

Using mobile phone big data to discover the spatial patterns of rural migrant workers' return to work in China's three urban agglomerations in the post-COVID-19 era

EPB: Urban Analytics and City Science
2022, Vol. 0(0) 1–17
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23998083211069375
journals.sagepub.com/home/epb



Kai Liu , **Pengjun Zhao** and **Dan Wan**

Human Activities and Regional Governance Big Data Analytics Lab, School of Urban Planning and Design, Peking University Shenzhen Graduate School, Shenzhen, China

Xiaodong Hai

Smart Steps Digital Technology Co., Ltd, Beijing, China

Zhangyuan He and **Qiyang Liu**

Human Activities and Regional Governance Big Data Analytics Lab, School of Urban Planning and Design, Peking University Shenzhen Graduate School, China

Yonghui Qu, **Xue Zhang** and **Kaixi Li**

Smart Steps Digital Technology Co., Ltd, Beijing, China

Ling Yu

Human Activities and Regional Governance Big Data Analytics Lab, School of Urban Planning and Design, Peking University Shenzhen Graduate School, China

Abstract

Knowing how workers return to work is a key policymaking issue for economic recovery in the post-COVID-19 era. This paper uses country-wide time-series mobile phone big data (comparing monthly and annual figures), obtained between February 2019 and October 2019 and between February 2020 and October 2020, to discover the spatial patterns of rural migrant workers' (RMWs') return to work in China's three urban agglomerations (UAs): the Beijing–Tianjin–Hebei Region, the Yangtze River Delta and the Pearl River Delta. Spatial patterns of RMWs' return to work and how these patterns vary with location, city level and human attribute were investigated using the fine-scale social sensing related to post-pandemic human mobility. The results confirmed the multidimensional spatiotemporal differentiations, interaction effects between variable pairs and

Corresponding author:

Pengjun Zhao, School of Urban Planning and Design, Peking University Shenzhen Graduate School, No. 2199, Lishui Road, Shenzhen 518055, China.

Email: pengjun.zhao@pku.edu.cn

effects of the actual situation on the changing patterns of RMWs' return to work. The spatial patterns of RMWs' return to work in China's major three UAs can be regarded as a comprehensive and complex interaction result accompanying the nationwide population redistribution, which was affected by various hidden factors. Our findings provide crucial implications and suggestions for data-informed policy decisions for a harmonious society in the post-COVID-19 era.

Keywords

China's three urban agglomerations, mobile phone big data, post-COVID-19 era, rural migrant worker

Introduction

The 2019 coronavirus (COVID-19) pandemic has affected on workers' migration, labour supply (Pullano et al., 2020), capital flow (UN News, 2020) and human behaviours (Li et al., 2021). The crisis has interrupted supply chains and brought job losses, insolvencies and declines in revenue across the world. The travel bans, border closures and quarantine measures caused by COVID-19 have meant that migrant workers cannot return to their workplaces.

Population movements by rural migrant workers (RMWs), as a proxy for the resource optimization of industrialization and urbanization (Gu et al., 2007), can bring a significant potential for socioeconomic development in cities (Wong et al., 2007), which is a kind of representation of the interaction intensity between cities under urban extra growth and spatial pattern evolution (Ma and Zhang, 2020). An RMW is a worker with a rural household registration who is employed in an urban workplace far away from his or her hometown (Wang et al., 2019), and it is a feature of the new social stratum (Gu et al., 2007). In China's context, an agricultural registered permanent residence still holds cultivated land, which means that a RMW bears the obligation to earn money to support other members of his/her original family, seek his/her opportunity of career development or temporarily gain a foothold in cities for becoming a potential resident in a new city. Additionally, many factors like rent pressure in cities, farming business and demand to return home prompt RMWs' periodic migration behaviour, which is a social phenomenon in China's specific development with Chinese characteristics, which are slightly different from other countries of the Oriental world, such as Japan (Asian Growth Research Institute, 2020), Vietnam (Nguyen et al., 2015) and Indonesia (Resosudarmo et al., 2010).

A correct understanding of RMWs' return to cities along with China's urbanization is strategically important in planning urban, rural economic and social development. Furthermore, knowing how workers return to work is also a key policymaking issue for economic recovery in the post-COVID-19 era.¹ However, it is a serious challenge to do so in large countries, particularly, in urbanizing countries with millions of RMWs who migrate across the whole country for work. There were an estimated 291 million (for 2019) and 286 million (for 2020) RMWs in China (askci Corporation, 2021), which accounted for more than one third of the entire working population (Wang et al., 2019).

In particular, the mobile phone big data approach brings opportunities to observe, measure and shape the details of migrants' migration at a finer and wider scale. As the acquisition of mobile phone big data is based on the spatial behaviour of massive groups, it can be better applied to perceive the behaviour patterns of typical subgroups and to mine the underlying geographic spatial rules (Wu et al., 2020). Moreover, using mobile phone big data for social sensing can capture the spatial distribution/interaction of human behaviour in a coupling relationship with the geographical environment at multidimensional levels (Liu et al., 2015). In this sense, establishing a data-driven approach to achieve the fine-scale social sensing related to post-pandemic human mobility

supported by a comprehensive geospatial geodatabase is necessary. Furthermore, good data-driven approaches can support policymakers in formulating related policies to describe the influence of the lockdown measures, achieving sustainable resilient cities and seeking the safe development in the post-epidemic era (Esposito et al., 2021).

Despite the COVID-19 crisis, China maintained its growth momentum, and its normalization of interregional migration benefitted from strong control measures (such as quarantine, social distancing and the isolation of infected populations) amid pandemic- and coronavirus-induced lockdowns. Since the 13th National People's Congress of 25 May 2020, the epidemic in China has effectively been brought under control, and China should make efforts to resume work and production (cf. China Talk, 2020). This impact of the COVID-19 response in China is encouraging, and the drastic control measures implemented in China have substantially mitigated the spread of COVID-19 (Kraemer et al., 2020). Policymakers from other countries should seriously consider broader-scale social distancing like China, especially in countries that still suffer from the coronavirus-induced pandemic (Anderson et al., 2020). The good situation of RMWs returning to work is important for China's economic recovery, and meaningful for the social welfare department and labour security department. Understanding China's evidence can help other countries to address the work resumption problem in the post-COVID-19 era. As extensive case finding and isolation are needed to relax lockdown measures gradually and to reduce the socioeconomic pressures, it seems that Chinese RMWs had enough courage and willingness to leave their hometowns to go to the megacities for work when the suggestions that work and production should begin to resume. The authors argue that this hypothesis needs to be confirmed using mobile phone big data to understand the interaction between spatial differentiation and human attribute characteristics in the case of RMWs' return to work in China.

An urban agglomeration (UA) refers to a city group with multiple urban cores within a specific region, which is formed by the developed infrastructure network tied to close traffic transportation and economic communication between urban cores. It requires a form of compact spatial organization that ultimately achieves a high degree of urban homogenization and integration. Within the framework of China's UA development strategy, researchers should focus more on the pattern of RMWs within as comprehensive an urban system as possible (e.g. at a national scale) to obtain more arguments on the background of planning guidance for urban integration.

