Automation impacts on China's polarized job market

Automation impacts on China's polarized job market

Automation Impacts on China's Polarized Job Market

Haohui "Caron" Chen[1,2*†], Xun Li[3,4*†], Morgan Frank[5,6], Xiaozhen Qin[3,4], Weipan Xu[3,4], Manuel Cebrian[5,7], Iyad Rahwan[5,7*]

Affiliations:

[1]Data61, Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia.

[2]Faculty of Information Technology, Monash University, Australia.

[3]Sun Yat-sen University, China.

[4]China Regional Coordinated Development and Rural Construction Institute, Sun Yat-sen University, China

[5]Max Planck Institute for Human Development, Germany.

[6]Institute for Data, Systems, and Society (IDSS), Massachusetts Institute of Technology, USA.

[7]Media Laboratory, Massachusetts Institute of Technology, USA.

*Correspondence to: Xun Li (lixun@mail.sysu.edu.cn), Haohui "Caron" Chen (caron.chen@data61.csiro.au), and Iyad Rahwan (rahwan@mpib-berlin.mpg.de).

† These authors contributed equally to this work.

Abstract: When facing automation threats, a large Chinese city worker might not be as lucky as a worker in a large U.S. city. Empirical studies found that large U.S. cities exhibit resilience to automation impacts because of the increased occupational and skill specialization. However, in this study, we observe polarized responses in large Chinese cities to automation impacts. The polarization might be attributed to the elaborate master planning of the central government. Cities are assigned with different industrial goals to achieve globally optimal economic success and a fast-growing economy. By dividing Chinese cities into two groups based on their administrative levels and premium resources allocated by the central government, we find that Chinese cities follow two distinct industrial development trajectories: one trajectory involves government support leading to a diversified industrial structure and, thus, a diversified job market, and in the other trajectory an absence of government support leads to specialty cities and, therefore, a specialized job market. By estimating the impact of automation on a polarized job market, we observe a Simpson's paradox: overall, there is not a statistically significant relationship between city size and resilience to automation, but when we disaggregate by the level of job diversity, we find larger cities with a diversified job market exhibit greater resilience, while larger cities with specialized job markets are more susceptible. These findings inform policymakers to deploy appropriate policies to mitigate polarized automation impacts.

Keywords: automation impacts, polarized job market, diversity, occupation space, China

1 Introduction

Advances in artificial intelligence (AI) and robotics technologies have revived the concerns of technological unemployment. Frey and Osborne (1) estimated that more than 47% of U.S. jobs were at high risk of computerization, and an alternative OECD study (2) found a more modest 9% of U.S. jobs were at stake. The former research estimated automation effects based on the occupational level, while the latter based on the task level. Although these two studies' results differ significantly, they have led to a broad discussion of automation's impact on the labor market and economic development. Many scholars have estimated different countries' automation rates (3, 4). Usually, developing countries like China, Vietnam, and Malaysia are more susceptible to automation (5,6). World Bank claimed that 1.8 billion jobs or roughly two-thirds of the labor force in developing countries would be replaced by automation (5). According to Frey et al. (7), specific to China, it's expected that accounted for 77% of China's employment is at risk, approximately 550 million jobs (8, 9). The impacts on China's job market expect to be devastating. Even though these results are highly disputed, they foresee unavoidable effects on the upcoming decades' global job market.

The varying automation impacts among countries are directly related to the national industrial structure (10). The developed countries such as the United States and European countries are more engaged in the service, finance, research, and development sectors. In contrast, developing countries have more manufacturing occupations, which are more sensitive to automation. For example, China has become the world's factory, receiving manufacturing orders worldwide; simultaneously, increasing domestic demand has elevated its manufacturing capacities. However, jobs in the manufacturing sector are among the most susceptible to automation (1). In addition, although the urbanization rate has exceeded 60%, there is still a significant labor force in the farming sector, which is also susceptible to automation. Both aggravate the overall impact on China's job market. Furthermore, insufficient and unbalanced development is the biggest obstacle to China (11, 12). Thus, the effects on the job market would undoubtedly exacerbate inequality, and the resulting destabilization would be a significant problem for not just China but also the world.

In recent years, China has begun to set off a wave of automation, and the impact of automation on the labor market has also gradually appeared (13). A few existing studies have focused on the effects of robots on China's labor market. For example, Giuntella and Wang used the IFR data to find a negative impact of robot exposure on Chinese workers' employment and wages at the prefectural city level (14). Although these studies have paid attention to the effects of automation on the labor market, most of them didn't comprehensively quantify the impact. They only focus on robots, a kind of automation technology. Therefore, this paper uses a more general measure of automation and considers the potential effects of automation on the labor market at the city level in China.

Similar to the different automation rates among countries, the impact of automation also varies from city to city. Frank et al. (10) found that large U.S. cities are more resilient to the effects of automation because increasing numbers of specialized occupations and skills that are resilient to automation were observed in larger cities. Under the market economy, cities in the United States develop organically. As the size of cities increases, so does the diversity of cities. But, China experienced a long period of top-down planned economic development before the reform and opening up, which directly affected Chinese cities' development trajectories. Under the background of the planned economy, whether the city can obtain more development resources determines the city's subsequent development path. The underlying industrial structure might constitute a distinct job market and, as a result, cause Chinese cities' polarized responses to automation impacts. Therefore, Chinese cities' reflections on automation may show different characteristics from those in the United States. Cities with additional access to development resources have higher resistance to automation due to more diverse industrial structures. Therefore, this paper hopes to explore the influences of the urban development path on the corresponding automation in China based on the automation rate analysis at the city level.

Based on the results of Frey and Osborne (1), we estimate job impact rates for 102 cities in China by mapping American occupation to Chinese occupation. We found polarized responses along with the growing size: some large cities show resilience to automation, while others are also estimated to be susceptible to automation (Fig. 1). Most of these vulnerable cities usually dominate in a specialized industrial structure. These "specialty cities" could be explained as the legacy of China's planned economy that the central government assigned specific development targets to these particular regions (15). After the economic reform, regional governments gain greater autonomy in choosing industrial strategies. For example, between 2016 and 2017, the central government approved 403 local government applications to initiate "specialty towns"—focused on one particular industry—and aims to have 1,000 of them by 2020 (16). However, given the regional endowments and the central government's preferences, China's cities have grown in size along two trajectories concerning industrial structure. Some regions adapted a more diversified industrial structure, similar to U.S. cities (17). Still, some adapted a path-dependent development trajectory (18) to become "specialty cities." Given industrial polarization, China's cities develop a polarized job market. In the context of automation, a polarized labor market may cause significant threats and inequality to non-advantaged cities. Therefore, this study aims to understand the emerging impacts of automation technologies on China's polarized job market.

2 Methods and Data

2.1 Materials and data sources

There are four types of data sources used in the paper. The first one is the 702 occupations' automation rate of the United States from Frey and Osborne (1). It's one of the primary data to estimate the China cities' automation rate. The second one is the Sixth National Population Census in 2010 (hereafter referred to as the 2010 Census). The Census data is the only data that identifies the employment distribution of 413 sub-sectional occupations and 95 industrial subsectors for now in China. This paper needs precisely the employment data with 413 subsections of occupation to estimate the automation impact. However, only a few cities have published this kind of detailed data. Therefore, we only obtained 102 cities' occupation data, and the research is also based on 102 cities. The third is the data used to divide the cities into premium and non-premium cities, that is, the number of universities funded by the projects "985" and "211" and the bullet trains. By 2010, 39 and 109 universities have been supported by projects "985" and "211," respectively (some are funded by both concurrently). The data are available from the official website of the Ministry of Education of China: http://www.moe.gov.cn/srcsite/A03/moe_634/. The daily operating frequency of bullet trains in each city can be found at the online bullet train ticket office: http://www.gaotie.cn/. Both university and bullet train statistics are available in Table. S1. The last one is some data at the city level from the China City Statistical Yearbook 2010, including vocational teachers and fixed asset investments.

2.2 Estimation of job impacts in China

Frey and Osborne (1) estimated the automation possibility at the occupation level and studied the substitution impact of automation on the United States' labor market. Frey and Osborne obtained 70 occupations as the training samples by asking experts whether the technology would replace the occupations. 0 represents the occupation would not be replaced while 1 illustrates would. The training samples were used to model the relationship between the occupation attributes and whether it will be automated or not. Based on the model, 702 occupations' automation rate has been predicted. This research has enabled scholars to estimate the automation substitution effect in other countries via mapping the American occupational codes to domestic codes (3, 4). Although this method of occupation code mapping is subject to error, it makes it possible to conduct cross-country comparative studies of the effects of automation.

Based on Frey and Osborne's (1) results, this study also uses a matching method to estimate the Chinese occupation's automation rate. To utilize the estimated job impact rate of (1), we build a correspondence table between China's Grand Classification of Occupations (GCO) and the U.S.'s Standard Occupational Classification (SOC). There are 8 chapters, 66 sections, and 413 subsections of occupations in the GCO. The GCO's 413 sub-sectional occupations are mapped to one or multiple occupations from 702 SOC 6-digits occupations based on the titles and corresponding descriptions. The matching process is as follows.

(1) Three paid students ($N = 3$) with backgrounds in human geography mapped GCO subsectional occupations ($J = 413$) to as many SOC 6-digits occupations ($K = 702$) as appropriate based on titles and corresponding descriptions, respectively. This step generates an array of mappings for individual GCO occupations: $occupation_{GCO}(j) \in R^{702}$, $j \in J$, where R is a vector with element $occupation_{SOC}(k) \in \{0, 1, 2, 3\}, k \in K$. The value of $occupation_{SOC}$ indicates the number of occurrences of such an occupation as mapped by students.

(2) If a GCO occupation has all candidates $occupation_{SOC} \leq 1$, the authors decided on the SOC to which that occupation should be mapped, the corresponding $occupation_{SOC} \leftarrow 1$, and the rest $occupation_{SOC} \leftarrow 0$.

(3) The commonly matched SOC occupations ($occupation_{SOC} \geq 2$) are kept and the corresponding value is changed $occupation_{SOC} \leftarrow 1$, and the rest $occupation_{SOC} \leftarrow 0$. The final output of the correspondence table is a matrix $R^{413\times702}$, in which each element is $occupation_{SOC} \in \{0,1\}$.

