

Article

# Spatial Distribution Pattern of the Headquarters of Listed Firms in China

Ling Zhang <sup>1,\*</sup> , Hui Zhang <sup>1</sup> and Hao Yang <sup>2,\*</sup>

<sup>1</sup> School of Economics, Peking University, 5 Summer Palace Road, Haidian District, Beijing 100871, China; nk94zhang@pku.edu.cn

<sup>2</sup> Beijing Academy of Social Sciences, 33 North Fourth Ring Road, Chaoyang District, Beijing 100101, China

\* Correspondence: zhangling19891101@163.com (L.Z.); yanghaobjdx@aliyun.com (H.Y.)

Received: 2 July 2018; Accepted: 19 July 2018; Published: 22 July 2018



**Abstract:** This study chose 4667 listed firms with headquarters in China as the research objects. The spatial aggregations of headquarters in sub-periods and subsectors were investigated based on prefecture-level cities using Global Moran's I statistic, the Getis-Ord  $G_i^*$  statistic, and kernel density estimation. The spatial association and frequent patterns of headquarters belonging to different sectors were analyzed using the Apriori algorithm. The results showed the following: (1) Most headquarters of listed firms agglomerate around megacities and concentrate in the coastal regions. (2) Many prefer that headquarters be located in an area within a 5 km radius, forming the pivotal local industry cluster. (3) Industrial, information technology, and consumer discretionary firms show a higher degree of agglomeration than materials and health care firms. Materials and consumer discretionary industries have a stronger spillover effect on surrounding cities compared to industrial, information technology, and health care firms. (4) Service industries show a strong spatial association with other sectors at all distance thresholds and frequent items. These results can provide a reference for future considerations of headquarters locations, industrial development, and the reshaping of urban networks.

**Keywords:** headquarters; listed firms; spatial distribution; Apriori algorithm

## 1. Introduction

A market is composed of a large number of small firms and a small number of large firms [1]. Starting a new business, establishing headquarters, and opening joint ventures, branches, and subsidiaries contribute to expansions in regional employment, provide earning opportunities, and promote local economic growth [2,3]. In general, a very small number of large firms account for most of the economic output and are crucial to the development of the regional economy [4]. Moreover, unforeseen shocks to a single sector that is strongly interconnected with others in an economic system can have dramatic effects on the overall economy [4,5]. Analyzing the behavior of microeconomic entities, especially large-scale or strategically important companies, can thus enable a better understanding of the macroeconomy.

Compared to non-listed companies, listed companies benefit from the competitive advantages of a larger scale, better performance, smoother information transmission, and a stronger ability to source finance, making them the most dynamic and high-potentially enterprise organizations in the modern market economy [6–8]. Listed companies with strong market performance and a good social reputation can easily attract and retain talents, which in turn promotes scientific and technological progress and innovation, forming a virtuous circle. When emergencies arise in the market, listed companies can recapitalize themselves and spread risk with the aid of the capital market [9,10]. Headquarters outsource a large amount of work, such as advertising, consulting, accounting, and

financing, which stimulates the development of the producer services industry and creates more jobs and tax revenue [11–15]. In addition, employees at these headquarters, who are regarded as elite members of society, can help enhance levels of local consumption. Cities in which more headquarters of large firms are located are often considered the “commanding” centers of the economy, occupying a dominant strategic and economic position [16]. The “headquarters economy” is also considered “a path of industrial upgrade”. Examples of municipalities seeking to attract headquarter activity can be found across North America, Europe, and the Asia-Pacific region [13,17–20]. Therefore, the headquarters of listed firms can be considered an ideal research object for microeconomic studies, and the location and behavior of these headquarters have drawn a great deal of attention from both the public and private sectors [17].