Against this background, this paper uses country-wide time-series mobile phone big data (comparing monthly and annual figures) to discover the spatial patterns of RMWs' return to work in China's three UAs: the Beijing–Tianjin–Hebei Region (BTH), Yangtze River Delta (YRD) and Pearl River Delta (PRD) in the post-COVID-19 era. Furthermore, this paper reveals how these patterns vary with location, city level and human attribute for better understanding the spatial patterns of RMWs' return to work within the unified layout of China's urban system.

Literature review

Focussing on the characteristics of population movement and its related factors, Malmberg and Clark (2020) showed that residential classification by income in megacities was associated with human migration and ethnic concentration. Pullano et al. (2020) argued that the reduction in human mobility due to the COVID-19 was associated with active population, the employments situation and the socioeconomic levels of the regions. Zhou et al. (2021) examined the patterns of interprovincial and intraprovincial migration using statistical data. However, they did not discuss the interaction effects between variable pairs on the patterns of population movement, such as city level, location and human attribute.

In the context of the oriental world, the situations of labour division, family responsibility, human lifestyle and migration are usually accompanied by great differences in gender and age (e.g. Liu 2018; Wako, et al., 2001). As mentioned in the *Introduction* section, Chinese RMWs' periodic

migration behaviour has some specific features. This paper focuses on the interaction effects between spatial differentiation and human attribute characteristics in the case of RMWs' return to work in China to provide a better understanding of RMWs' return to work, especially in the post-epidemic scenario.

Literature abounds on understanding the spatial patterns of population flow revealed by using mobile phone big data. The use of such data in the relevant literature has focused on performing network analysis (e.g. Wang et al., 2019; Wei et al., 2018), measuring urban polycentricity (e.g. Liu et al., 2021; Wei et al., 2020) and evaluating urban planning (Cao et al., 2021; Liu et al., 2021). However, there is limited literature on how the patterns of RMWs' returns to cities vary with human attributes in a uniform and comparable time series. Although some existing scientific reports monitored and examined the spatial patterns of RMW numbers from 2008 to 2019 using survey data and statistical yearbooks, they only discussed the statistical characteristics of RMW numbers with respect to human attributes without sufficient in-depth quantitative analyses to explore the interaction effects between different factors. Additionally, studies seldom focus on the time-series patterns of population movements affected by COVID-19 on a national scale within a comparative analytical framework. Our work not only uses the mobile phone big data to achieve a good data-driven policy decision support within a comprehensive multidimensional and multi-scenario analysis framework, but also provides detailed insight into RMWs' return to work at different scales and provides accurate policy guidance on China's labour recovery for each city in the three UAs.

Referring to these achievements, most previous studies did not fully pay attention to the spatial interaction between the UAs within a comprehensive urban system, due to the insufficient geo-spatial integration of big data and an imperfect analysis framework. The authors argue that some validation and exploratory analyses to discover the results of the coronavirus-induced pandemic and to reconfirm the regional bias are also necessary. For decision-makers and politicians, the use of big data analytics and a data-driven approach should contribute to optimizing the labour supply policy on a national scale in the post-COVID-19 era. Using a set of time-series geo-tagged big data, this study adds some insights concerning the spatial patterns of RMWs' return to work and how these patterns vary with other factors, for example, human attributes, city level and location, aiming to fill the existing research gaps by focussing on a country-wide spatial scale and by performing a series of multi-variable analysis.

Data and methodology

China's three major urban agglomerations

The three vast UAs, that is, BTH, YRD and PRD, are the most developed and populated urban areas in China, as they accounted for 9.0% (BTH), 19.8% (YRD) and 8.6% (PRD) of the national GDP and 7.9% (BTH), 17.1% (YRD) and 9.3% (PRD) of the country's total densely inhabited district population in 2018 (China.org.cn, 2019). Based on our calculation and the official figures (cf. National Earth System Science Data Center, 2021), the numbers of RMWs (cross-domain) working in three UAs in 2019 accounted for 44.2% of the total. Choosing the three major UAs in China as the target area gives us a high chance of achieving our research purpose. Figure 1(a) shows the location of China's three major UAs. Figure 1(b)–(d) demonstrate the location of the 75 cities that belong to these three UAs. For this study, the PRD is regarded as the Great Pearl River Delta.²

Method of identifying rural migrant workers from mobile phone big data

This paper uses the mobile phone big data provided by a certain telecom operator in China with 445 million users. Based on these samples and their corresponding regional census-based population

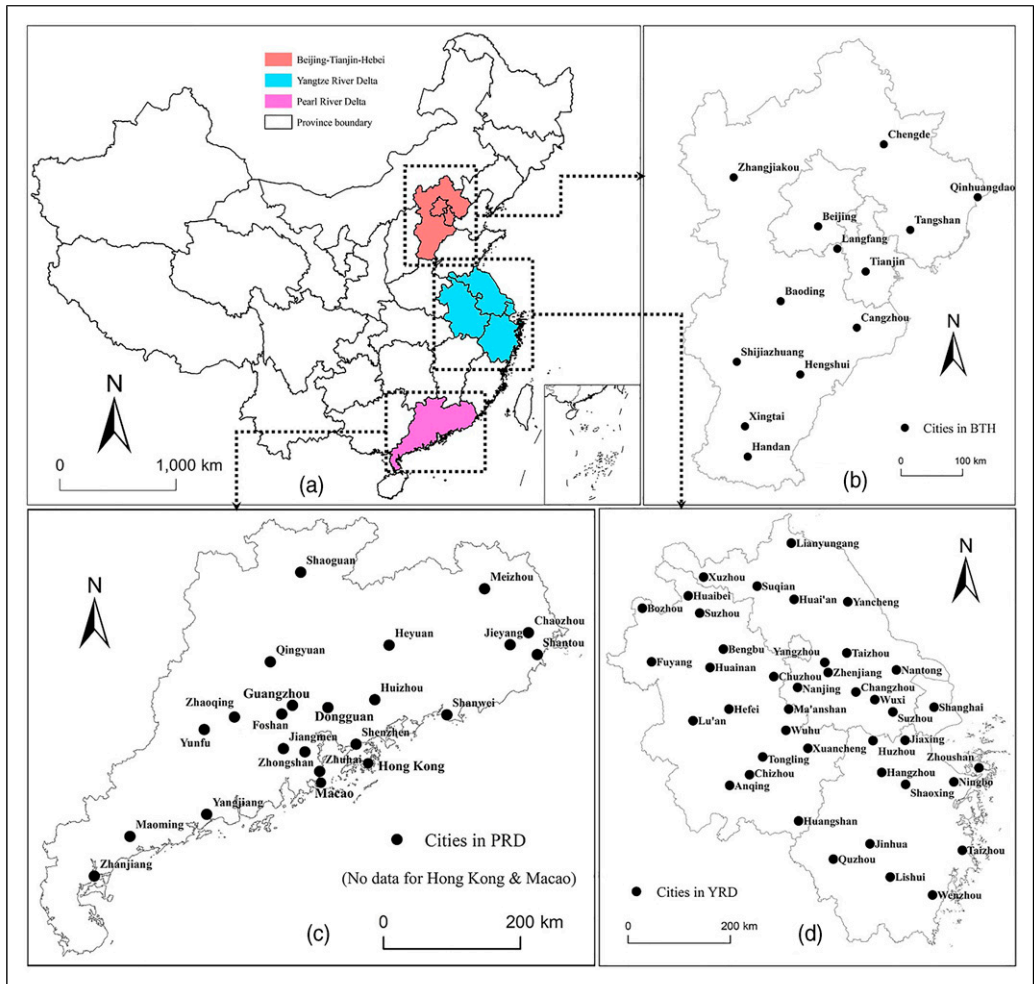


Figure 1. Study area. (a) Location of BTH, YRD and PRD. (b) Cities in Beijing–Tianjin–Hebei. (c) Cities in PRD. (d) Cities in YRD. YRD: Yangtze River Delta; PRD: Pearl River Delta.