We cannot find a mapping for occupations in the "leader of political or state-owned entities" SOC sector; therefore, we assign a zero impact rate for them given the special nature of this type of occupation. Subsequently, we compute the mean impact rate of all $occupation_{SOC} = 1$ and assign the mean value as the computerization probability to the corresponding GCO occupations.

After obtaining the automation rate of 413 Chinese occupations, we estimated the impacts of automation on China's job market at the city level using [1] as adapted from (10).

$$E_m = \frac{\sum_{j\in Jobs_m} p_{auto}(j)\cdot f_m(j)}{\sum_{j\in Jobs} f_m(j)} \quad [1]$$

E_m is the expected job impact rate for city m, $f_m(j)$ denotes the number of workers in city m with job j, and $Jobs_m$ is the set of job types in city m. $p_{auto}(j)$, the probability of computerization for job j is adapted from (1).

2.3 The division of the city types in China

In this study, we adopt the prefecture-level cities as the analysis units. Unlike the cities developed under the United States market economy, Chinese cities' development path is not entirely free growth under the planned economy's influence. As the hypothesis claimed before, we think cities' different development paths will lead to various automation effects on the city labor market. Based on the characteristics of urban development, we try to categorize Chinese cities in two ways.

In the first method, we group cities into two groups by their administrative levels. We call sub-provincial and above cities elite cities and the general prefecture-level cities non-elite cities. In this study, there are a total of 19 elite cities and 83 non-elite cities. The central government successively imposed the administrative levels of each city after the 1950s until the 1990s. Therefore, this long-lasting administrative division might also unveil the master planning of the government over the past decades and affect the cities' development.

In the context of the planned economy, the richness of the development resources available to cities is different, such as the research facilities and transport infrastructure, directly affecting cities' subsequent development trajectories. Therefore, in the second method, we divided cities into premium and non-premium cities according to the richness of cities' development resources. To develop innovation capabilities and international competitiveness, China launched two national education projects—"985" and "211"—in the 1990s to enable a limited number of universities to become world-class research facilities (19). As a result, these premium research facilities have significantly improved the innovation competitiveness of corresponding cities (20). By 2010, 113 universities have been funded by the "985" and "211" projects. Another significant premium resource is the bullet train railway network invested in and managed by the central government. The allocation of bullet train stations and the operating frequency represent the importance of a corresponding city as the regional service center, such as a freight forwarder. Moreover, bullet trains also exhibited strong effects on the overall competitiveness of a city (21). In this study, we use the bullet train's daily operating frequency as a proxy for infrastructure investments by the central government). Using k-means clustering ($k = 2$) on those two features, we got two groups of cities, and we name them premium ($N_{premium} = 20$) and non-premium cities ($N_{non-prem} = 82$). Among them, 16 elite and 4 non-elite cities appear on the list of premium cities (see Table. S1 for the complete list of both division systems).

It is not a complete surprise that the majority of premium cities are elite cities, like direct-controlled municipalities and sub-provincial level municipalities. These cities enjoy administrative powers over their peers. One of the exclusive advantages is direct communications with the central government; therefore, state-owned and multinational companies prefer to reside in them to reduce communication costs and to stay informed of volatile policies and regulations. Fixed asset investments can be a good proxy for representing the preferences of those companies. Fig. S1 shows that fixed asset investments grow linearly with city size in premium cities and cities of higher administrative power. In contrast, they grow sub-linearly in non-premium cities and cities of lower executive power.

2.4 Diversity and occupation space

It has been proved that the effect of automation on the city's labor market is related to the industrial structure, calculated by the specialization or diversity, in the empirical study of the United States (10). To explore the relationship between the Chinese cities' automation rate and the industrial structure, we compute the industry and job diversity using normalized Shannon entropy (22), as follows. Normalized Shannon entropy has been widely used in urban science (10, 23).

$$H_{job}(m) = -\sum_{j\in Jobs_m}\left[p_m(j)\times log\left(p_m(j)\right)\right]/log(Jobs_m) \quad [2]$$

$$H_{industry}(m) = -\sum_{k\in Industries_m}[p_m(k)\times log(p_m(k))]/log(Industries_m) \quad [3]$$

$p_m(j)$ denotes the probability of a worker in city m having job j, $p_m(k)$ denotes the probability of a worker in city m working in industry k, and $Industries_m$ is the set of industry sectors in city m. In the computation of job diversity $H_{job}(m)$, we use $Jobs_m \in 413$ $sectional\ jobs$, and in the computation of industry diversity $H_{industry}(m)$, we use $Industries_m \in 95$ $industry\ sectors$.

Besides the Shannon entropy calculated above, this study also used the complex network method to quantify urban occupational structure, widely used in urban research on the export network, industry network, and innovation network (24). The study further analyzes the impact of automation on the urban labor market based on the occupation network. We adapt Hidalgo et al.'s (24) methodology of product space to construct the occupation space. This method builds the product space by calculating the probability that the two products have a comparative advantage in the same country's export structure, which is the proximity between them. Similarly, the occupation space is constructed by calculating the probability that two occupations have a comparative advantage in the city.

Firstly, we calculate the occupations' revealed comparative advantage. $RCA_{m,j}$ measures, for city m, occupation j's relative level to the national average. $x_{m,j}$ is the total number of jobs of occupation j in city m. When $RCA_{m,j} > 1$, city m has more jobs of occupation j as a share of its total numbers of jobs than the national average, and vice versa. An occupation with $RCA_{m,j} > 1$ is considered one of the city's advantaged occupations.

$$RCA_{m,j} = (x_{m,j}/Jobs_m)/(Jobs_m/\sum_m Jobs_m) \quad [4]$$

Secondly, we calculate the proximity of each occupation pairs. The proximity between two occupations i and j, denoted as $\phi_{i,j}$. The minimum of the pairwise conditional probabilities of a city accommodates an occupation given that it also accommodates another. Closer proximity indicates a higher correlation between two occupations and closer spacing of the two occupation nodes in the occupation space (Fig. 4a). The value of $\phi_{i,j}$ is between 0 and 1.

$$\phi_{i,j} = min\left\{Pr\left(RCA_{m,i}\geq 1|RCA_{m,j}\geq 1\right), Pr\left(RCA_{m,j}\geq 1|RCA_{m,i}\geq 1\right)\right\} \quad [5]$$

Finally, we construct the occupation space. First, we build a 413 * 413 occupation proximity matrix. Second, we generate a maximum spanning tree (MST) based on the proximity matrix to generate the network's skeleton and complement it with additional proximity links whose $\phi_{i,j} > 0.66$. Lower proximity values led to crowded network representations, while higher values resulted in a sparse network. According to Hidalgo et al. (24) rule of thumb for the proximity threshold should lead to a network of an average degree equals 4, i.e., the number of links is twice one of the nodes. In this case, we achieve such a network representation threshold. For better visualization, we use a force spring layout to visualize the network.

3 Result

3.1 The impacts of automation

Based on the Frey and Osborne (1) results and the mapping of China and the US occupations, we estimated the automation rate of China and 102 cities. On average, 79% of jobs are expected to be at high risk in China, which is close to the estimation of (*3*) (see Table. S1 for results for all 102 cities). Well-known large cities, such as Beijing, Shanghai, Guangzhou, and Shenzhen, exhibit resilience to automation technologies, with expected impact rates (the percentage of employment at high risk of being lost due to automation) of 64%, 67%, 69%, and 72%, respectively, whereas other large cities such as Zhumadian and Nanyang are the opposite, both with job impacts as high as 83% (Fig. 1). Unlike the findings of (10), the expected job impact rate doesn't merely change linearly with city size. It depends on the type of cities, elite cities and non-elite cities, premium cities, and non-premium cities defined in part 2.3.

We build regression models between city size and the job impact rate of cities. The estimations of the models' results are presented in the inset table. In the model of whole samples, the relationship between the job impact rate and the city size doesn't pass the significant test ($p_{value} = 0.86$), but does within groups in Fig. 1b and Fig. 1c. The x-axis is the city size, represented by the number of employees, and the y-axis is the expected job impact. The scatter diagrams show that both large premium and elite cities exhibit resilience to automation impacts relative to the others. Here, we observe a Simpson's paradox (25): overall, there is not a statistically significant relationship between city size and resilience to automation, but when we divided these cities into two groups, we find that larger cities on the advantaged side (premium and elite) have greater resilience to automation impacts while the larger cities on the non-advantaged side are more susceptible. In this paradox, small non-advantaged cities (non-premium and non-elite) are resilient to automation impacts, which is the opposite of the findings for the United States (10). We also compare city size effects between Chinese and U.S. cities (see Fig. S2). Even though the overall impact in China is significantly higher than in the United States, Chinese cities on the advantaged sides exhibit more potent size effects than U.S. cities ($|\beta_{premium}| > |\beta_{elite}| > |\beta_{US}|$). Therefore, cities enjoying more premium resources or higher administrative power help mitigate the automation impacts to a great extent.

Susceptible large cities have long been regarded as "specialty cities," which are either specialized in farming (e.g., Nanyang), mining (e.g., Pingdingshan), or manufacturing (e.g., Taizhou). In contrast, resilient large cities are either innovation hubs (e.g., Beijing) or financial (e.g., Shanghai), or regional services centers (e.g., Guangzhou). The former industrial sectors tend to employ higher computerization risks, whereas the opposite is true for the latter. Therefore, it is worth studying the relationships between the cities' industrial structure and the automation impacts.

