scale, so that comparison between flows is meaningful with a clear interpretation; (b) GI is better at capturing the spatial variations of the studied distribution; (c) GI can be computed for regions of different sizes; and (d) most migration data today are in the form of migration matrices, from which GI can be readily computed. Hence, the authors used GI as the major source of statistics. More details of the comparisons are in Table 1.

GI was first proposed by Gini (1926), an Italian economist, in the 1920s for measuring income inequality. The index system then underwent further adaptation and generalization, and it found its application in multiple subjects at different scales. On one hand, GI has been adapted to different scenarios in the study of income distribution, including improvement of computational efficiency (Lerman and Yitzhaki 1984; Yitzhaki 1983), extended application of grouped data (Lyon et al. 2016) and decomposition of effects (Lau et al. 2017). On the other hand, GI has been applied into industrial agglomeration and population migration. Krugman (1991) proposed using location GI to study the concentration of industrial space, classifying it into absolute GI and relative GI. White (1986) introduced a GI system to demographic research, aiming to measure the concentrations of population migration flows. More recently, Janská et al. (2014) applied GI to track how migration changes and migrants settled over time in Czechia. To sum up, since its application in measuring income inequality, GI has been constantly adapted and increasingly applied in different fields, whether a quantity of interest is concentrated or spread over the studied spatial coordinates.

In terms of drawbacks, GIs is sensitive to missing data in censuses or surveys, which leads to deviations in the inferred spatial-temporal trends. Therefore, in this study, the authors specifically address the accurate inference of the spatial concentration of population migration from incomplete data by proposing a GI system that is corrected at origin, at destination and from a global perspective. This study entails: (a) taking missing data into consideration in population migration and proposing the GIc system, and (b) applying the GIc to analyse spatial-temporal features of population migration in multiple periods. The most significant contribution of this paper is that we re-standardize the GI in topological structure, through which we recognize that the concentrations of population migration may be underestimated. The authors hope that

Table 1 Comparisons of different indexes for spatial concentration of population migration

Index	Advantage(s)	Disadvantage(s)
CV	Able to satisfy the transferring criterion and scale criterion	If migration increases entirely, but much faster at the higher end, the CV will change sharply to show different rankings and trends.
IMD	Less impacted by the type of data (event or transition) or the length of time	Unable to describe high spatial concentrations properly
GI	Provides multivariate aspects, as GI can be calculated in different perspectives; suitable for comparison of variables with different scales and time; easy for comprehension; also satisfies the scale criterion and the transferring criterion	Complicated for calculation
TEI	Able to compare variables with different numbers of groups; satisfies decomposition criterion	Cannot satisfy the scale criterion
Atkinson's Index	Able to compare variables with different numbers of groups	A parameter needs to be set, which may be artificial or random (because it needs extensive experience), and hence the calculation is not robust, or it may cause bias.

Summarised from White (1986), Plane and Mulligan (1997) and Bell et al. (2002); scale criterion refers to the invariance of proportion when the numbers of migrants increase; and transferring criterion refers to the increase in SC if migrants transfer to another destination with a higher spatial concentration

this proposed GI correction will broaden and deepen the study of the spatial concentration of population migration and population migration in general.

This paper is organized as follows: after describing the methodology and data in the next section, especially comparison of GIs and introduction to Chord Diagram. The following section details the main results of this research; and finally, the authors offer conclusions and discuss the results.

Research Methodology and Data Source

Research Methodology

Raw GI and its Correction

GI is a well-known and widely applied statistics for measuring the spatial concentration of population migration, offering direct comparisons between flows (M_{ij} and M_{kl}, see below). After corrections and extensions, Bell et al. (2002) proposed a standardized GI system that can be used to compare population migration in different regions, and which includes three components: global GI, origin GI and destination GI. These indexes have been widely applied since then (Bell and Muhidin 2009; Bell et al. 2015). The formal definitions of raw GIs read:

Global GI:

$$GG = \frac{\sum_i \sum_{j \neq i} \sum_k \sum_{l \neq k} |M_{ij} - M_{kl}|}{2[n(n-1)-1]\sum_i \sum_{j \neq i} M_{ij}} \tag{1}$$

Origin GI:

$$OG_i = \frac{\sum_{j \neq i} \sum_{l \neq i,j} |M_{ij} - M_{il}|}{2(n-2)\sum_{j \neq i} M_{ij}} \tag{2}$$

Destination GI:

$$DG_j = \frac{\sum_{i \neq j} \sum_{k \neq j,i} |M_{ij} - M_{kj}|}{2(n-2)\sum_{i \neq j} M_{ij}} \tag{3}$$

Where n denotes the number of regions, i, j denote different origins and k, l denote different destinations. $M_{ij(kl)}$ denotes migration flows from i to j (or k to l). In addition, GG, OG_i, $DG_j \in [0,1]$. A value of 1 indicates a complete concentration in which all the migration concentrates in a single flow. If the index is 0, it implies a complete balance in which the migration population is evenly distributed among the flows.