weights, 167 million RMWs (cross-domain) from China's 367 cities in 2020 are identified. This figure has a coincidence rate of 98.98% with the official figure (cf. [askci Corporation, 2021](#)). From February 2019 to October 2019 and from February 2020 to October 2020, the time-series numbers of RMWs from 367 cities working in the three UAs were identified and measured, using a range of machine learning and data mining technology, such as the total population extrapolation algorithm,³ spatiotemporal human mobility and labour migration models, and the return-to-city/return-to-work model. [Figure 2](#) shows the method of generating the mobility data for RMWs.⁴

Briefly, the technical section of this telecom operator employed a method involving spatio-temporal constraints and human behaviour identification to handle the raw data (including a desensitization and encryption process), while setting specific rules (such as lower income, periodic rural-urban migration and typical mobility patterns before/after the Spring Festival, non-college student) to distinguish the RMWs from the total population (cf. [Figure 2](#) and the Section S1 of [Supplementary Materials](#)). More detailed information about how to identify the typical subgroups of the actual population (RMW, in this case) is in [Wu et al. \(2020\)](#).

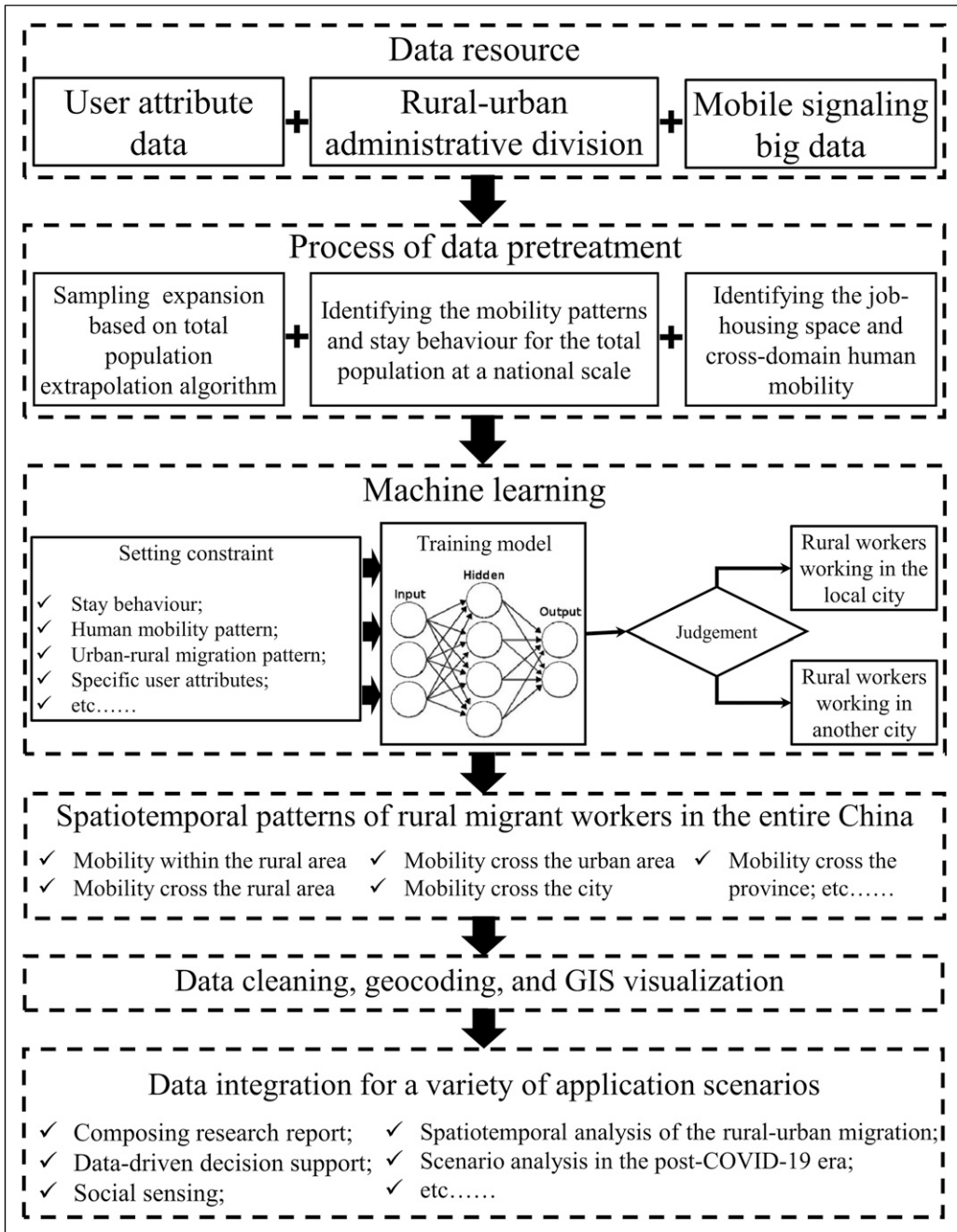


Figure 2. Method of identifying rural migrant workers from mobile phone big data.

Newest grade classification of China's 367 cities

As one of our targets is to reveal the spatial differences in RMWs' movement (from the viewpoint of both hometown and workplace) by city level, this study considers the newest rule of grade classification (cf. [Research Institute of New First-tier City \[RINFC\], 2020](#)) of China's 367 cities (see the Section S2 of [Supplementary Materials](#)). This study considers five kinds of city level (Class 1 to 5: C1, C2, C3, C4 and C5). That is to say, C1 corresponds to 19 first-tier or new first-tier cities and C2 corresponds to 30 second-tier cities. The locations and grade classifications of China's 367 cities are shown in [Supplementary Figure S1](#).

Analytical framework

To discover the spatial characteristics of RMWs' return to work in the three UAs focussing on three aspects: quantitative trend, spatial differentiations and the interaction effects between variable pairs, we design an original analytical framework to achieve the research purpose, as shown in [Figure 3](#). First, geo-coding and data cleaning technology with the MATLAB program are applied to get several groups of time-series origin-destination (OD) matrices summarized by location (i.e. seven geographical divisions, three UAs, 367 cities as hometowns and 75 cities in three UAs as workplaces), city level and human attribute (i.e. gender and age of the RMWs). In this case, the