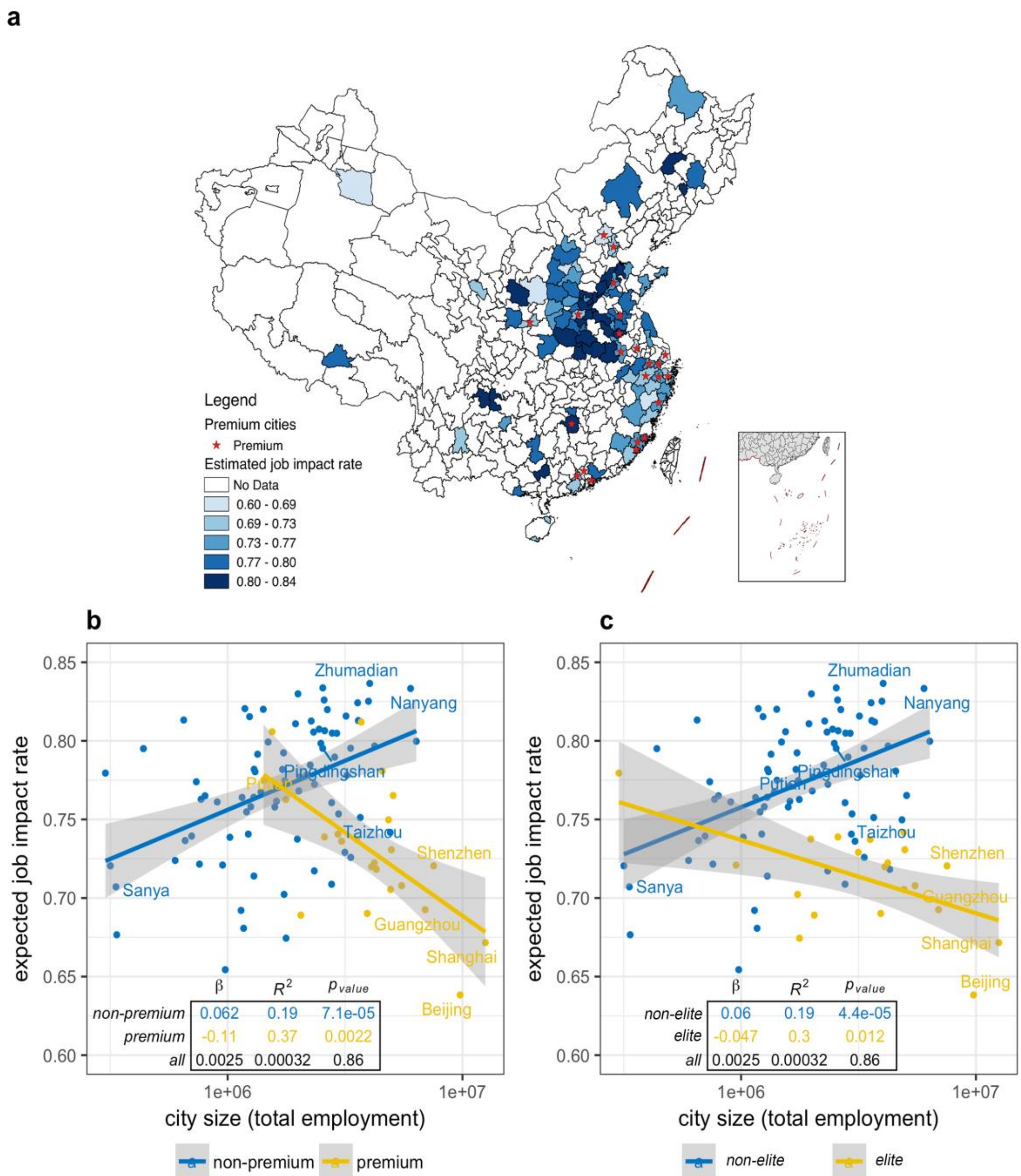


Fig. 1. Expected job impact rate over city size. (a) Geographical distributions of expected job impact from automation across China's cities. Premium cities are marked with red stars. (b) and (c) Expected job impact rate over city size. In (b), cities are grouped using k-means clustering according to the educational resources and transport infrastructures funded by the central government. In (c), cities are grouped based on their administrative levels.

3.2 Polarized job market

To understand China's industrial structure, we perform a principal component analysis (PCA) (26) on the employment distributions across 20 industrial sectors (see Fig. 2). PCA has been used in urban science (27, 28). The PCA result shows that two leading principal components (PCs) accumulate more than 85% of the total variances. Therefore, we can characterize China's industrial structure by analyzing two PCs' industrial employment compositions and cities' PC scores. A group of vectors consisting of tertiary industries pointing in the 9 o'clock direction on PC1 corresponds to the coexistence of the tertiary industries in the same set of cities. However, farming and mining both have minor contributions regarding PC1 but significant regarding PC2, indicating that both industries tend to co-locate but dislocate to the tertiary industries. Manufacturing contributes to both PCs but affects inversely against farming and mining regarding PC2. Thus, we can interpret two PCs as a servicing index and a natural resources index from the composition of PCs, respectively. Cities' positions on this plot unveil their industrial structure. For example, Beijing and Shanghai have significant servicing but minor natural resources indexing scores, addressing their advantages in servicing roles in the Chinese economy. In this plot, two types of large cities emerge. One type is servicing centers, and the other type is natural resources centers, which also correspond to two types of large cities that face polarized automation impacts. Interestingly, suppose we set one of the smallest cities in this study, Sanya, as the starting point. In that case, we can observe that cities tend to follow two trajectories toward those two extreme industrial structures, along with increasing size (dashed vectors in Fig. 2).

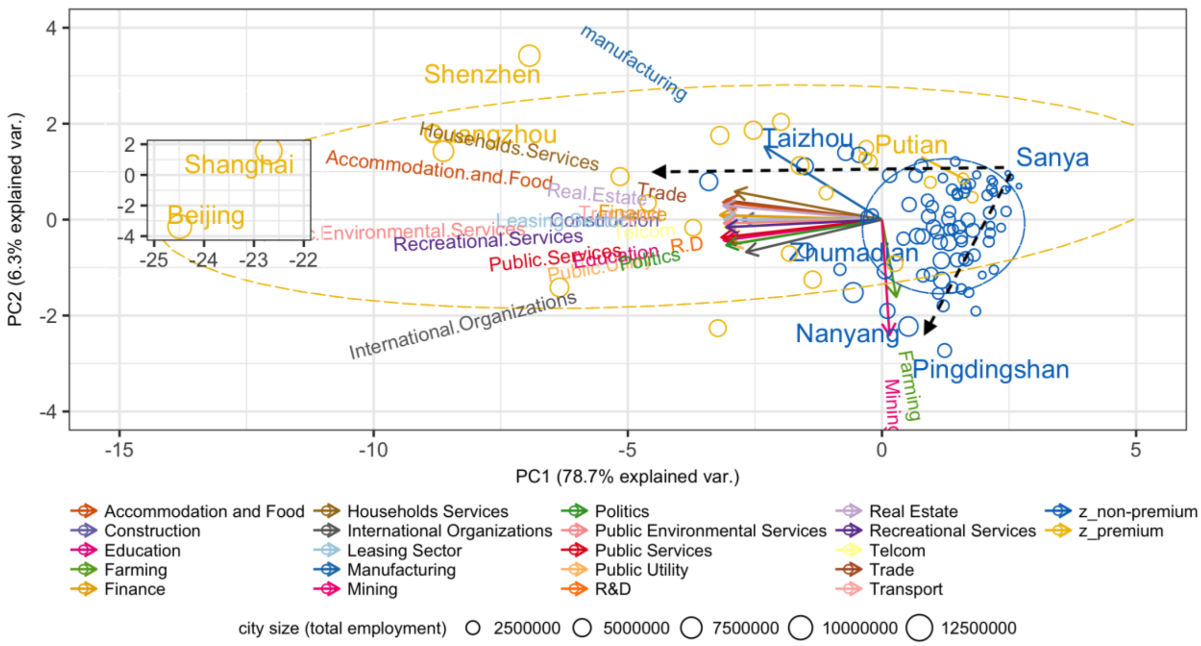
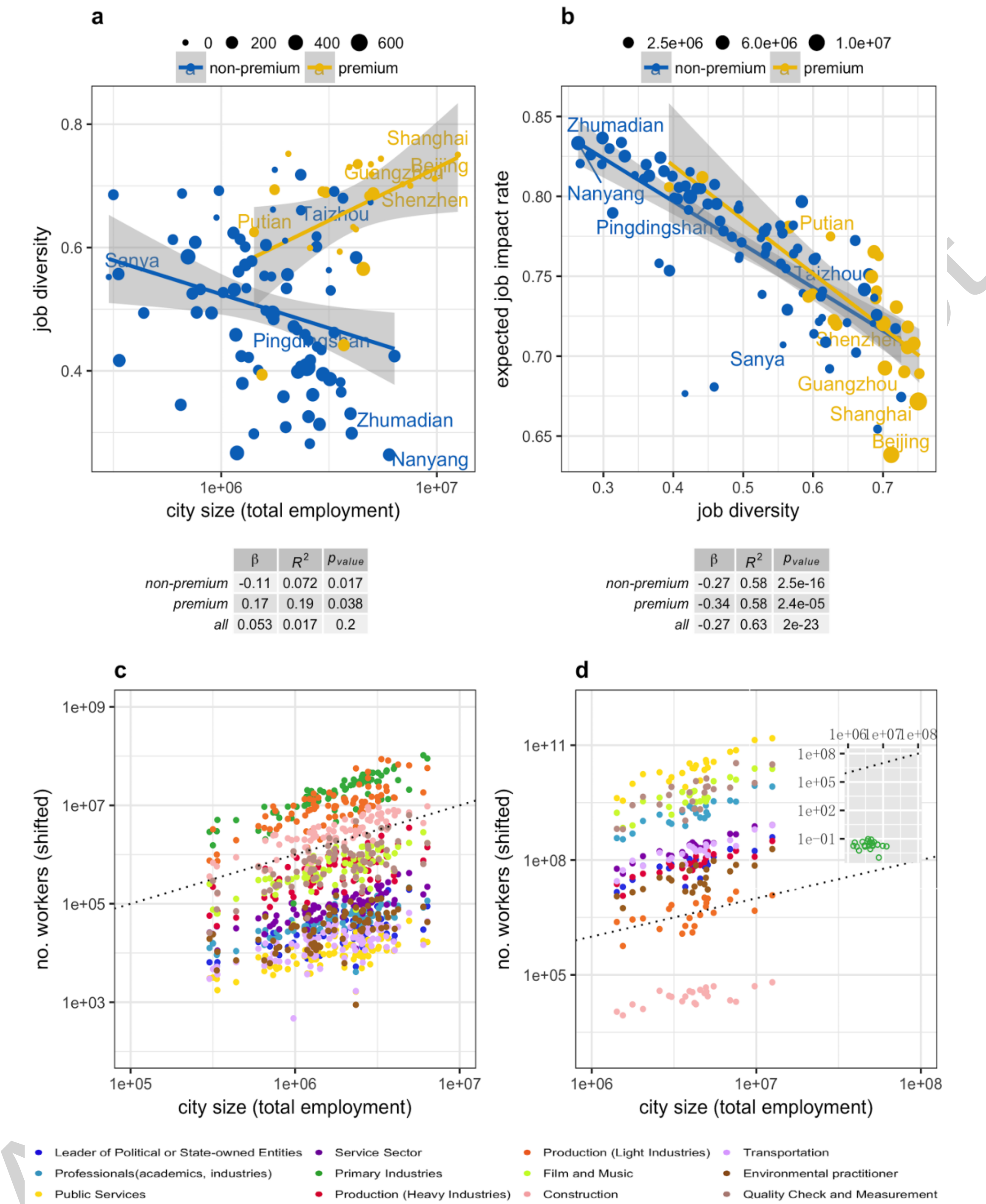