However, these defined GIs ignore missing data (*null*). In this situation, the numerator and denominator, especially the counts of flows, are affected. Therefore, GI cannot reach 1 even in the case of complete concentration, violating the original interpretation of the formula. Under this circumstance, the authors try to standardize GI better, focusing particularly on the denominator. Accordingly, the authors propose GIc as below:

Global GIc:

$$GG^* = \frac{\sum_i \sum_{j \neq i} \sum_k \sum_{l \neq k} |M_{ij} - M_{kl}|}{2[n(n-1)-1-p]\sum_i \sum_{j \neq i} M_{ij}} \tag{4}$$

Origin GIc:

$$OG_i^* = \frac{\sum_{j \neq i} \sum_{l \neq i,j} |M_{ij} - M_{il}|}{2(n-2-p_i)\sum_{j \neq i} M_{ij}} \tag{5}$$

Destination GIc:

$$DG_j^* = \frac{\sum_{i \neq j} \sum_{k \neq j,i} |M_{ij} - M_{kj}|}{2\left(n-2-p_j\right)\sum_{i \neq j} M_{ij}} \tag{6}$$

Here n, i, j, k, l and $M_{ij(kl)}$ are the same as they are in Eqs. 1–3. p denotes the number of flows missing from the migration data, p_i denotes the number of migration flows missing from origin i and p_j is the count of migration flows missing from destination j.

To illustrate the effect of this correction, consider a randomly generated 5×5 migration data matrix M (see Table 2).

In this matrix, the rows and columns correspond to five demographic regions A-E, between which there are 20 pairwise migration flows (excluding the internal migration, expressed by diagonal elements). A randomly chosen population migration flow is set to population m (as A_{41} in the matrix) and five randomly chosen migration flows are set to missing (*null*), whereas all other off-diagonal elements are set to 0. According to the definition of complete concentration (Bell et al. 2002, 2015), when migrant populations concentrate into one flow, the index is 1. In this situation, overall GI = emigration GI of region D = immigration GI of region A = 1.

After calculations, $OG_D = 0.667$, $DG_A = 0.667$, $GG = 0.737$, and $OG_D^* = DG_A^* = GG^* = 1$. Note that GI is supposed to equal 1, since all the migration flow is carried by population migration from region D to region A, which is not the case in the raw GI due to the missing value here. In contrast, the GIc is indeed equal to 1, and it represents the concentration of the migration flow well, as it is more standardized.

Chord Diagram

Since population migration flows are essentially directed edges between population nodes in mathematical terms, a Chord Diagram (CD) is particularly useful in its visualization due to its topological structure (Gu et al. 2014). The main elements of a CD are nodes and arc lengths corresponding to nodes and chords among nodes, which can reveal the origin, destination, flow direction and tie intensity (volume). In the diagram, arc length indicates the total tie intensity of nodes; migration direction is presented based on the colouring of origins, and arrows make this clearer. Nodes are connected by chords; chord width indicates tie intensity among nodes; width and tie intensity are directly proportional (Abel and Sander 2014). Moreover, with colouring, a CD is clearer, brighter and more distinct than a traditional diagram.

The first application of a CD was in visualizing how expressions of genes on the topologically continuous genome are correlated (Zahn et al. 2017). Thereafter, CDs were extensively applied to visualize weighted connections between entities with topological constraints, including highway passenger flows (Chen et al. 2017) and population migrations (Qi et al. 2017), with great success. CDs are complementary to traditional spatial heatmaps, and a combination of both is particularly powerful in revealing patterns hidden in the interactions between the nodes of study.

Table 2 A 5×5 random migration data matrix

	A	B	C	D	E
A		0	0	null	0
B	0		0	0	0
C	0	null		null	0
D	m	0	0		null
E	null	0	0	0	

Data Sources

To see the proposed GI system in action, the authors applied it to analyse the concentration of interprovincial population migration in China from 1985 to 2015.

The interprovincial migration data in China came from (a) the 4th to 6th population censuses from 1985 to 2015 and (b) nationwide 1% population sample surveys in 1995, 2005 and 2015. The authors aggregated migration data in five-year periods, to capture the long-term features of migration, as opposed to the short-term features of population flow. The data cover 31 provinces, municipalities and autonomous regions in China (excluding Hong Kong, Macao and Taiwan). Except for the data from the 5th and 6th population censuses, the data of other years are incomplete to different extents: there are 30, 44, 15 and 4 missing migration flows in the 4th population censuses and 1% population sample surveys in 1995, 2005 and 2015, respectively. Most of the missing values are related to Tibet, while some other missing flows are related to Ningxia and Qinghai. In addition, considering the conversion ratio of data and the non-negative integer nature of the migrant population, the migration data are restored by years based on the sampling ratio and rounded off. Of note, the census data is referred to by the period during which it is collected e.g., "data from 2005-2010" refers to the 6th population census, when the migrant population was recorded.

Results

With MATLAB, the authors calculated the raw GIs and GIcs, and they compared the results. The calculated values reflect the concentration of interprovincial population migration from the global perspective, the origin perspective and the destination perspective, collectively depicting a spatially imbalanced migration.