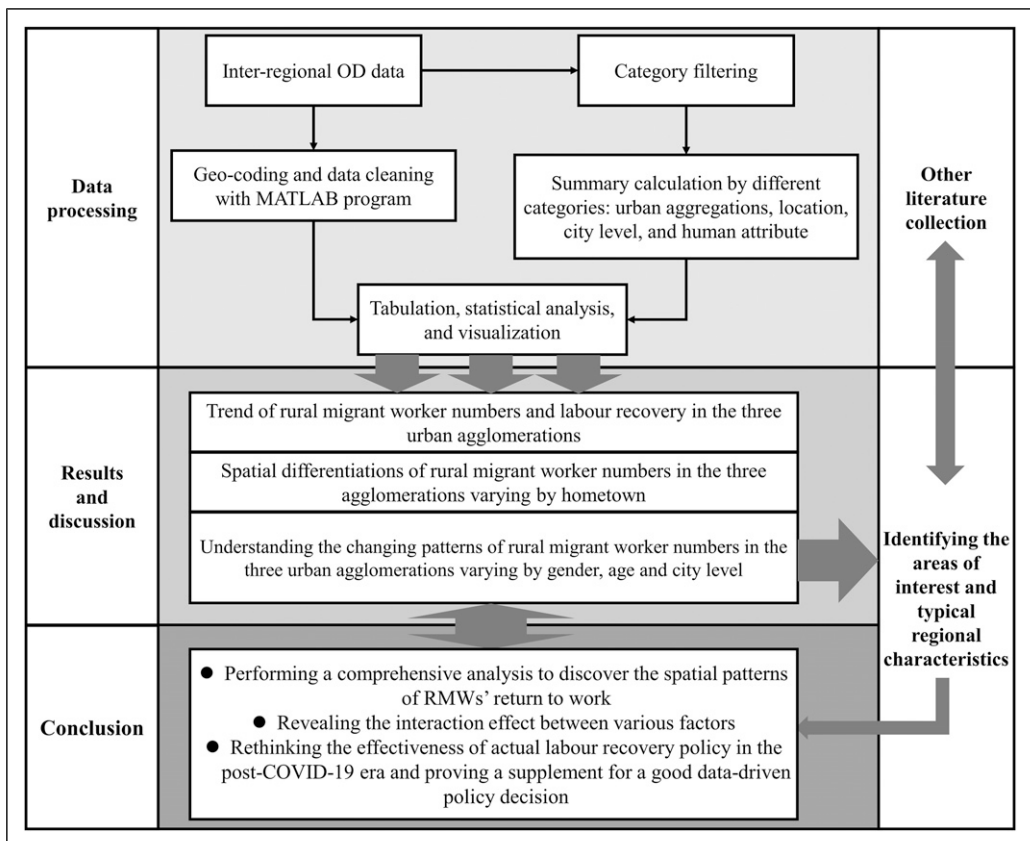


Figure 3. Analytical framework.

origin means one of China's 367 cities as a hometown for the RMW, while the destination means one of the 75 cities located in three UAs as a workplace for the RMW. In particular, this study uses a kind of geo-tagged population-based data at 18 points in time from February 2019 to October 2019 and February 2020 to October 2020, not flow-based OD data. Then, the tabulation, statistical analysis and Geographic Information System (GIS) visualization are performed based on the results of the summary calculation by different categories.

Next, this study discusses in detail the spatial characteristics of RMWs' return to work in the three UAs focussing on three aspects: quantitative trend by labour supply area (i.e. RMWs' hometown) and RMWs' workplace for three UAs, spatial differentiations of RMW numbers varying by hometown, and changing patterns varying by gender, age and city level. Alongside the process of GIS visualization and statistical analysis, other literature (such as news information and government documents) has also been collected to confirm the key findings when the areas of interest and typical regional characteristics are identified from our results. Finally, the conclusion follows a presentation of the results to reveal the hidden mechanism and to provide a good data-driven decision support (cf. [Figure 3](#)).

Results and discussion

Trend of rural migrant worker numbers and labour recovery in the three urban agglomerations

From February 2020 to October 2020, the results show that the ratio of RMW numbers working in the three UAs in March 2020 had drastically reduced to 53.29% (BTH), 68.42% (YRD) and 68.18% (PRD) of their numbers in March 2019 (cf. [Figure 4\(a\)](#)). This reflects that the Chinese Government promulgated strong control measures amid pandemic and coronavirus-induced lockdowns and that the rural-to-urban migration during the 2020 Spring Festival travel rush⁵ was severely suppressed. Since then, the RMW numbers in the three UAs have shown a gradual recovery trend. By October 2020, the RMW numbers in YRD had completely recovered to the same level as those in October 2019. The year-over-year ratios of RMW numbers in the other two UAs had also recovered by October 2020 to over 95% of their values in October 2019 (cf. [Figure 4\(a\)](#)).

Referring to [Figure 4\(b\)](#), due to the policy and psychological implications of COVID-19, in March 2020, just after the 2020 Spring Festival, remarkable numbers of RMWs preferred to stay in their hometowns or decided to go to workplaces in cities other than the three UAs, which led to a substantial decline in the numbers of RMWs. Human employment behaviours in the three UAs completely changed in March 2020. When considering the quantitative differences in RMW numbers by hometowns in China's seven geographical divisions in March 2020 (cf. [Figure 4\(b\)](#)), we found that: (1) the RMWs whose hometowns were in North China could not easily return to the BTH UA; (2) the numbers of RMWs coming from East and Central China decreased remarkably in the YRD UA; and (3) the PRD UA failed to attract enough RMWs to work there at that time, especially from Central and South China.

To indicate where the RMWs went and to speculate how many RMWs postponed their returns to the large cities due to COVID-19 after the 2020 Spring Festival on March 2020, we also checked the year-over-year increasing RMW numbers staying in their hometowns in 367 cities in March 2020 as compared to March 2019 (not shown here). Based on our calculation, the populous provinces (i.e. Henan, Hubei and Chongqing) kept more RMWs who would otherwise have gone away to work in March 2020 than in March 2019. Hubei Province, as the site of the initial outbreak in China (cf. [Supplementary Figure S1](#)), still struggled with strictest lockdown measures to tackle the spread of COVID-19 in March 2020 (cf. [The Paper, 2020](#)). Similarly, [Figure 4\(c\)](#) explains the year-over-year numbers of RMWs by hometown in China's seven geographical divisions (October 2020 vs