Fig. 2. China's industrial structure obtained from PCA on the employment distributions across 20 industry sectors. The solid vectors represent the coefficients of industrial employments on the PCs, the nodes represent cities' corresponding PC scores, and the dashed vectors starting from Sanya indicate two arbitrary directions that illustrate how small cities transform their industrial structure along with increasing size. The size of the nodes is proportional to the city size. Two leading PCs explain 78.7% and 6.3% of the overall variances, respectively. Two eclipses show the 95% confidence interval (CI) of PC scores of corresponding cities.

In Fig. 2, along with the two distinct industrial trajectories that we hypothesized, premium cities tend to develop a more diversified industrial structure along with increasing size, which is similar to U.S. cities (17). In contrast, non-premium cities are adapted to be specialized in either farming and mining or manufacturing. Even though we cannot investigate how cities are assigned with specific missions given the lack of historical planning materials, we can infer from the previous finding that the central government has put in force two distinct industrial strategies over these two types of cities. Given that China started its reform as an emerging country, to achieve a fast-growing economy, limited premium resources can only be deployed in a few trailblazer cities and become innovation hubs and financial and regional services centers. The other cities might need to specialize in a few industries to benefit from the scale economies (29). Moreover, we also find that non-advantaged cities tend to develop a more diversified industrial structure when geographically close to elite cities (Fig. S3). The spatial proximity effect might aggravate polarization because both innovations and diversified job markets exist in elite cities (see Fig. 3a). Skilled and well-educated workers tend to migrate to these cities to earn higher wages (30). As a result, non-advantaged cities farther from elite cities have lower chances of attracting those workers. Therefore, because "specialty cities" might lack skilled and well-educated workers to develop high-tech industries, they continue to specialize in low-tech industries, such as farming and mining. We confirm the population loss of non-advantaged cities with the growing distance to the closest elite cities (Fig. S4). We use the distance to the nearest elite cities instead of the premium cities because elite cities are more geographically distributed than premium cities (mainly located in eastern China). Thus, non-premium cities in western China could benefit more from their elite neighbors than distant premium cities in the east of China. Moreover, the distance computation involves not just the elite cities in this study but also those beyond this study, given the lack of detailed census records.

Because a diversified industrial structure constitutes a diversified job market and vice versa (strong correlation between both are observed in Fig. S5), polarized job diversities of two groups of cities over city size are also observed (Fig. 3a). There are some large non-premium cities also developing a diversified job market, such as Taizhou. Those cities might have been taking similar industrial trajectories as those of premium cities, given that they are closer than their peers to diversified elite cities. More importantly, we find that job diversity significantly affects expected job impacts (Fig. 3b). Similar to findings in the United States (10), a diversified job market can mitigate automation impacts to a great extent.

By studying occupation growth over city size, we can understand how job market composition affects the expected job impact rate (see Figs. 3c & 3d, and see Fig. S6 for results under the administrative division). The occupation growth patterns of premium and non-premium cities are distinct. Susceptible occupations (e.g., primary industry, production, and construction) grow super-linearly with city size in non-premium cities (that is, the growth rate of city size exceeds the growth rate of a specific occupation). In contrast, they grow either sub-linearly, linearly, or even negatively with city size in premium cities. On the other hand, the most resilient occupations, such as public services, quality checks, and measurements, professionals, and film and music, grow super-linearly with city size in premium cities but on the contrary in the non-premium cities, which are the most distinct regression coefficients between premium and non-premium cities.



	β	R^2	p_{value}
non-premium	-0.11	0.072	0.017
premium	0.17	0.19	0.038
all	0.053	0.017	0.2

	β	R^2	p_{value}
non-premium	-0.27	0.58	2.5e-16
premium	-0.34	0.58	2.4e-05
all	-0.27	0.63	2e-23

Fig. 3. Job diversity, impact rate, and occupation growth between premium and non-premium cities. (a) Job diversity over city size. The node size is proportional to the distance to the closest elite cities. (b) Expected job impact rates over job diversity. The node size is proportional to the city size. Linear regression results are shown in the tables. City size is log-transformed in the model. Panels (c) and (d) show patterns of occupation growth over city size in non-premium and premium cities, respectively. Points are vertically shifted according to linear fit in a log scale, and the black dashed line has a slope of 1 for reference.

11

3.3 Evolution of occupational structure

The product space of Hidalgo et al. (24) offers a compelling illustration that addresses the relationships between world trade products and the roles played by different countries. In the product space, products are connected based on their co-location probability, and the core area consists of sophisticated products such as metal products, machinery, and chemicals. In contrast, the periphery consists of fishing, tropical, and cereal agriculture. Industrialized countries are dominant in exporting products in the core area, whereas non-industrialized ones dominate the periphery. Also, empirically demonstrated was that one developing country could traverse through links of the network to the core area to gain relative advantages with some sophisticated products, which can later constitute the base of further traversing and industrialization. Inspired by the product space, we hypothesize that advantaged cities grow to be hubs of innovation, finance, arts, and services, and become service centers for surrounding non-advantaged cities that develop into “specialty cities” master planning of the central government. Therefore, we construct China’s first occupation space based on the co-location probability of any two occupations (Fig. 4).

Similar to the United States’ occupation space (31), China’s occupation space also has a service sector and professional occupations in the core area and production and farming at the periphery (Fig. 4a). We build an OLS regression to explore the relationship between the occupation’s location in the occupation space and the estimated automation rate. Closeness centrality is calculated to represent the occupation’s location in the space. The result shows a negative correlation between them, and it passes a 1% significant test. That means occupations at the core express resilience to the impacts of automation, whereas occupations at the periphery tend to be susceptible to automation (Fig. 4b).

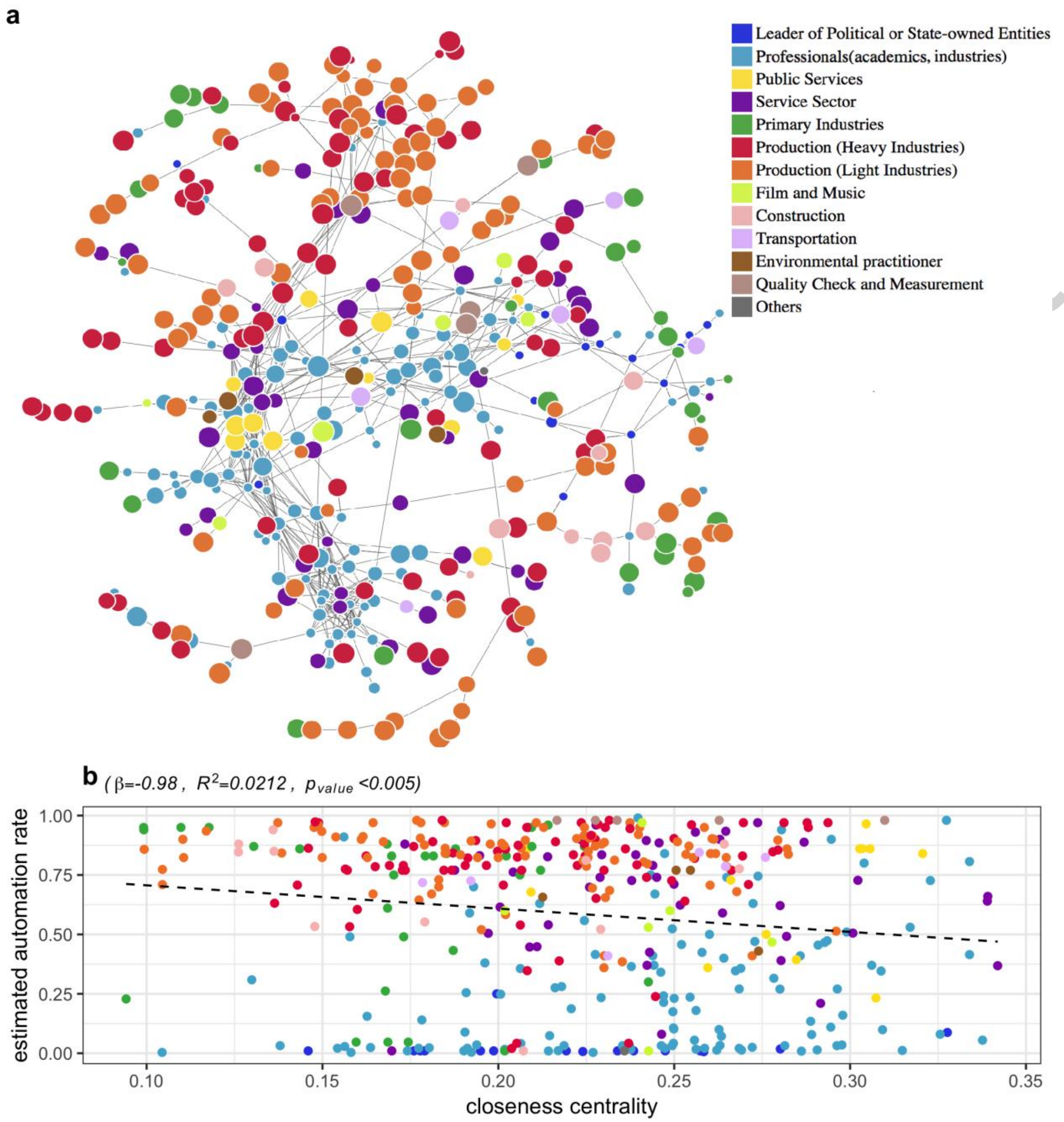


Fig. 4. Occupation space across 102 Chinese cities. Panel (a) shows the occupation space; panel (b) shows the relationship between an occupation's closeness centrality and automation rate. The dashed line indicates the linear regression's best fit, which reports a negative relationship between both.