Comparison of Raw GI and GIc

According to the results (Fig. 1), the raw GI underestimates migration concentration, and the GIc better reveals the spatial imbalance of population migration. The raw GI has a smaller value when migration data are incomplete for a particular period, hence underestimating the migration concentration. Whereas GI deviation is positively correlated with the number of population migration flows missing, global GI and local coefficient are impacted differently. From the global perspective, the missing 4–44 migration flows (a missing rate of around 0.4%–5%) gives a GI absolute deviation of 0.004–0.051. From the local perspective, based on origins and destinations, the authors calculated the missing rate ranking stratified by the period, which also suggests a positive correlation with the corresponding GI deviation. Hence, the authors conclude that raw GI suffers from underestimation when there is missing data, regardless of the perspective of GI, be it global or local.

Spatial Concentration of Population Migration Based on the Global Perspective

Figure 1 shows that a reduction occurred after the increase of the interprovincial migration concentration in China in 2005. In other words, it shows an inverse U-
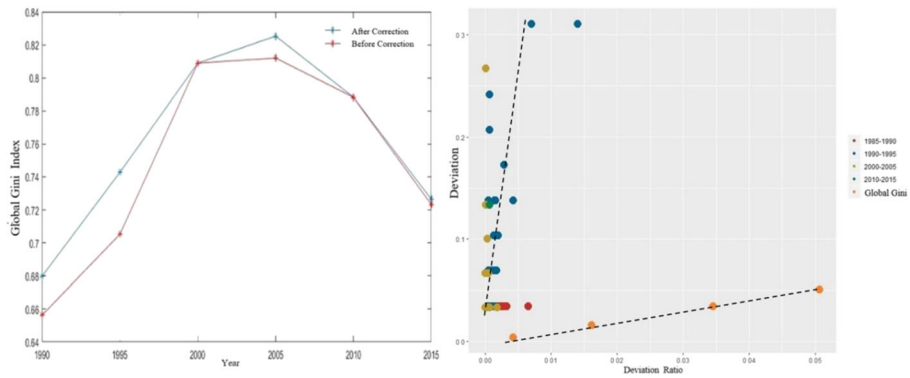
Fig. 1 Effect of GI correction

shaped trend. The rate at which the spatial concentration of population migration increases gradually reduced in the concentration phase, whereas the rate of reduction slightly increased during the decline stage.

Based on our analysis, Stage I refers to the developing stage of reform and opening up, when Guangdong and the Yangtze River Delta (especially Guangdong), the door of the economic opening up, attracted a great number of migrants looking for better jobs and higher incomes. Meanwhile, there was also population migration from well-populated regions with less developed economies to more developed regions. In Stage II (after 2005), some origins (except some traditional provinces of emigration) of migration flow received a backflow from the developed regions, as previous migrants moved back to their hometowns or nearby. From the destination perspective, some undeveloped regions underwent improvements that promoted investment and infra-structure as a result of policy inclination, and they reached a higher level of urbaniza-tion. Therefore, many regions other than the traditional destinations like the Pearl River Delta, the Yangtze River Delta and the Beijing-Tianjin region are increasingly attractive to the migration population, resulting in a diversification of population migration destinations. As a result of changes at both origins and destinations, population migration is becoming less concentrated.

Strong Linkages between Origins and Destinations with High Spatial Concentration

Origins and destinations are correlated, and they need to be viewed collectively. Here, the authors point out two definitions: (a) if an origin is highly concentrated, the emigrants from such an origin will often choose certain destinations, and the population flow will become larger; (b) similarly, for a destination with high concentration, immigrants often share certain origins.