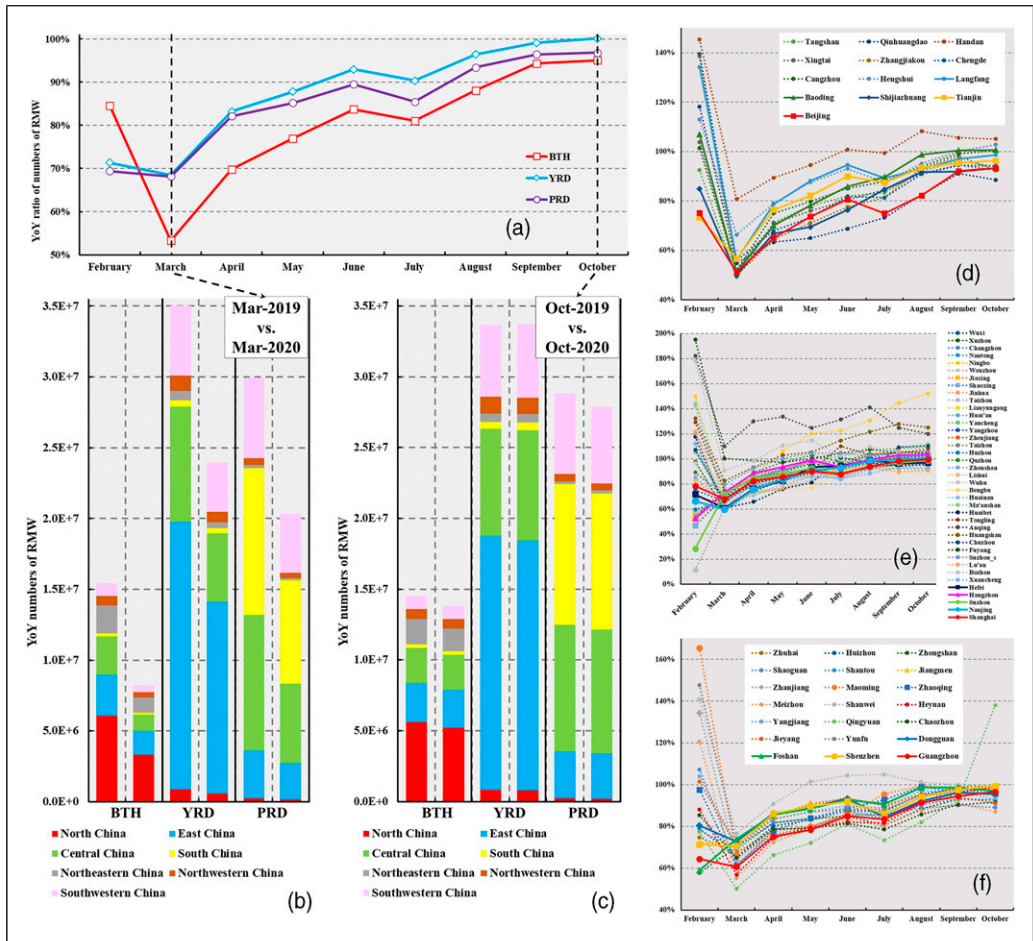


Figure 4. Trend of RMW numbers and labour recovery in the three UAs from February 2020 to October 2020. (a) Year-over-year ratios of RMW numbers in the three UAs. (b) Year-over-year numbers of RMWs by hometowns in China’s seven geographical divisions (March 2020 vs. March 2019). (c) Year-over-year numbers of RMWs by hometowns in China’s seven geographical divisions (October 2020 vs. October 2019). (d–f) Year-over-year ratios of RMW numbers (labour recovery) in the three UAs by city. RMW: rural migrant workers; UA: urban agglomerations.

October 2019), which indicates that total RMW numbers have basically recovered in three UAs and that the structures of RMWs’ hometown working in three UAs have become slightly different. Additionally, the situation of labour recovery in the three UAs reflected in the RMW numbers has various different features in the 75 cities (cf. Figure 4(d)–(f)). Some related descriptive findings can be found in the Section S3 of [Supplementary Materials](#).

Although the total number of RMWs in the three UAs had basically recovered to the same level by October 2019 (cf. Figure 4) when it comes to structure and quantity, it is apparent that there are still some spatial differentiations of RMW distribution in the three UAs with hometowns in the post-COVID-19 era (cf. Figure 4(b) and (c)), such as RMWs coming from North China to the BTH, those coming from East China to the YRD and those coming from Central China to the PRD. This argument is considered in the next section.

Spatial differentiation of rural migrant worker numbers in the three urban agglomerations varying by hometown

The *Trend of Rural Migrant Worker Numbers and Labour Recovery in the Three Urban Agglomerations* section focused on a quantitative trend of RMW in three UAs from February 2020 to October 2020 as compared with the period from February 2019 to October 2019. As a supplement, this study also examines how the spatial differentiation of RMW numbers in the three agglomerations varied with hometown, to confirm whether RMWs' lifestyle and career choices were changed by the COVID-19. For RMWs, hometown locations and workplaces are affected by various spatial constraints and socioeconomic drivers, such as spatial proximity, traffic convenience, capital and personal pursuit of urban life. This section makes the following visualizations to investigate this issue.

Figure 5(a), (c) and (e) show the OD cases (hometown to workplace) in blue lines with decreasing numbers of RMWs in October 2020 as compared with October 2019 in the three UAs. Figure 5(b), (d) and (f) express the OD cases in red lines with the increasing RMW numbers. For instance, for the BTH UA, the numbers of RMWs coming from closer areas significantly decreased, such as Northeastern China, Hebei, and the south part of Shanxi (cf. Figure 5(a)). However, in the increasing case of the BTH UA, the numbers of RMWs coming from the populous provinces, such as Henan, Sichuan and Chongqing, slightly increased (cf. Figure 5(b)). This might be a result of the long-term effects of strict epidemic control measures in the BTH UA, changing human mobility patterns in some places and changing human employment behaviour in others.

Similarly, for the YRD UA, we find that the employment gap has been filled by newcomers (i.e. RMWs) from closer areas, such as Anhui, Hunan, Guizhou and Henan. This implies that the YRD UA itself has more domestic-oriented recruitment companies (cf. *Zhejiang News, 2020*), which means that it has a better RMW labour recovery, as mentioned in *Trend of Rural Migrant Worker Numbers and Labour Recovery in the Three Urban Agglomerations*. Compared with the BTH UA (13 cities) and the PRD UA (21 cities), the YRD UA has a relatively large number of cities (41 cities). Additionally, the human mobility associated with the cities of the YRD UA is also relatively greater (cf. Figure 5(c)). Therefore, from this viewpoint, the good recovery of RMW numbers in the YRD UA was helped by its good urban integration.

The situation in the PRD was more complicated. For example, RMWs who were supposed to work in Guangzhou (central city) chose to go elsewhere, such as Shenzhen and Zhuhai (cf. Figure 5(e) and (f)). In other words, the internal flow, RMWs' workplace choices, RMWs' lifestyle and labour supply-demand structure changed somehow among the cities of the PRD UA.

Patterns of rural migrant worker numbers in the three urban agglomerations varying by gender, age and city level

This section discusses how the changing patterns of RMW numbers vary by human attributes (i.e. gender and age) for the five city levels (cf. the map shown in [Supplementary Figure S1](#)) due to COVID-19 and investigates the interaction effects between variable pairs on these patterns. To highlight how the human attributes affect RMW numbers and to present the results more effectively, we offer five boxplots and five heat maps using the data from March 2020 and March 2019 (cf. Figure 6). Based on this kind of visualization, we can clearly grasp the changing patterns of RMW numbers in the 367 cities in terms of five human attribute categories: male, female, youngster (16–29 years old), adult (30–49 years old), and middle aged and elderly (50–65 years old).

By March 2020, when the Spring Festival ended, human attributes had important, complex and varying impacts on the population redistribution (cf. Figure 6). Generally speaking, male and adult RMWs coming from C3 and C4 cities are the main source of labour supply for the three UAs.