To illustrate cities' dominance in certain parts of the occupation space, we highlight a few symbolic cities based on their relative advantages. Beijing (Fig. 5b) is dominant in the professional sector at the core, whereas the small premium city of Putian (Fig. 5a), known as the "shoes-making city," is dominant in the production sector at the periphery. The large non-premium city Nanyang (Fig. 5e) is dominant in 28 occupations, mainly in farming and production at the periphery, whereas the small, non-premium city of Sanya (Fig. 5d), known as the tourist city, is dominant in the service sector at the core. To explore the relationship between the city's occupation structure and city size, we estimate the average closeness centrality coefficient and the city size in two groups. The results show that the closeness centrality increases with the increase of the city size in the premium cities but decreases in the non-premium cities (Fig. 5c). Furthermore, it indicates two distinct evolution paths of the job market for premium and non-premium cities: premium cities transit from the periphery to the core, and non-premium cities transit in the opposite direction. This confirms that premium cities grow to be service centers, whereas non-premium cities become "specialty cities." Polarization might achieve economic optima for China as a whole; however, it might cause significant threats to non-advantaged cities and inequality for the entire job market in the context of technological unemployment.

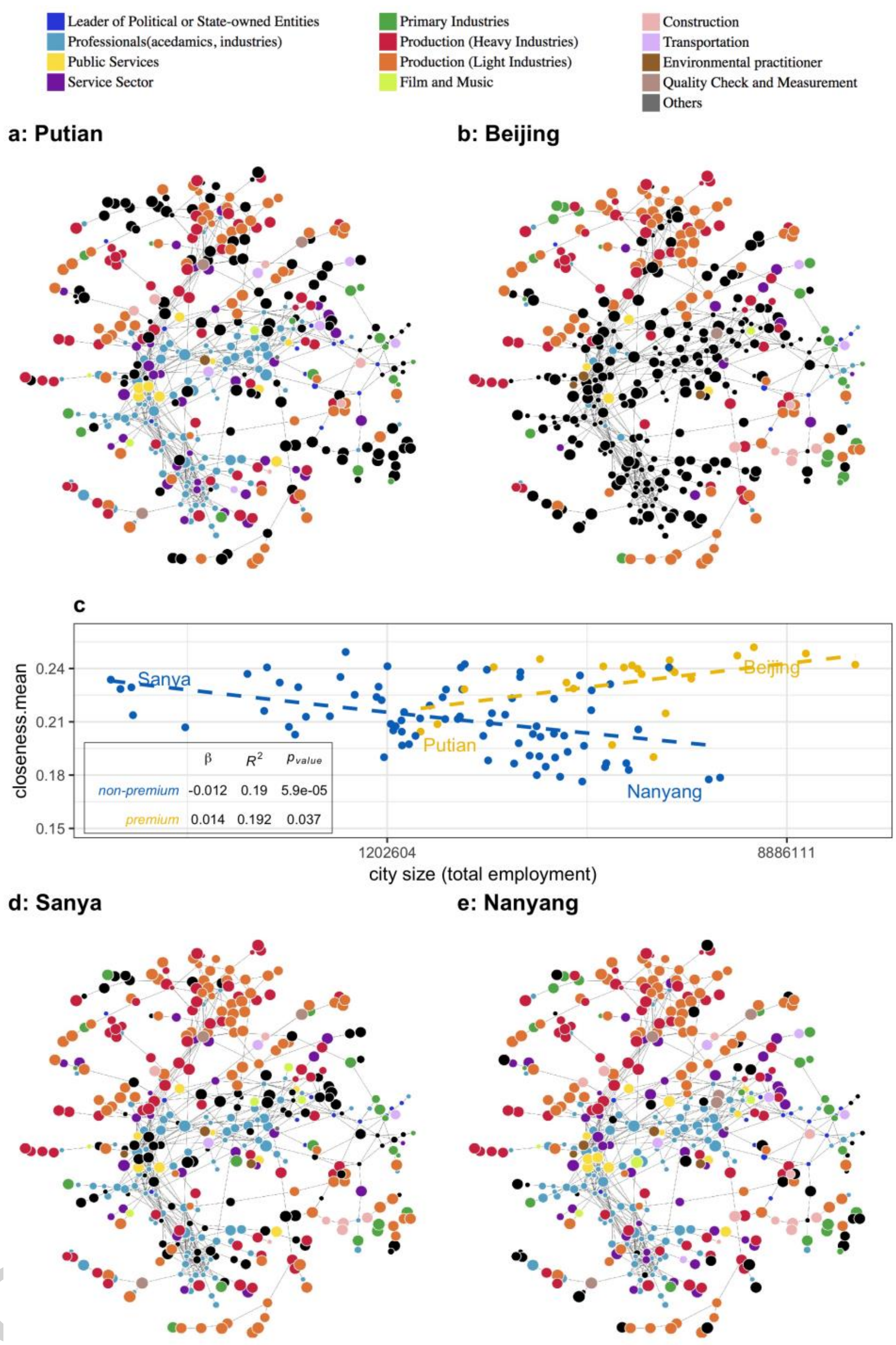


Fig. 5. Evolution of positions in occupation space along with growing size. (a), (b), (d), and (e) represent the positions of Putian, Beijing, Sanya, and Nanyang in the occupation space. (c) represents two distinct evolution paths of premium and non-premium cities in the occupation space. The result of the linear regression model is presented in the inset table, and the city size is log-transformed.

4 Discussion

The automation impacts on the job market have never been as tensional as it is today. The main concern is that the loss of jobs to automation could outstrip the demand created by increasing productivity (32). Therefore, appropriate policies should be deployed to help mitigate rapid changes to the job market. However, few existing studies have understood the emerging differences at the city level for China's most populous country. This study attempts to compute the job impacts for 112 Chinese cities and addresses the polarized responses' increasing city size.

Given a lack of access to historical planning materials, we cannot investigate how missions were assigned to each city. However, by grouping cities based on their administrative level and allocation of premium resources, we find two distinct industrial development trajectories that unveil the central government's master planning regarding the pursuit of the global optima of economic success. China's polarized industrial structure has constituted a polarized job market in which susceptible occupations appear more in non-advantaged cities, and resilient occupations strive in advantaged cities and polarize responses to emerging automation impacts. Moreover, non-advantaged cities can benefit significantly from their neighbor advantaged cities of a diversified industrial structure, which might further exaggerate the polarization of automation impacts. In all, sufficient grounds exist to believe that cities' lack of administrative power or premium resources and distance from advantaged cities could be left behind in the second machine age (32).

Up-skills would be one of the most feasible ways to mitigate automation impacts (33) but involve appropriate allocation of educational resources, especially lifelong learning. We find that China's existing educational resources allocation might not match where they are needed the most (Figs. S7 & S8). The growth of vocational teachers and vocational schools is linearly and sub-linearly correlated with size in non-advantaged cities. However, because large, non-advantaged cities could suffer the most from automation impacts, vocational education facilities should at least super-linearly grow to the city size in non-advantaged cities. Thus, the central government should play an essential role in motivating or subsidizing appropriate policies to support vocational education where it is most needed.

To the best of our knowledge, this study is the first to explore the city division system in China and find the polarization of industrial structure and the job market. The city division system could change urban research in planned economies (e.g., scaling laws (34) and agglomeration economies (35)) given that non-advantaged cities have not been following organic growth, whereas advantaged cities enjoy the fruits of their non-advantaged peers. Empirical studies tend to investigate cities as an entire population and overlook the fundamental differences between organic and planned economies. For example, in (35), China was found to exhibit smaller estimates of urban agglomeration elasticities than in other countries. However, we might expect a different finding when Chinese cities are considered to be different from each other. Two distinct urban regression coefficients were found between eastern and western Europe (36); therefore, we believe that it is worth revisiting China's city size change using this city division.

Indeed, journalists and even expert commentators have successfully portrayed technological unemployment as an unsettling picture of the future of work. However, most overlook the complementarities between automation and labor. As Autor (37) addressed, automation does substitute tasks but also increases productivity and earnings and, as a result, augments a higher demand for labor. In China, future research about the complementarities should take into account the polarized job market.

Due to the data limitation, this study has some deficiencies. Firstly, we can only access the census data for 102 cities made available to the public by the local governments. Even though they are geographically distributed and have different sizes and GDPs, they are just above one-third of the entire population, 295 cities. In this regard, audiences should be cautious in using the conclusions drawn from this limited number of city samples. Secondly, due to the lack of temporal data, this study only uses the cross-section data, and it is challenging to explore the urban development evolution further. Lastly, Similar to (10), many limitations inherent in occupation-level estimations apply to this study as well. Moreover, China lacks statistics on skill and task distributions across occupations, and we cannot perform the task-based approach (2, 38) in estimating the jobs impact in this study. China's job impacts could be significantly weaker than the estimation by (7) and this study. Therefore, policymakers are encouraged to interpret the impact rates as relative values for comparing the impacts between cities.

References

1. Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation?. Technological forecasting and social change, 114, 254-280.

2. Arntz, M., Gregory, T., & Zierahn, U. (2016). The risk of automation for jobs in OECD countries.

3. Fuei, L. K. (2017). Automation, computerization and future employment in Singapore. Journal of Southeast Asian Economies, 388-399.