CDs were plotted to visualize these observations topologically. For each period, the top three origins and destinations in terms of the originated or received population migration flows were highlighted, and the three strongest population migrations were plotted for each of the highlighted region (Fig. 2). In the figure, black arrows indicate migration direction; purple areas indicate the top three origins of this period; green areas indicate the top three destinations; and grey areas indicate other regions that are strongly related to the top three origin or destination regions.
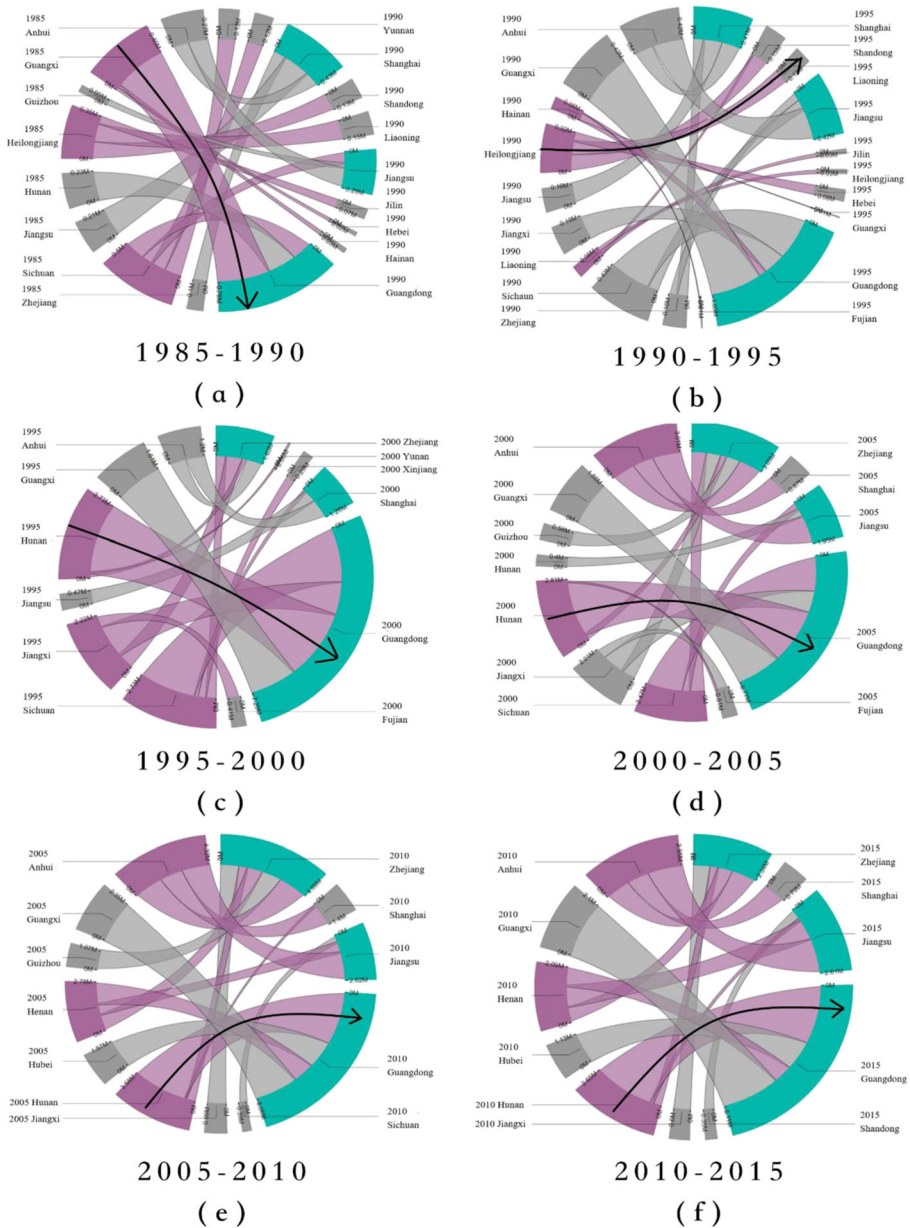
Fig. 2 Correspondent migration linkages between origins and destinations with high concentration (1985–2015)

The results suggest that the origins and destinations of the highest GIs are often strongly related by population migration flows. These strong linkages lead to imbalanced spatial distributions of population migration in China. Temporally speaking, such population migration flows weakened from 1985 to 1995 but they strengthened afterwards, and they contributed a large fraction of the overall migration flow of the origin or destination.

Most of the origins in 1990 were traditional emigration provinces, except Hainan Province, and the top four destinations, Guangdong, Jiangsu, Zhejiang and Shanghai are constantly receiving large immigrant populations, indicating an active ongoing population migration. In the regions with high destination GIcs, large numbers of immigrants are coming from provinces with high origin GIcs. On zoning, origins with high population migration concentrations are often located in the middle and west of China; destinations with high population migration concentrations are often in the east of China and near the coast. Of note, the observed population migration is also reminiscent of the long-term flow from the middle or west to the east of China since the 1980s (Wang and Pan 2016), reflecting the highly spatially directed nature of the migration flow.

Spatial Concentration Pattern of out-Migration

Generally, from the perspective of origins (namely, based on the origin GIc), the spatial structure of migration tends to be stable: emigration from the middle and some regions of the west of China is concentrated.

A high emigration concentration means that the emigrants from a region often move to certain areas, whereas a low concentration means the emigrants are scattered extensively across different destinations, which is calculated according to Eq. 5. Based on the overall spatial distribution, a pattern of high migration in middle China and some western regions has gradually taken shape since 1995.

Temporally speaking, the emigration flows from the northeast were mostly concentrated in nearby provinces. Later, as socio-economic development slowed down, the migrating population was attracted to the eastern coastal regions with better job opportunities and transportation, reducing the population migration concentration.

Population migration originating from the middle regions is highly concentrated, showing a steady increase due to the dependency of population migration on its destination. Emigrants from the middle often move to eastern coastal regions. For a given origin, the emigrant population will concentrate around a certain destination (Pu and Ge 2016). For instance, from 2005 to 2010, 3.7 million emigrants moved from Hunan to Guangdong, Fujian, Shanghai and Zhejiang, accounting for 80.76% of the total population migration from Hunan. Note that these four destinations are in close demographic contact. Generally, emigrants from middle China tend to concentrate in eastern coastal regions.

As for population migration from the west, two different scenarios are observed. The southern provinces including Sichuan, Chongqing, Guangxi and Guizhou are in contact with the aforementioned middle provinces, and they exhibit a similarly large and highly concentrated emigrant population. In ethnic minority regions like Tibet, Xinjiang, Gansu and Ningxia, however, the emigration concentration has been low for a long time, with small fluctuations occurring in some of the regions.

Spatial Concentration Pattern of in-Migration

A destination with a high population migration concentration often receives immigrants from certain origins, and vice versa. The degree of concentration is calculated using Eq. 6. Generally, immigration is highly concentrated in the eastern coastal regions, with a decreasing concentration in Xinjiang.

