1. *Method of identifying rural migrant workers from mobile phone big data*

First, some specific rules for identifying the RMWs were set before the machine learning process. The typical samples which have a specific kind of features and constraints, such as staying behaviour and periodic rural-urban migration before/after the Spring Festival, lower income, difference of hometown and workplace were carefully trained (cf. Figure 2). Additionally, a college student is not regarded as a RMW but must be firstly excluded before the identification of RMWs began, as one of the premise constraints. Second, a neural network model was used to training the fittest model to judge and identify a RMW which better meet the following conditions: such as staying in his/her hometown during the Spring Festival but working in another city, specific user attributes, typical urban-rural migration pattern, etc. Finally, the data producer (a third party) employed a method involving spatiotemporal constraints and human behaviour identification and a range of machine learning and data mining technology to handle the raw data (including a desensitization and encryption process), while setting specific rules to distinguish the RMWs from the total population. Besides the patent of the method of identifying RMWs using mobile signaling big data, the data producer had also used the total population extrapolation algorithm, spatiotemporal human mobility and labour migration models, and the return-to-city/return-to-work model to derive the data available for this study.

2. *Newest grade classification of China's 367 cities*

Based on a comprehensive weighted indicator calculated using five dimensions: business resource aggregation, city centrality, urban resident vitality, lifestyle diversity, and future plasticity, the Research Institute of New First-Tier City (RINFC), Shanghai Media Group provided the newest grade classification of China's 337 major cities (cf. RINFC, 2020). Thirty other outlying cities, which are mainly located in frontier provinces and ethnic minority autonomous regions, are also considered in this study and are regarded as fifth-tier cities. Figure S1 illustrates, including four first-tier cities, 15 new first-tier cities, 30 second-tier cities, 70 third-tier cities, 90 fourth-tier cities, and 158 fifth-tier cities.

Using this grade classification makes it possible to identify how the patterns of RMWs' return to work in the three UAs vary with different city levels of workplace and hometown. To describe and interpret our results better, we also present a map of China's seven geographical divisions in Figure S1(b): North, East, Central, South, Southwestern, Northwestern, and Northeastern China. As an exceptional case, Hubei Province is specially marked.
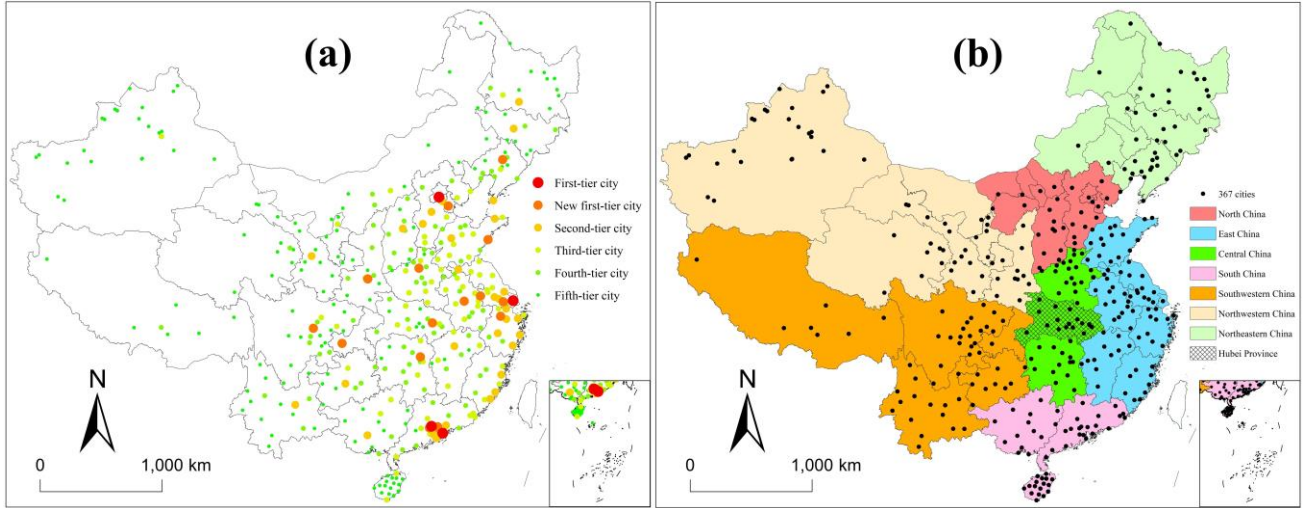
Fig. S1. Grade classifications and geographical divisions of China's 367 cities. (a) Six-level grade classification of China's 367 cities. (b) China's seven geographical divisions. Note: There is no available data for Taiwan, Hong Kong and Macao. Cities in these three areas are not shown.

3. Time-series patterns of labour recovery in 75 cities of three urban agglomerations

This study investigates the time-series patterns of labour recovery (i.e., year-over-year ratios of RMW numbers) in 75 cities of three UAs from February 2020 to October 2020 (cf. Figures 4d, 4e, and 4f), Here, the meaning of labour recovery does not include working remotely or working from home. In fact, after February 20, 2020, except Hubei Province, almost all parts of China had fully launched the policy of work-production resumption (cf. Zhihu, 2021). By May 22, 2020, the numbers of newly confirmed COVID cases had dropped zero in China. In October 2020, the number of tourists during the Chinese National Day Holiday was basically the same as in 2019 (Zhihu, 2021). That is to say, the adverse effects of the lockdown on the labour recovery have been quickly curbed in China's context, so this paper emphatically needs to reveal other hidden factors (e.g., RMWs' hometowns, workplace locations, and human attributes) related to RMWs' return to work along a long-term timeline.

In particularly, in the case of labour recover in C1 cities (cf. Section 3.2) by October 2020, YRD's Shanghai, Nanjing, Suzhou, and Hangzhou had the best situations (around 98.6-100%), following by PRD's Guangzhou, Shenzhen, Foshan, and Dongguan (around 95.5-99.3%). The results also indicated that tourist cities (e.g., Suzhou, Hangzhou, Huangshan, and Qingyuan) and hub cities (e.g., Lianyungang and Xuzhou) tended to have better labour recoveries. However, the labour recovery of BTH was not so good, which might be due to the strict control of the capital. Some edge cities or satellite cities (e.g., Handan, Meizhou, Shantou) might have experienced insufficient

labour release (because of coronavirus-induced retention) towards the central cities of three UAs, and RMWs coming from these places might have found it easier to change their workplace choice around February 2020 (cf. Figures 4d, 4e, and 4f).