4. Bowles, J. (2014). The computerisation of European jobs-Who will win and who will lose from the impact of new technology onto old areas of employment?. Bruegel, 2014.

5. Peña-López, I. (2016). World development report 2016: Digital dividends.

6. Chang, J. H., Rynhart, G., & Phu, H. (2016). ASEAN in transformation: How technology is changing jobs and enterprises.

7. Frey, C. B., Osborne, M., Holmes, C., Rahbari, E., Garlick, R., Friedlander, G., ... & Chalif, P. (2016). Technology at work v2. 0: The future is not what it used to be. CityGroup and University of Oxford, 338.

8. Chui, M., Manyika, J., & Miremadi, M. (2016). Where machines could replace humans—and where they can't (yet). McKinsey Quarterly, 30(2), 1-9.

9. Manyika, J., Chui, M., & Miremadi, M. (2017). A future that works: AI, automation, employment, and productivity. McKinsey Global Institute Research, Tech. Rep, 60.

10. Frank, M. R., Sun, L., Cebrian, M., Youn, H., & Rahwan, I. (2018). Small cities face greater impact from automation. Journal of The Royal Society Interface, 15(139), 20170946.

11. Xie, Y., & Zhou, X. (2014). Income inequality in today's China. Proceedings of the National Academy of Sciences, 111(19), 6928-6933.

12. Brelsford, C., Lobo, J., Hand, J., & Bettencourt, L. M. (2017). Heterogeneity and scale of sustainable development in cities. Proceedings of the National Academy of Sciences, 114(34), 8963-8968.

13. Cheng, H., Jia, R., Li, D., & Li, H. (2019). The rise of robots in China. Journal of Economic Perspectives, 33(2), 71-88.

14. Giuntella, O., & Wang, T. (2019). Is an Army of Robots Marching on Chinese Jobs?.

15. Yeh, A. G. O., & Wu, F. (1999). The transformation of the urban planning system in China from a centrally-planned to transitional economy. Progress in planning, 51(3), 167-252.

16. T. Economist, (2017), China pushes towns to brand themselves, then regrets it. The Economist (available at https://www.economist.com/news/china/21732826-officials-beijing-fret-local-boosters-are-getting-carried-away-china-pushes-towns-brand).

17. Duranton, G., & Puga, D. (2000). Diversity and specialisation in cities: why, where and when does it matter?. Urban Studies, 37(3), 533-555.

18. Zhu, S., He, C., & Zhou, Y. (2017). How to jump further and catch up? Path-breaking in an uneven industry space. Journal of Economic Geography, 17(3), 521-545.

19. Lixu, L. (2004). China's higher education reform 1998–2003: A summary. Asia Pacific education review, 5(1), 14.

20. Jiang, Y., & Shen, J. (2010). Measuring the urban competitiveness of Chinese cities in 2000. Cities,27(5), 307-314.

21. Zheng, S., & Kahn, M. E. (2013). China's bullet trains facilitate market integration and mitigate the cost of megacity growth. Proceedings of the National Academy of Sciences, 110(14), E1248-E1253.

22. Kumar, U., Kumar, V., & Kapur, J. N. (1986). Normalized measures of entropy. International Journal Of General System, 12(1), 55-69.

23. Eagle, N., Macy, M., & Claxton, R. (2010). Network diversity and economic development. Science, 328(5981), 1029-1031.

24. Hidalgo, C. A., Klinger, B., Barabási, A. L., & Hausmann, R. (2007). The product space conditions the development of nations. Science, 317(5837), 482-487.

25. Simpson, E. H. (1951). The interpretation of interaction in contingency tables. Journal of the Royal Statistical Society: Series B (Methodological), 13(2), 238-241.

26. Lovric, M. (2011). International Encyclopedia of Statistical Science. Springer.

27. Chiesura, A. (2004). The role of urban parks for the sustainable city. Landscape and urban planning, 68(1), 129-138.

28. Zhu, J. (1998). Data envelopment analysis vs. principal component analysis: An illustrative study of economic performance of Chinese cities. European journal of operational research, 111(1), 50-61.

29. Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. The American Economic Review, 70(5), 950-959.

30. Bessen, J. (2015). Learning by doing: the real connection between innovation, wages, and wealth. Yale University Press.

31. Muneepeerakul, R., Lobo, J., Shutters, S. T., Goméz-Liévano, A., & Qubbaj, M. R. (2013). Urban economies and occupation space: Can they get “there” from “here”?. PloS one, 8(9), e73676.

32. Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. WW Norton & Company.

33. Spitz-Oener, A. (2006). Technical change, job tasks, and rising educational demands: Looking outside the wage structure. Journal of labor economics, 24(2), 235-270.

34. Bettencourt, L. M., Lobo, J., Helbing, D., Kühnert, C., & West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. Proceedings of the national academy of sciences, 104(17), 7301-7306.

35. Melo, P. C., Graham, D. J., & Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies. Regional science and urban Economics, 39(3), 332-342.

36. Strano, E., & Sood, V. (2016). Rich and poor cities in Europe. An urban scaling approach to mapping the European economic transition. PloS one, 11(8), e0159465.

37. David, H. J. J. O. E. P. (2015). Why are there still so many jobs? The history and future of workplace automation. Journal of economic perspectives, 29(3), 3-30.

38. Dengler, K., & Matthes, B. (2018). The impacts of digital transformation on the labour market: Substitution potentials of occupations in Germany. Technological Forecasting and Social Change, 137, 304-316.

Acknowledgments: We thank Esteban Moro (Universidad Carlos III de Madrid) and Lijun Sun (McGill University) for helpful suggestions and comments on this study. We also thank the anonymous reviewers for their valuable suggestions.

Author contributions: X.L., H.C., M.C. and I.R. conceived of the research question. X.L., P.X. and X.Q. processed the Census data. X.L., H.C., M.F. and M.C. interpreted the result. X.L., H.C., M.F., X.Q., M.C. and I.R drafted the manuscript and compiled supplementary information. All authors edited the manuscript and supplementary information and aided in concept development.

Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and the Supplementary Materials. Additional data used in our study, including raw Census data and city statistics, will be publicly available for research purposes.

Supplementary Materials

Fig. S1. The fixed asset investments over city size. Panel (a) shows the result under the resources division and (b) the administrative division. The result of the linear regression is presented in tables, and the predictor city size is log-transformed.

	β	R^2	p_{value}
non-premium	0.66	0.55	3.7e-15
premium	1	0.55	5e-05
all	0.84	0.65	1.8e-24

	β	R^2	p_{value}
non-elite	0.62	0.55	2.4e-15
elite	1	0.84	1.3e-08
all	0.84	0.65	1.8e-24

Notes: Data comes from the 2010 Census and the China City Statistical Yearbook 2010.

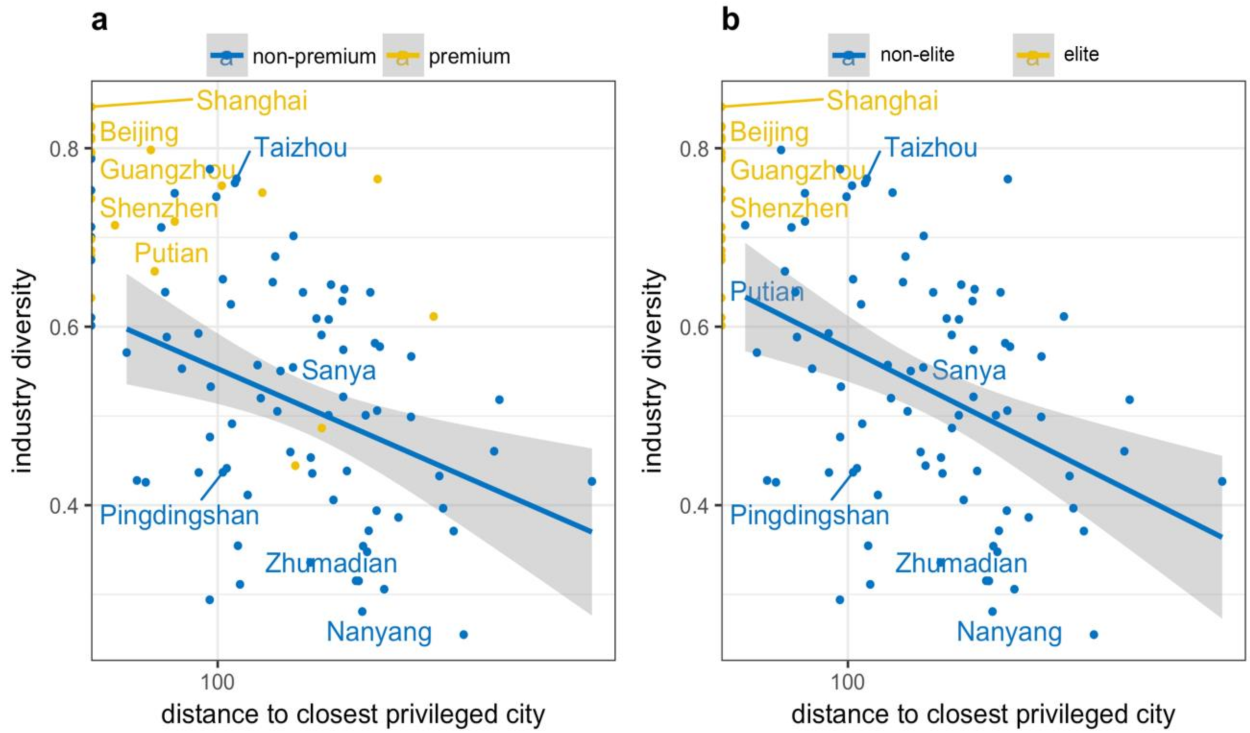
Fig. S2. The comparison of expected job impact rate across Chinese and US cities. The predictor city size is log-transformed. The job impact rate and city size are obtained from (10).

	β	R^2	p_{value}
non-premium	0.062	0.186	7.1e-05
premium	-0.11	0.367	0.0022
US cities	-0.032	0.279	1.1e-28

	β	R^2	p_{value}
non-elite	0.06	0.189	4.4e-05
elite	-0.047	0.304	0.012
US cities	-0.032	0.279	1.1e-28

Notes: Data of city size comes from the 2010 Census.