Unlike other areas, the Pearl River Delta and the Yangtze River Delta have experienced concentrated immigration in all the observed periods. Specifically, Guangdong received the most concentrated immigration in all periods, and two out of Zhejiang, Jiangsu and Shanghai came second and third depending on the period. In these regions, a large immigrant population forms the basis of a highly concentrated population, which is shared between several origin provinces. These regions are typically: (a) highly developed, with more job opportunities and higher salaries; (b) coastal, hence offering more openness and freedom, as well as being more receptive and fairer to immigrants; (c) nationwide traffic hubs with convenient traffic and developed systems, hence they are within easy reach of immigrants.

On the other hand, immigrant population varies according to origins. For a specific destination, the immigrant population usually originates in a certain region and its neighbourhood; hence, it is spatially concentrated (Chun 2008; Chun and Griffith 2011). From 1985 to 2015, these origins were usually Hunan, Hubei, Anhui, Jiangxi, Sichuan and Guizhou, most of which are neighbouring provinces, showing the route dependency of origins. More precisely, from 2005 to 2010, the immigrant population from Henan, Hunan and Hubei accounted for almost half of the total immigrants to Guangdong (49.34% of 6.84 million), compellingly suggesting that a strong migration flow tends to drive the population in its surroundings to move to the same destination (e.g., the population moved from Hunan and Hubei to Guangdong from 2010 to 2015).

Notably, immigration concentration in Xinjiang reduced, owing to the policy support it received and the land resources available there attracting a great number of immigrants, mainly from Henan, Gansu and Sichuan (Fan 2005b; Shen 2013). After that, the number of immigrants kept increasing, and the source provinces extended to include Shandong and Jiangsu. This diversification of origins translated to a reduced population migration concentration. Figure 3 maps the spacial concentration of on the origin GIc basis, and Fig. 4 maps it from the perspective of destination GIc.
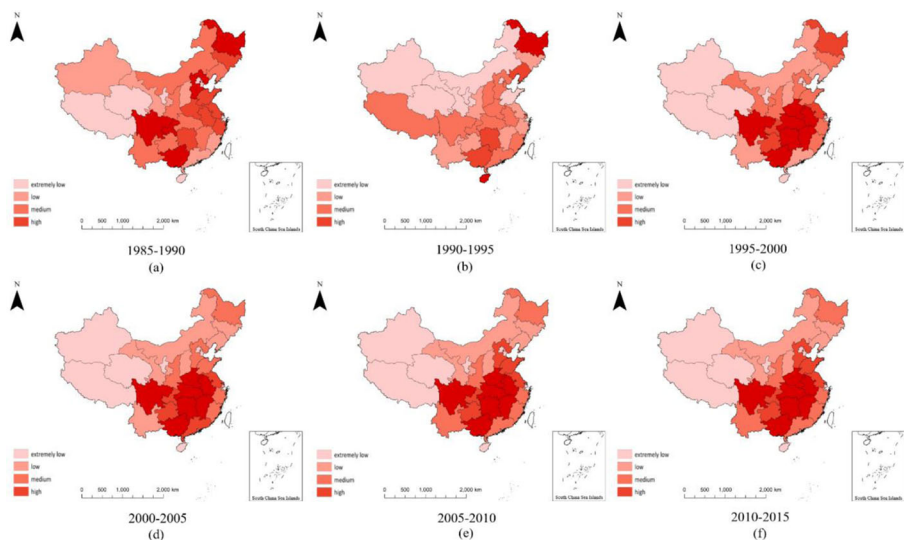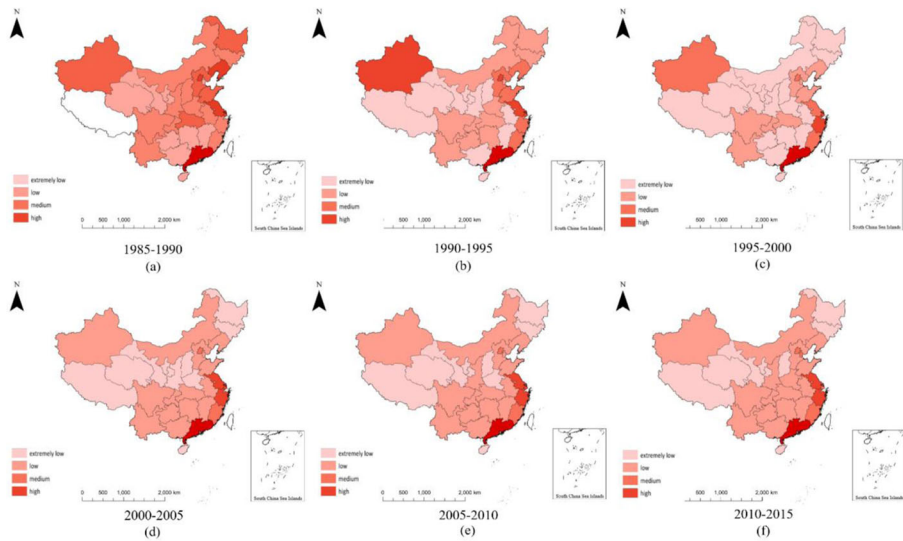

Fig. 3 Map of origin GIc (1985–2015)