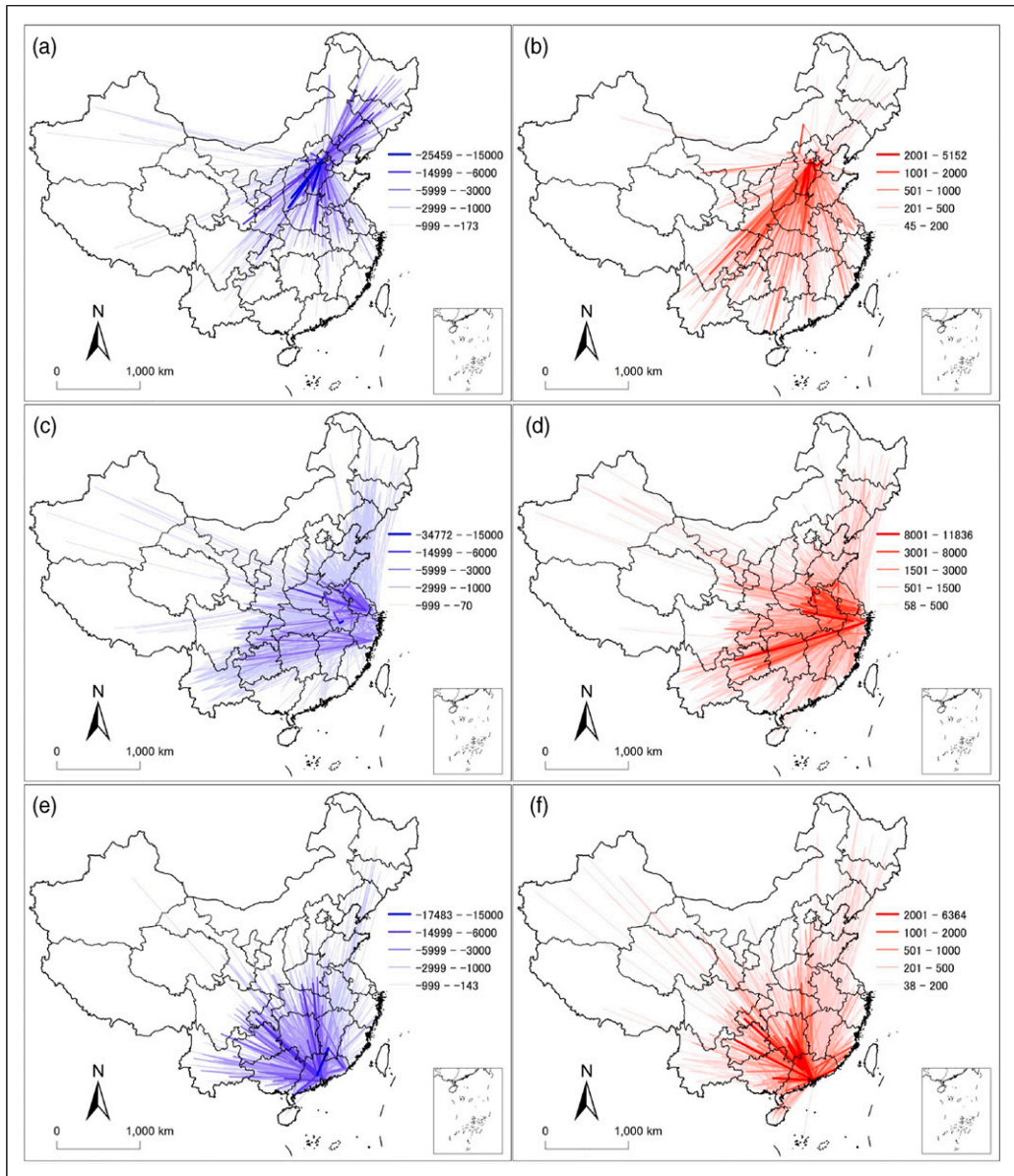


Figure 5. Spatial differentiation of rural migrant worker numbers in the three UAs varying with hometown. (a, c and e) Decreasing OD cases in three UAs. (b, d and f) Increasing OD cases in three UAs. UA: urban agglomerations; OD: origin-destination.

However, the reduction in rates of female and young RMW numbers to the three UAs tends to be higher. Specifically speaking, RMW numbers coming from mainland first-tier cities (e.g. Chongqing and Hefei) have larger quantitative decreases (cf. Figure 6(a)). Differences in human attributes mainly affected females and youngsters (cf. Figure 6(b)). Some satellite cities, such as Foshan and Dongguan, as satellite cities of Guangzhou, also had noteworthy decreases in numbers of RMWs.

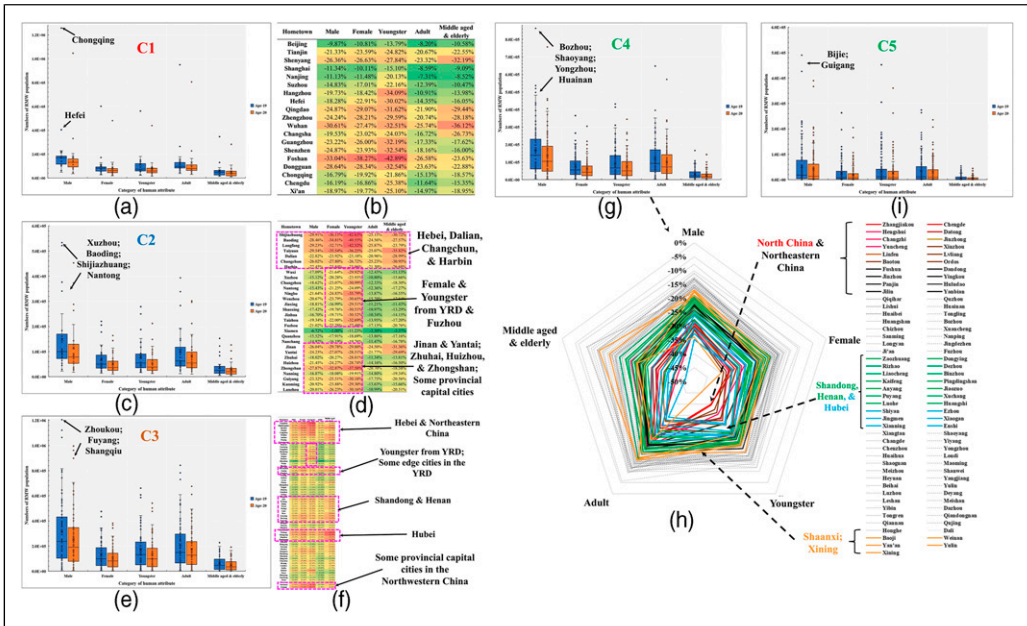


Figure 6. Pattern of rural migrant worker numbers in the three urban agglomerations by five city levels varying with human attributes. (a, c, e, g and i) Boxplots of changing pattern of RMW numbers in terms of different human attributes by five city levels. (b, d, f and h) Heat maps revealing the changing percentages of RMW numbers in terms of different human attributes by hometown. RMW: rural migrant workers. Note: Results for C5 cities are omitted: low RMW numbers but large city amounts.

In the case of C2 cities (cf. Figure 6(c)), edge cities (e.g. Shijiazhuang and Baoding) or hub cities (e.g. Xuzhou) within the three UAs tended to attract more RMWs. Moreover, C2 cities in North and Northeastern China restrained many local RMWs from going to the three UAs (cf. Figure 6(d)). Besides, female and young RMWs coming from C2 cities within the YRD also had an obvious impact. In addition, some provincial capital cities (e.g. Jinan and Kunming) and some local medium-sized cities (e.g. Zhuhai and Zhongshan) also kept a lot of the local RMW population, who might be supposed to work in the three UAs in the non-epidemic scenario (cf. Figure 6(d)).

As shown in Figure 4(a), the year-over-year ratios of RMW numbers in the three UAs almost completely reached the same level (95%) as September 2019; therefore, this study only considered and analysed September's results, eliminating the impact of China's National Day on the human mobility. Our results for September 2020 showed that quantitative changing patterns displayed more spatial differentiation (similar to Figure 5), but less human attribute differentiation (the boxplot and the heat map of September results are omitted here).