Fig. S3. Industry diversity over the distance to the closest elite cities. (a) The resources division, (b) The administrative division. The predictor distance to the nearest elite city is log-transformed.

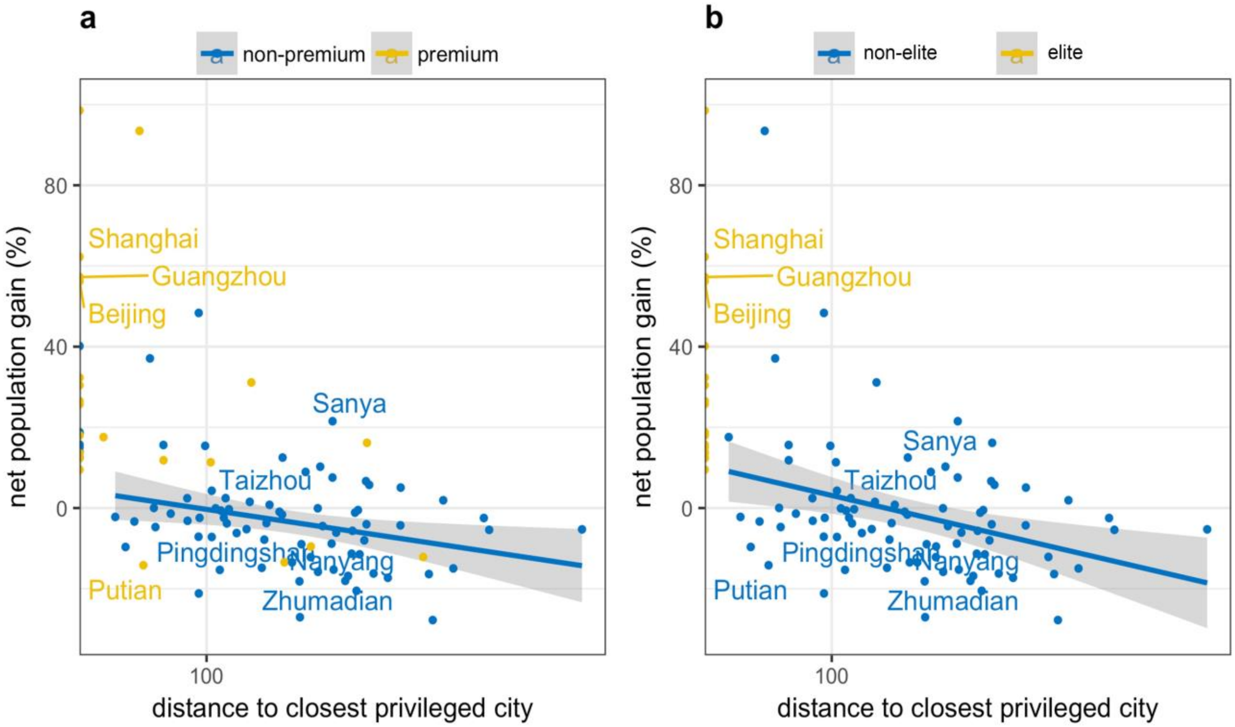


	β	R^2	p_{value}
non-premium	-0.22	0.12	0.0023

	β	R^2	p_{value}
non-elite	-0.26	0.15	0.00026

Notes: Data comes from the 2010 Census.

Fig. S4. Population net gain over the distance to the closest elite cities. (a) The resources division, (b) The administrative division. The net population gain represents the number of the population coming into the city. The predictor distance to the closest elite city is log-transformed.



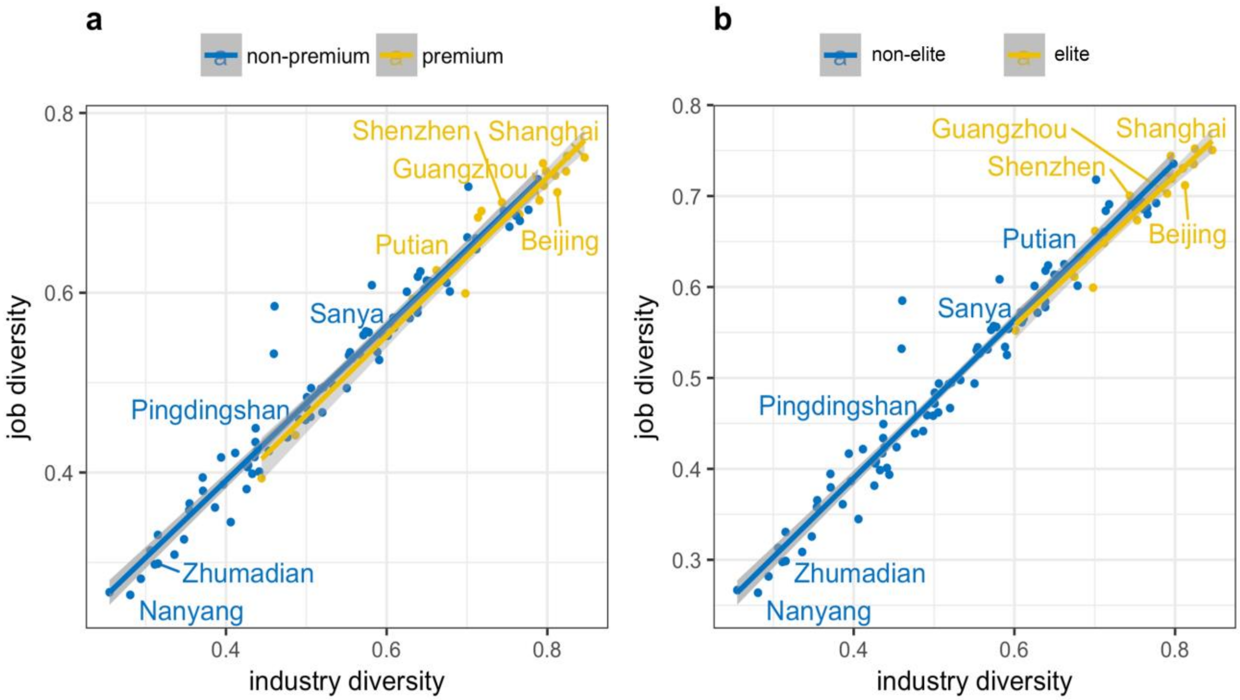
	β	R^2	p_{value}
non-premium	-17	0.081	0.015

	β	R^2	p_{value}
non-elite	-26	0.11	0.0021

Notes: Population data comes from the 2010 Census.

24

Fig. S5. Job diversity over industry diversity. (a) The resources division, (b) The administrative division.



	β	R^2	p_{value}
non-premium	0.86	0.948	2.7e-51
premium	0.88	0.96	3.7e-16
all	0.85	0.962	4.8e-73

	β	R^2	p_{value}
non-elite	0.87	0.951	2.6e-54
elite	0.82	0.941	1.6e-12
all	0.85	0.962	4.8e-73

25

Fig. S6. Job growth over city size between elite and non-elite cities.