Fig. 4 Map of destination GIc (1985–2015)

In China, the spatial pattern of interprovincial population migration depends on topography and climate, and this pattern results in regional living conditions and differences in economic development. These relatively stable macro factors also contribute to the relative stability of population distribution and the pattern of internal population migration (Wang et al. 2012). Since 1978, China has taken action to spur economic vitality, including reform and opening policies, a western development policy and a strategy for the rise of central China. These policies, however, have not changed the basic spatial pattern of regional development and economic growth; on the contrary, they have even aggravated regional imbalance, along with interprovincial migration. For instance, some regions like the Yangtze River Delta and the Pearl River Delta have stepped into a new era with high economic growth since 1978, having attracted many immigrants from less developed areas. Therefore, eastern coastal regions have had high spatial concentrations of immigration for a long period. From the beginning of the reform and opening-up policy to the late 1990s, interregional migration in China was mainly influenced by economic development, job opportunities and policy. However, in the twenty-first century, apart from the above factors, population migration is being more and more impacted by social networks and living conditions (Gu et al. 2019; Liang 2016; Liu and Xu 2017), and hence migration decision making is a comprehensive behaviour with multivariate factors. Moreover, strategies of regional development and the new urbanization planning (from 2014 to 2020) are promoting a dwindling trend of spatial concentration of population migration in China, with more and more migrants migrating to different regions, which are more attractive than before.

Conclusion and Discussion

Since population migration is an important social phenomenon in China, whose spatial pattern impacts multiple socioeconomic elements, it needs to be carefully studied.

Based on the GIc system, the authors measured the spatial concentration of interprovincial population migration in China from 1985 to 2015 to uncover its imbalanced nature. The research reveals that:

First, the GIc system measures the spatial concentration of interprovincial population migration with less bias, whereas the traditional raw GI is biased, and it underestimates the spatial concentration of population migration. Specifically, this bias is positively correlated with the quantity of missing data, and it is more pronounced for global GI than for local GI. Use of the GIc should fuel more accurate analyses of spatially imbalanced population migration, and it should provide a better reference for policy making.

Second, from the perspective of global GIc, migration shows an inverse U-shaped trend, indicating that there is a dilution of migration after a concentration. However, global GIc exceeds 0.64 for all the six periods, showing that interprovincial migration in China from 1985 to 2015 was highly imbalanced, although there was a decline in GIc after 2005.

Third, origins and destinations with highly concentrated population migration are often connected by a strong migration flow that carries a large population. Given that the population migration flows studied involve a quite large population, a migration tie is more likely to be established between the top origins and destinations ranked by spatial concentration. These observations also indirectly confirm the long-term tendency of migrating from middle and west to the eastern developed areas.

Fourth, generally speaking, the emigration population concentration is high in middle China and some western regions (Chongqing, Sichuan and Guangxi). Moreover, the middle regions are associated with a constantly highly concentrated emigration density. The size of the emigrant population also evidently depends on the destination, and most migration is directed to the east and coastal regions, which are contiguous.

Fifth, destinations with highly concentrated immigration are in the Yangtze River Delta and the Pearl River Delta. The top three destinations are always Guangdong and two out of Zhejiang, Jiangsu and Shanghai. Such destinations have high concentrations of migrants and high immigrant populations, which means not only that the origins are shared, but also that relevant migration flows concentrate large populations. What is more, high concentration is also impacted by destination-related specifications. Based on a specific destination, the immigrants are sourced from a certain origin and its neighbourhood. Notably, with policy support and abundant resources, Xinjiang has gradually become a new source of growth. However, as the affected areas have extended, relevant emigration GI has gradually dropped.

One of the main findings of this paper is that the global GIc of China's population migration shows an inverse U-shaped trend. After 1985, with the implementation of reform and the opening up policy, household registration in China was liberalized to some extent, and thus the binding of land to the labour force gradually declined, leading to the beginning of large-scale population migration. Due to the imbalanced regional economic development and the uneven supply of public services, the spatial concentration of population migration increased rapidly. The imbalanced supply of labour led to a further increase in the imbalance of regional economic and social development, thus forming a cyclic, cumulative causality effect. After entering the twenty-first century, the government realized this problem, and it began to implement a strategy

of equalization of public services, encouraging labour to settle down into small and medium-sized cities, thus forming an orderly urbanization. Meanwhile, some strategies for balanced regional development were proposed and implemented, such as the development of the western regions, the rise of the central regions and the revitalization of the northeast regions. For provinces and regions that attract large numbers of migrants, a seller's market for labour has emerged, and a decline in wages and an increase in the cost of survival have also blocked further increases in immigration. The above reasons have led to a decline in the spatial concentration of population migration in recent years.

However, from the overall perspective, the GIc of China's inter-provincial population migration is still higher than 0.6, indicating that the population migration pattern is still highly imbalanced. According to our results, governors of destination provinces need to provide better urban space as well as improved environmental capacity for the highly concentrated immigration population. For instance, governments can supply high-quality and diversified public services, and they can enhance the urban built environment to provide citizens with comfortable urban lives. Also, owing to the significance of transportation in urban spaces, governments should develop public transportation networks to improve commuting convenience.