Referring to Figure 6(e), the C3 cities provide the larger RMW numbers. From Figure 6(f), we find that the numbers of RMWs coming from C3 cities in Northeastern China, some edge cities within the YRD UA, RMW supply areas (e.g. Hebei, Shandong, Henan and Hubei) and some provincial capital cities in the Northwestern China had remarkable decreases. From Figure 6(g), (i) and (j), we find that the spatial biases by human attribute gaps tend to have basically consistent patterns of change. The changing patterns of RMW number from cities in North and Northeastern China, and some populous areas such as Henan, Hubei and Shaanxi, follow these patterns, as do C3 cities (cf. Figure 6(h)).

In particular, this section also considered that the situation of several some representative RMW supply areas like Henan, Hubei, Sichuan and Chongqing could provide some helpful findings to

urban planners for better data-driven decision support. For example, by September 2020, based on our calculation, the numbers of RMWs coming from Shanxi Province to the three UAs decreased by about 15.36% as compared to 2019, followed by three provinces in Northeastern China with a decrease of 9.33%, and then Yunnan Province with a decrease of 6.24%. RMWs from these areas were likely to choose to work in their hometown cities or closer cities except the three UAs. Moreover, although these representative RMW supply areas kept significant numbers of RMWs, due to the epidemic control measures and personal psychological factors in March 2020, by September 2020, RMWs were still choosing to work in the three UAs, since the decreases in RMW numbers from there were all less than 5%. This phenomenon in Henan, Hubei, Sichuan and Chongqing was different from the three provinces in Northeastern China and Sichuan. This might be due to the differences in human behaviour, lifestyle, human mobility connection and their relative spatial locations.

Conclusion

At a fine GIS-enabled level, we discovered various statistical characteristics and year-over-year changing patterns of RMW numbers in the three UAs from February 2020 to October 2020 as compared to the period from February 2019 to October 2019. Furthermore, we investigated how these patterns vary with location, city level and human attribute due to COVID-19. Particularly, within three scenarios, that is, before, amid and after the COVID-19 outbreak, our work describes RMWs' long-term movement across China, giving multidimensional recognition of spatiotemporal patterns of human mobility, and social sensing of the rural-urban migration flow (cf. [Figure 2](#)). The descriptive results indicate that: (1) compared to the BTH and PDR UAs, the labour recovery in the YDR UA was better; (2) changing patterns of RMW numbers working in three UAs have spatial differentiations by location, RMWs' hometown and city level; and (3) the interaction effects of human attributes, RMWs' hometown and city level on RMWs' return to work in three UAs have some specific differences and commonalities.

In a word, in the post-COVID-19 era, the spatial patterns of RMWs' return to work in China's major three UAs can be regarded as a comprehensive and complex interaction result accompanying a nationwide population redistribution, related to [Jia et al. \(2020\)](#)'s findings. The patterns are affected by various hidden factors, such as workplace location, city level, RMWs' hometown, human attributes, social psychology, etc. Additionally, future work should focus on this issue by adding more social-psychological and spatiotemporal-behavioural geographical contexts.

Our findings contribute to describing the geospatial characteristics and differences of RMWs' return to work in three UAs, and supporting policymakers across various levels of local governments or departments by providing feedback for the optimization of related decisions. Moreover, our achievements may promote the usage of mobile phone big data for social sensing. Measuring detailed indexes using mobile phone big data has a promising future, as it makes it easier to integrate various geographical characteristics within urban computing. This study used China's seven geographical divisions, RINFC's six grade classifications and five categories of human attribute to group the RMWs coming from 367 cities, which enabled us to uncover the unknown heterogeneity, complexity and regularity of human migration more comprehensively on a national scale. This study can add finer information on how location, city level and human attribute are associated with the spatial patterns of RMWs' return to work in the three UAs from multiple viewpoints.

From the aspect of human factors, future studies should focus on the impact of COVID-19 on workplace choice, lifestyle and psychological influence. From the aspect of employment, we deduced that the labour supply-demand structure and coronavirus-induced policies within an urban system can have a certain impact on RMWs' return to cities. For the geographical aspect, the relative spatial location within an urban system, city level and transportation convenience should be

considered in related studies on rural-to-urban migration. Additionally, as our future work, using the RMW data for the whole of China over a longer period is necessary to examine the changing patterns of spatial interaction and the competition between different UAs, not only the three major UAs.

The mobile phone big data we used undoubtedly has limited precision. Spatial uncertainty due to regional bias arising from the variable density of base stations might bring uncontrollable errors. Machine learning algorithms for identifying the actual RMW population might import some misjudgements. Furthermore, the size of actual RMW numbers presented here is an approximation based on an extrapolation (e.g. Wu et al., 2020). Additionally, uncertainty due to the market share varying with the location or user segments can always occur. On the premise of unified human attribute tags and a unified geo-coding method, future studies should promote the usage of geo-tagged big data with high credibility and few errors (e.g. Liu et al., 2021).

In conclusion, in view of the regional characteristics and differences in RMWs' return to work in the three UAs and their relationships with location, city level and human attribute, this study bridges the gap between the changing patterns of RMW numbers and multidimensional indexes, and it provides crucial implications and suggestions for data-informed policy decisions for a harmonious society in the post-COVID-19 era. The mobile phone big data analytic method for this paper may help policymakers to understand worker migrants more clearly and completely. The findings in this paper also provide new evidence on population migration behaviour in the post-COVID-19 era.

Acknowledgements

The authors gratefully acknowledge funding from National Natural Science Foundation of China (No. 41925003), UKRI's Global Challenge Research Fund (No. ES/P011055/1), and Ministry of Education Key Projects of Philosophy and Social Sciences Research (No. 18JZD029). This paper is for the Peak Urban Project from Peking University team.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study is supported by National Natural Science Foundation of China (No. 41925003), UKRI's Global Challenge Research Fund (No. ES/P011055/1), and Ministry of Education Key Projects of Philosophy and Social Sciences Research (No. 18JZD029).

ORCID iD

Kai Liu  <https://orcid.org/0000-0002-5788-9053>

Supplemental material

Supplemental material for this article is available online.

Notes

1. This expression is suitable in China's context, since the COVID-19 pandemic has largely subsided.
2. Hong Kong and Macao are not part of the analyses. The aggregated data in this study is in mainland of China.

3. The patent for this invention is: Method and device for calculating the market share of the telecom operator. Authorization code: CN110312213B. China National Intellectual Property Administration, People's Republic of China.
4. [Figure 2](#) contains the major framework for two patents. One has been explained in note 3. The other one concerns the method of identifying RMWs using mobile signalling big data. Application code: 2021108795492 (in processing).
5. The Spring Festival travel rush is famously known as the largest annual migration on earth. It takes place over a 40-day period calculated by the Chinese lunar calendar, when people from all corners of China travel to their hometowns to spend the Chinese Lunar New Year with their families.