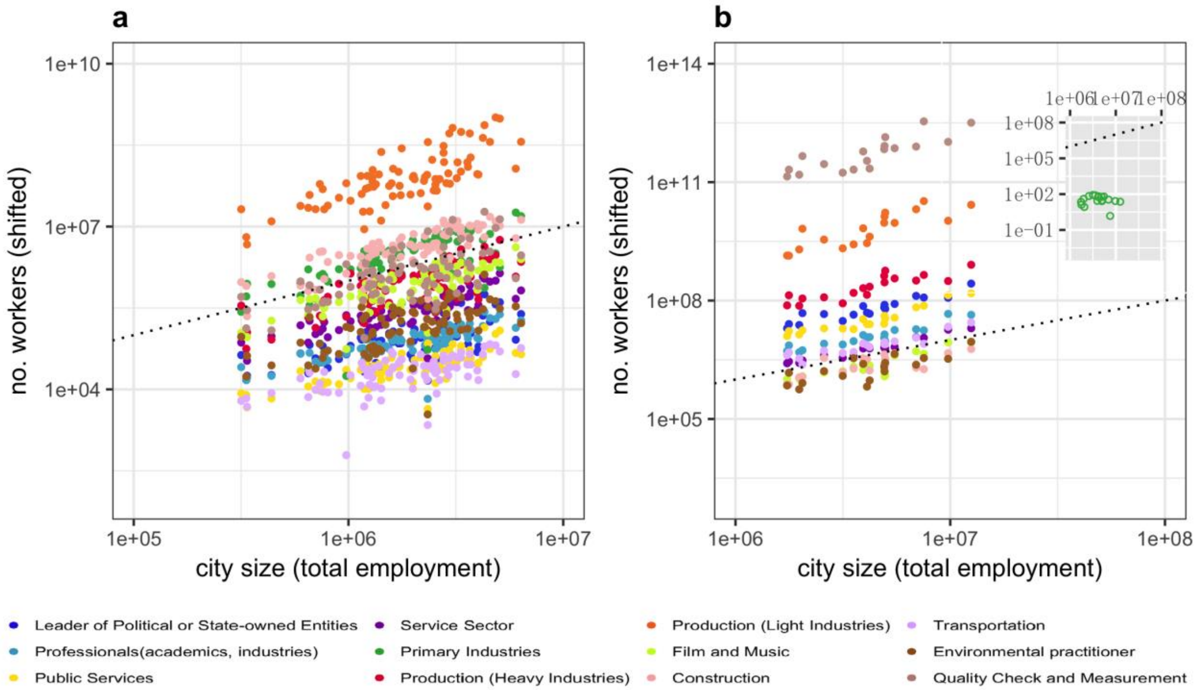
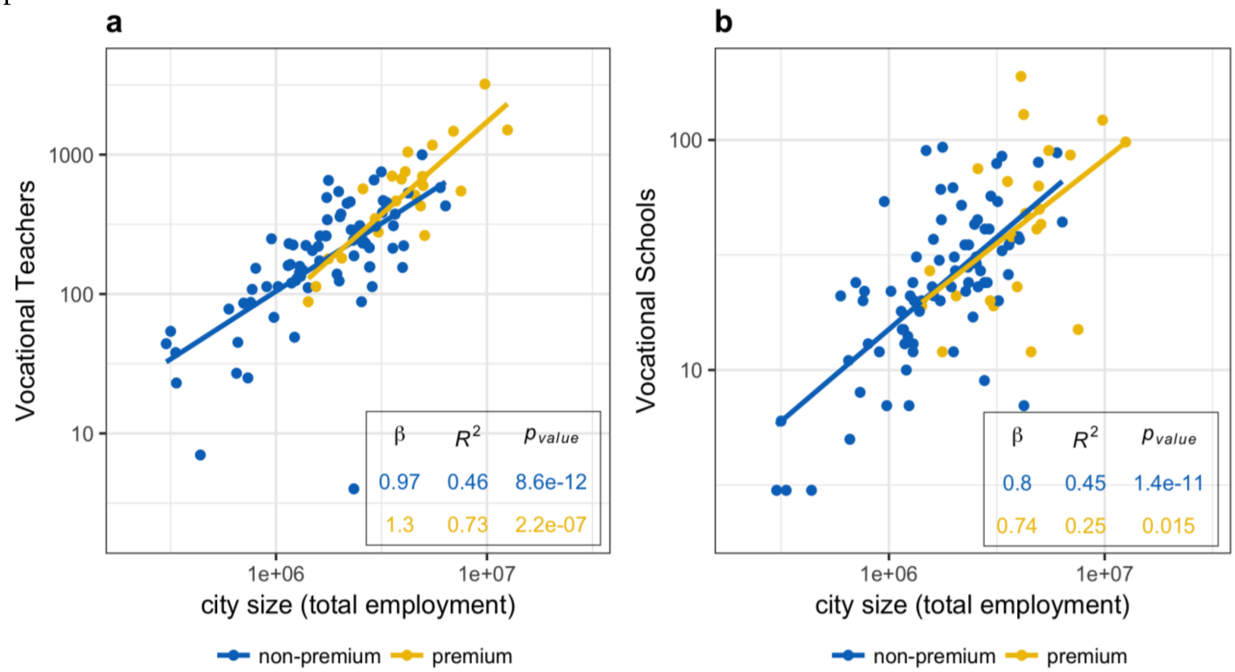
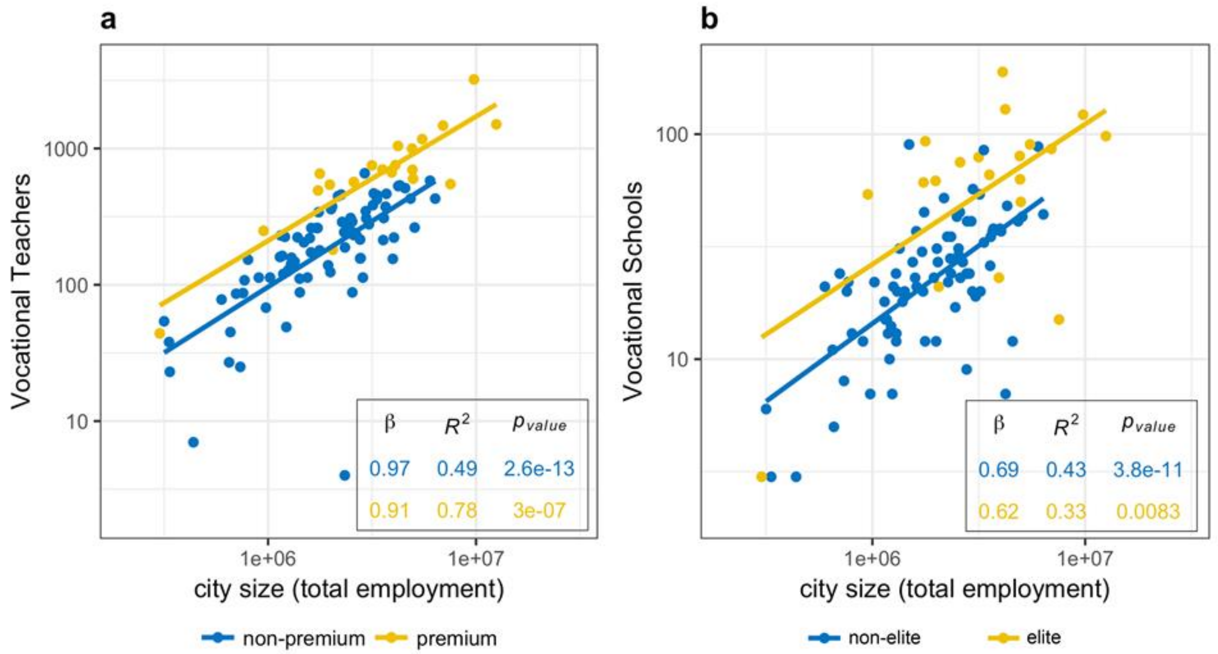


Fig. S7. Allocations of vocational education resources over city size under the premium resources division. (a) the vocational teachers grow superlinearly in premium cities and linearly in non-premium cities. (b) the vocational schools grow sublinearly in both premium and non-premium cities.



Notes: Data comes from the 2010 Census and the China City Statistical Yearbook 2010.

Fig. S8. Allocations of vocational education resources over city size under the administrative division. (a) the vocational teachers grow linearly in non-elite cities and sublinearly in elite cities. (b) the vocational schools grow sublinearly in both elite and non-elite cities.



Notes: Data comes from the 2010 Census and the China City Statistical Yearbook 2010.

Table S1. China's cities division system. The numbers of universities under projects '211' and '985' and the daily operating frequency of bullet trains are shown in columns 2, 3, 6, and 7. The estimation of job impact rates is shown in columns 4 and 8.

City	Universities Count	Bullet Trains	Expected Job Impact Rate	City	Universities Count	Bullet Trains	Expected job impact rate
Yibin	0	0	81.99%	Puyang	0	0	81.09%
Luzhou	0	0	81.26%	Kaifeng	0	58	82.60%
Lhasa[1]	1	0	77.95%	Zhumadian	0	93	83.65%
Heihe	0	0	73.94%	Luohe	0	97	79.15%
Beijing[1,2]	35	349	63.83%	Pingdingshan	0	0	79.52%
Shanghai[1,2]	10	372	67.16%	Sanmenxia	0	64	75.81%
Tianjin[1,2]	3	261	70.78%	Zhengzhou[1,2]	1	230	72.23%
Shenzhen[1,2]	0	246	72.05%	Lanzhou[1]	2	112	70.23%
Guangzhou[1,2]	4	377	69.26%	Qingyang	0	0	82.05%
Foshan[2]	0	148	71.83%	Qingdao[1]	1	67	74.17%
Zhuhai	0	80	65.44%	Weihai	0	46	76.16%
Huizhou	0	131	77.24%	Jinan[1,2]	1	248	73.73%
Jiangmen	0	15	71.72%	Dongying	0	0	75.48%
Kunming[1]	0	73	72.91%	Liaocheng	0	0	81.29%
Yuncheng	0	43	75.35%	Linyi	0	0	79.97%
Changzhi	0	0	77.47%	Tai'an	0	0	77.81%
Linfen	0	58	77.83%	Dezhou	0	124	82.42%
Lvliang	0	0	77.12%	Binzhou	0	0	80.74%

Shuozhou	0	0	76.28%	Chizhou	0	53	77.40%
Jincheng	0	0	76.40%	Bengbu[2]	0	143	80.58%
Xinzhou	0	0	79.92%	Suzhou	0	88	78.97%
Yangquan	0	39	72.39%	Tongling	0	95	72.05%
Taiyuan[1]	0	69	67.44%	Huainan	0	45	73.86%
Jinzhong	0	18	75.79%	Huaibei	0	0	76.11%
Nanjing[1,2]	8	335	69.02%	Huangshan	0	49	72.15%
Yancheng	0	0	79.67%	Liu'an	0	71	80.49%
Xuzhou[2]	1	273	78.07%	Xuancheng	0	0	78.19%
Taizhou	0	48	76.06%	Haozhou	0	0	83.37%
Chifeng	0	0	79.84%	Hefei[1,2]	2	213	73.89%
Suizhou	0	26	81.54%	Fangchenggang	0	7	79.50%
Shiyan	0	4	79.24%	Liuzhou	0	128	77.70%
Shaoxing[2]	0	168	74.06%	Guigang	0	111	82.99%
Jiaxing[2]	0	184	73.61%	Ningde	0	56	71.40%
Huzhou[2]	0	144	76.27%	Zhangzhou	0	0	70.87%
Wenzhou[2]	0	148	76.53%	Xiamen[1,2]	1	212	68.91%
Zhoushan	0	0	73.65%	Putian[2]	0	183	77.49%
Lishui	0	63	69.21%	Longyan	0	43	76.74%
Quzhou	0	97	74.07%	Quanzhou[2]	1	208	74.97%
Taizhou	0	106	75.12%	Nanping	0	75	76.41%
Jinhua	0	46	72.58%	Xi'an[1,2]	5	155	71.99%
Hangzhou[1,2]	1	247	70.54%	Yan'an	0	6	68.07%

Ningbo[1,2]	1	163	73.07%	Baoji	0	86	76.81%
Luoyang	0	120	79.53%	Turpan	0	31	67.66%
Shangqiu	0	98	82.52%	Jilin	0	94	78.46%
Xinyang	0	122	81.58%	Liaoyuan	0	0	81.32%
Xuchang	0	68	80.56%	Songyuan	0	0	82.01%
Nanyang	0	4	83.33%	Hengyang[2]	0	195	81.19%
Xinxiang	0	75	80.48%	Xiangtan	0	69	78.03%
Hebing	0	60	76.50%	Haikou[1]	0	57	72.10%
Anyang	0	81	80.64%	Sanya	0	51	70.71%
Jiaozuo	0	14	78.18%	Guiyang[1]	0	125	73.75%

Note: Cities marked with superscript 1 and 2 are elite and premium cities, respectively. It should be noticed that even though Lhasa isn't counted as a premium city due to the lack of bullet train passing-by, its railway investments are much higher than any other bullet train lines.