On the other hand, the central government needs to adhere further to the new-type urbanization strategy, to move further in relaxing the restrictions on the household registration system, and to encourage the floating population to settle down in small and medium-sized cities near their origins. In addition, economic resources need to be allocated properly to promote the rational distribution of labour; for example, shifting some labour-intensive industries to developing provinces to promote local employment rate and to attract migrant workers who have local hukous to return. Finally, the central government needs to implement the equalization strategy of public services, so that public service resources such as education and medical care can be dispersed to small and medium-sized cities to attract people to return by voting with their feet. Local governments in source provinces need to formulate relevant policies to encourage migrants to return to their hometowns and start businesses. Local governments in destination provinces need to control the size of the immigrant population by rationally arranging employment positions and resources.

Currently, population migration has two major impacts on economic and social development. On one hand, it drives element flow, resource allocation and information exchange among regions, while on the other, it promotes development imbalance among regions. When measuring the above effects, it is important to calculate and analyse the spatial concentration of population migration. Thus, GI is necessary. However, using traditional GI results in calculation bias, as it cannot solve the problem of missing data. One of the main contributions of this paper is that we have proposed GIc, and hence we can analyse the concentration of migration in a more accurate way. Compared to the traditional GI, the main improvements of GIc are listed as below: first, through re-standardization, GIc is restricted to the range of [0,1] under the circumstance of missing data, so that the comparability of results across time and regions is enhanced. Second, the GI may underestimate the spatial concentration of migration, and GIc can solve this problem and give more accurate measurements. Thus, our new indexes can better reveal the imbalanced and concentrated spatial nature of population migration, and they can provide guidance for further study of its mechanism analysis

and relevant policies research. In the future, the GI system should be extended to reflect the underlying population migration phenomena better, and it should be compared against other commonly used indexes in a more rigorous manner to develop an even more accurate indicator. In addition, the observed imbalanced nature of population migration should be studied with state-of-the-art quantitative models to reveal the major driving factors for mechanistic insights, which collectively provide this subject with a broad and deep meaning.

Compliance with Ethical Standards

Conflict of Interest   The authors declare that they have no conflict of interest.

References

Abel, G. J., & Sander, N. (2014). Quantifying global international migration flows. *Science, 343*(6178), 1520–1522.

Adey, P. (2010). *Mobility* (p. Routledge). London.

Allison, P. D. (1978). Measures of inequality. *American Sociological Review, 43*, 865–880.

Ammendola, G. (2005). The age of migration: International population movements in the modern world. *American Foreign Policy Interests, 27*(6), 537–542.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory, 2*(3), 244–263.

Bell, M., Black, M., Boyle, P., Duke-Williams, O., Rees, P., Stillwell, J., & Hugo, G. (2002). Cross-national comparison of internal migration: Issues and measures. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 165*(3), 435–464.

Bell, M., Charles-Edwards, E., Ueffing, P., Stillwell, J., Kupiszewski, M., & Kupiszewska, D. (2015). Internal migration and development: Comparing migration intensities around the world. *Population and Development Review, 41*(1), 33–58.

Bell, M., & Muhidin, S. (2009). *Cross-national comparison of internal migration*. New York, NY: United Nations Development Programme.

Bernard, A., Bell, M., & Charles-Edwards, E. (2014). Improved measures for the cross-national comparison of age profiles of internal migration. *Population Studies, 68*(2), 179–195.

Castles, S., De Haas, H., & Miller, M. J. (2013). *The age of migration: International population movements in the modern world* (5th ed.). New York, NY: Palgrave Macmillan.

Castles, S., & Miller, M. (2009). *The age of migration: International population movements in the modern world* (4th ed.). Basingstoke: Palgrave Macmillan.

Chen, W., Liu, W., Ke, W., & Wang, N. (2017). The spatial structures and organization patterns of China's city networks based on the highway passenger flows. *Acta Geographica Sinica, 72*, 224–241 in Chinese.

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

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

Ding, J., Liu, Z., Cheng, D., Liu, J., & Zou, J. (2005). Areal differentiation of inter-provincial migration in China and characteristics of the flow field. *Acta Geographica Sinica, 1*(11), 106–114 in Chinese.

Duan, P., & Liu, C. (2012). Influence of China's inter-provincial migration on regional disparity. *China Population Resources and Environment, 22*(11), 60–67.

Duncan, O. D., & Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review, 20*(2), 210–217.

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. *The Professional Geographer, 57*(2), 295–311.

Gini, C. (1926). The contributions of Italy to modern statistical methods. *Journal of the Royal Statistical Society, 89*(4), 703–724.

Griffith, D. A., & Amrhein, C. G. (1991). *Statistical analysis for geographers*. Englewood Cliffs, NJ: Prentice Hall.

Gu, H., Shen, T., Liu, Z., & Meng, X. (2019). Driving mechanism of interprovincial population migration flows in China based on spatial filtering. *Acta Geographica Sinica, 74*(2), 222–237 in Chinese.

Gu, Z., Gu, L., Eils, R., Schlesner, M., & Brors, B. (2014). Circlize implements and enhances circular visualization in R. *Bioinformatics, 30*(19), 2811–2812.