References

- Anderson RM, Heesterbeek H, Klinkenberg D, et al. (2020) How will country-based mitigation measures influence the course of the COVID-19 epidemic? *Lancet* 395(10228): 931–934.
- askci Corporation (2021) Analysis of the number and income of migrant workers in China: 2015-2020. Available at: <https://www.askci.com/news/data/hongguan/20210119/0908281332761.shtml> (accessed 8 December 2021).
- Asian Growth Research Institute (2020) Survey report from Asian growth research institute. Available at: <http://www.agi.or.jp/reports/report2019-06.pdf> (accessed 8 December 2021).
- China Talk (2020) China's solution in the post-epidemic era. Available at: http://fangtan.china.com.cn/node_8018517.html (accessed 8 December 2021).
- China.org.cn (2019) China's three major megalopolises compared. Available at: http://www.china.org.cn/china/2019-12/30/content_75563387.htm (accessed 8 December 2021).
- Cao J, Li Q, Tu W, et al. (2021) Resolving urban mobility networks from individual travel graphs using massive-scale mobile phone tracking data. *Cities* 110(2021): 103077.
- Esposito D, Dipierro G, Sonnessa A, et al. (2021) Data-driven epidemic intelligence strategies based on digital proximity tracing technologies in the fight against COVID-19 in cities. *Sustainability* 13(2): 644.
- Gu S, Zheng L and Yi S (2007) Problems of rural migrant workers and policies in the new period of urbanization. *China Population, Resources and Environment* 17(1): 1–6.
- Jia JS, Lu X, Yuan Y, et al. (2020) Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* 582: 389–394.
- Kraemer MUG, Yang C-H, Gutierrez B, et al. (2020) The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 368(6490): 493–497.
- Li H, Zhang Y, Zhu M, et al. (2021) Impacts of COVID-19 on the usage of public bicycle share in London. *Transportation Research Part A: Policy and Practice* 150: 140–155.
- Liu K (2018) Spatiotemporal analysis of human mobility in Manila metropolitan area with person-trip data. *Urban Science* 2(1): 3.
- Liu K, Murayama Y and Ichinose T (2021) Exploring the relationship between functional urban polycentricity and regional characteristics of human mobility: a multi-view analysis under the layout of the Tokyo metropolitan area. *Cities* 111(2021): 103109.
- Liu Y, Liu X, Gao S, et al. (2015) Social sensing: a new approach to understanding our socio-economic environments. *Annals of the Association of American Geographers* 105(3): 512–530.
- Ma X and Zhang ZG (2020) A study on the spatial evolution characteristics and trend of the yangtze river delta region based on population migration. *Urban Planning Forum* 2020(5): 47–54.
- Malmberg B and Clark WAV (2020) Migration and neighborhood change in Sweden: the interaction of ethnic choice and income constraint. *Geographical Analysis* 53(2020): 259–282.
- National Earth System Science Data Center (2021) Data visualization of China's floating population from 2008 to 2019. Available at: <https://www.udparty.com/index.php/detail/article/details/?id=5121> (accessed 4 July 2021).

- Nguyen LD, Raabe K and Grote U (2015) Rural-urban migration, household vulnerability, and welfare in Vietnam. *World Development* 71: 79–93.
- Pullano G, Valdano E, Scarpa N, et al. (2020) Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the COVID-19 epidemic in France under lockdown: a population-based study. *The Lancet Digital Health* 2(12): e638–e649.
- Resosudarmo BP, Suryahadi A, Purnagunawan RM, et al. (2010) The socio-economic and health status of rural-urban migrants in Indonesia. In: Meng X, Manning C, Shi L, et al. (eds) *Rural-Urban Migration in China and Indonesia*. Cheltenham, UK: Edward Elgar Publishing.
- Research Institute of New First-tier City (2020) Ranking of cities' business attractiveness in China 2020. Available at: <https://www.yicai.com/news/100648666.html> (accessed 8 December 2021).
- The Paper (2020). After Wuhan, many places in Hubei were blocked. Available at: https://mp.weixin.qq.com/s/vTztWSZk8QUz4vfoA_Vv2Q (accessed 8 December 2021).
- UN News (2020) Uncertain future for migrant workers, in a post-pandemic world. Available at: <https://news.un.org/en/story/2020/09/1072562> (accessed 8 December 2021).
- Wang Y, Wang F, Zhang Y, et al. (2019) Delineating urbanization “source-sink” regions in China: evidence from mobile app data. *Cities* 86: 167–177.
- Wako H, Takahashi N and Matsui K (2001) Spatial characteristics of women's tourism activities in terms of life-stage in northern residential area of Sendai city, Japan. *The New Geography* 49(3): 12–33.
- Wei L, Luo Y, Wang M, et al. (2020) Multiscale identification of urban functional polycentricity for planning implications: an integrated approach using geo-big transport data and complex network modeling. *Habitat International* 97(2020): 102134.
- Wei Y, Song W, Xiu C, et al. (2018) The rich-club phenomenon of China's population flow network during the country's spring festival. *Applied Geography* 96: 77–85.
- Wong DFK, Li CY and Song HX (2007) Rural migrant workers in urban China: living a marginalised life. *International Journal of Social Welfare* 16: 32–40.
- Wu YM, Wang L, Fan LH, et al. (2020) Comparison of the spatiotemporal mobility patterns among typical subgroups of the actual population with mobile phone data: a case study of Beijing. *Cities* 100(2020): 102670.
- Zhejiang News (2020) New development pattern: how to deeply understand the status and role of the yangtze river delta. Available at: https://zjnews.zjol.com.cn/zjnews/zjxw/202009/t20200928_12324506.shtml (accessed 8 December 2021).
- Zhou C, Li M, Zhang G, et al. (2021) Spatiotemporal characteristics and determinants of internal migrant population distribution in China from the perspective of urban agglomerations. *PLoS One* 16(2): e0246960.

Author Biographies

Dr. Kai Liu is a post-doctoral fellow at Peking University Shenzhen Graduate School. He received his PhD from the University of Tsukuba. His research focuses on urban planning and time geography using geo-tagged big data.

Prof. Pengjun Zhao is a Full Professor and the Dean of School of Urban Planning and Design of Peking University. He is the Editor-in-Chief of *Cities*. His research mostly focuses on urban and transport planning, urban modelling.

Mr. Dan Wan works at Peking University. His research interests include integrated transport-land use model and urban traffic network.

Dr. Xiaodong Hai works at Smart Step Digital Technology Co., Ltd. She is a post-doctoral fellow from School of Economics in Peking University. She is an expert in combined innovative practice

for telecom operators and multi-source integrated big data. Her studies focus on regional economy and urban development using mobile signaling data.

Dr. Zhangyuan He is a post-doctoral fellow at Peking University. He received his PhD from Chair of Maritime Business and Logistics at the Universität Bremen. His research interests include sustainable urban logistics, logistics network optimization, and transport geography.

Dr. Qiyang Liu is a post-doctoral fellow at Peking University. He received his PhD from Institute for Transport Studies at the University of Leeds. His research interests include transport policy in the Global South, policy transfer, and transport-related social inequities.

Miss Yonghui Qu works at Smart Step Digital Technology Co., Ltd as a data analyst. She specializes in database construction, data mining and machine learning based on mobile signaling.

Miss Xue Zhang works at Smart Step Digital Technology Co., Ltd as an industry research experts. She concentrates on innovation practice promotion of mobile signaling data.

Miss Kaixi Li works at Smart Step Digital Technology Co., Ltd as an industry research experts. She specializes in database construction and statistical analysis based on mobile signaling.

Dr. Ling Yu is a post-doctoral fellow at Peking University Shenzhen Graduate School. Her research focuses on tourism geography based on human mobility analysis.