Janská, E., Čermák, Z., & Wright, R. (2014). New immigrant destinations in a new country of immigration: Settlement patterns of non-natives in the Czech Republic. *Population, Space and Place, 20*(8), 680–693.

Krugman, P. R. (1991). *Geography and trade*. MIT Press.

Li, Y., Liu, H., & Tang, Q. (2015). Spatial-temporal patterns of China's interprovincial migration during 1985-2010. *Geography Research, 34*(6), 1135–1148 in Chinese.

Liang, Z. (2016). China's great migration and the prospects of a more integrated society. *Annual Review of Sociology, 42*, 451–471.

Liu, Y., & Shen, J. (2016). Modelling skilled and less-skilled interregional migrations in China, 2000-2005. *Population, Space and Place, 23*(4), e2027.

Liu, Y., & Xu, W. (2017). Destination choices of permanent and temporary migrants in China, 1985–2005. *Population, Space and Place, 23*(1), e1963.

Lu, M., Ou, H., & Chen, B. (2014). Rationality or bubble? An empirical study on urbanization, migration and housing prices. *The Journal of World Economy, 37*(1), 30–54 in Chinese.

Lu, M., Gao, H., & Sato, H. (2012). On urban size and inclusive employment. *Social Sciences in China, 10*, 47–66 in Chinese.

Lyon, M., Cheung, L. C., & Gastwirth, J. L. (2016). The advantages of using group means in estimating the Lorenz curve and Gini index from grouped data. *The American Statistician, 70*(1), 25–32.

Lau, P. L., Koo, T. T., & Dwyer, L. (2017). Metrics to measure the geographic characteristics of tourism markets: An integrated approach based on Gini index decomposition. *Tourism Management, 59*, 171–181.

Lerman, R. I., & Yitzhaki, S. (1984). A note on the calculation and interpretation of the Gini index. *Economics Letters, 15*(3–4), 363–368.

Plane, D. A., & Mulligan, G. F. (1997). Measuring spatial focusing in a migration system. *Demography, 34*(2), 251–262.

Pu, Y., & Ge, Y. (2016). Multilateral mechanism analysis of interprovincial migration flows in China. *Acta Geographica Sinica, 26*(2), 205–216 in Chinese.

Qi, W., Abel, G., Muttarak, R., & Liu, S. (2017). Circular visualization of China's internal migration flows 2010–2015. *Environment And Planning A: Economy and Space, 49*(11), 2432–2436.

Rogers, A., & Raymer, J. (1998). The spatial focus of US interstate migration flows. *International Journal of Population Geography, 4*(1), 63–80.

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

Smith, D., Finney, N., & Halfacree, K. (2016). *Internal migration* (p. Routledge). New York.

Theil, H. (1972). Statistical decomposition analysis*; with applications in the social and administrative sciences* (No. 04; HA33, T4.). CA: North-Holland Pub. Co.

United Nations, Department of Economic and Social Affairs, population division (2017). *International Migration Policies: Data Booklet*(ST/ESA/ SER.A/395). Retrieved from UN Economic & Social Affairs: http://www.un.org/en/development/ desa/population/publications/pdf/policy/international_migration_policies_data_booklet.pdf. Accessed 1 May 2018.

Wang, G., & Pan, Z. (2016). The robustness of China's migration and Heihe-Tengchong line. *China Population Today, 4*(8) in Chinese.

Wang, G., Pan, Z., & Lu, Y. (2012). China's inter-provincial migration patterns and influential factors: Evidence from year 2000 and 2010 population census of China. *Chinese Journal of Population Science, 5*, 2–13 in Chinese.

White, M. (1986). Segregation and diversity measures in population distribution. *Population Index, 52*(2), 198–221.

Wu, H. X. (1994). Rural to urban migration in the People's Republic of China. *The China Quarterly, 139*, 669–698.

Yang, C., & Ning, Y. (2015). Evolution of spatial pattern of inter-provincial migration and its impacts on urbanization in China. *Geographical Research, 34*(8), 1492–1506.

Yang, L., Hui, L., & Qing, T. (2015). Spatial-temporal patterns of China's interprovincial migration during 1985-2010. *Geographical Research, 34*(6), 1135–1148 in Chinese.

Yitzhaki, S. (1983). On an extension of the Gini inequality index. *International Economic Review, 24*(3), 617.

Zahn, H., Steif, A., Laks, E., Eirew, P., VanInsberghe, M., Shah, S., et al. (2017). Scalable whole-genome single-cell library preparation without preamplification. *Nature Methods, 14*(2), 167–173.

Zang, Y. Z., Zhou, S. L., & Wu, Y. Y. (2016). The volume changes and spatial pattern dynamics of China's interprovincial migration: A perspective of social network analysis. *Human Geography, 4*, 16 in Chinese.

Zeng, M., Wu, R., & Zhang, S. (2013). Research on new situation of Chinese population redistribution and its socio-economic effects: Based on the analysis of Chinese sixth census data. *Population Journal, 35*(5), 15–25 in Chinese.

Affiliations

Ziliang Liu[1] · Hengyu Gu[2]

[1] School of Economics and Management, South China Normal University, No.378 Waihuanxi Road, Guangzhou 510006, China

[2] School of Government, Peking University, No.5 Yiheyuan Road, Beijing 100871, China