Analyzing the evolution of the spatial distribution and spatial aggregation of large-scale listed firms’ headquarters is fundamental for understanding the structure and constitution of each market and of the regional economy. This, in turn, is key to effective urban planning and strategic urban development. In previous studies, geographers and economic geographers have tended to use the concentration ratio of industry, the entropy index, the Herfindahl–Hirschman index, the Space Gini coefficient, and the EG (Ellison-Glaese) index to identify clustering in an industry. They have, moreover, conducted spatial autocorrelation analysis using the Gi-statistic or Moran’s I index, and used the ESDA (Exploratory Spatial Data Analysis) method to study the topological relationships of geospatial entities in a city or region [21,22]. However, these methods for measuring agglomeration are mainly cluster-based approaches based on administrative districts. A large difference in district size may be misleading and conceal important information related to small areas. Few studies have analyzed spatial aggregation and spatial location correlation among different industries based on point position relationships. With the advent of big data, the ability to access data and the potentially for data mining have greatly improved. The Apriori algorithm, a classic data mining method, offers the advantages of a simple theoretical basis, easy implementation, and good performance. This makes it well suited for calculating the spatial association rules and frequent patterns of listed firms from a microenterprise perspective.

The remainder of this paper is organized as follows: Section 2 conducts a literature review regarding patterns in the locations of listed firms’ headquarters and the evolution of their spatial distribution. Section 3 describes the methodology for data collection and analysis. In Section 4, the spatial aggregation of listed firm headquarters is analyzed in sub-periods and subsectors based on prefecture-level cities. Further, the spatial associations and frequent patterns of headquarters belonging to different sectors are analyzed based on coordinates. Finally, Section 5 provides concluding remarks.

## 2. Literature Review

Researchers in economic geography have an ongoing interest in defining the patterns in business location and identifying the complex interaction rules behind the evolution of the spatial distribution of company headquarters. Geographers began analyzing the locations of corporate headquarters in the 1960s. Goodwin provided the first analysis of the metropolitan distribution of US management activity [23]. By assessing changing levels in the geographic concentration of US headquarters’ activity, Semple and Phipps proposed a developmental sequence of headquarters activity as follows: “beginning with a period of increasing geographic concentration, a period of peak concentration in nationally dominant cities, and finally involving the growth of a dispersed, national system of major regional corporate centers” [24,25]. This conceptualization was gradually demonstrated by later research and became the early theoretical foundation for this field [16,26]. Studies on headquarters location specifically consider the benefits and costs of locational choices and the influences of state factors, policy interventions, and institutional background. On the one hand, firms prefer to locate their headquarters in the central business districts (CBDs) of larger metropolitan areas due to the regional advantages of high-quality infrastructure, high-end producer services, excellent university and research institutes, and convenient living environments. They choose to gather in only a few

larger metropolitan cities and set up close to each other to facilitate the sharing of common inputs and the same labor market, in addition to enjoying information spillover [14,27–29]. For example, Sydney and Melbourne accounted for 44.2% and 27.8% of the headquarters of the largest 300 listed firms in Australia in 2010 [30]. On the other hand, due to the growing cost of operating in CBDs and the development of information and communications technology, some firms have moved their headquarters outside of urban centers [31]. The evolution of the distribution of corporate headquarters appears to be linked to the economic, social, and political histories of particular places [32]. Some recent studies have shown that headquarters activities are geographically dispersed and configured in increasingly complex ways. Large firms (e.g., General Electric, IBM, Lenovo, and Royal Dutch Shell) have disaggregated and dispersed their financial, legal, and managerial headquarters activities across different locations [33–35]. Pan (2003) investigated the relocation of the headquarters of listed firms from 2001 to 2012 in China and found that Beijing and Shanghai were the most attractive cities for inter-province relocation. Firms in the mining and real estate industries have the highest relocation rates, and firms in the service industry have higher relocation rates than those in the manufacturing industry [36].

Furthermore, geographical research has not been limited to investigating the location and distribution of headquarters activity. Since the 1980s, quaternary research, which is the study of “corporate decision-making and high-level business”, has broadened to examine the nature of the interactions among headquarters, the internal elements of a firm (such as subsidiaries, divisions, and plants), and external elements in the firm’s competitive environment (such as suppliers, sources of finance, and customers) [16,37–39]. Research by Rice and Pooler (2013) [39] revealed the geographically evolving tracks of the largest subsidiaries in North America from 1996 to 2004, stressing that cities with a high subsidiary-parent ratio represent additional spatial focal points of power. Li et al. (2016) suggested that regulative and cultural distances between home and host countries can positively influence the quality of the headquarters-subsidiary relationship, in contrast to the previously held opinion that institutional distance increases the potentially for conflicts between headquarters and subsidiaries [40,41].

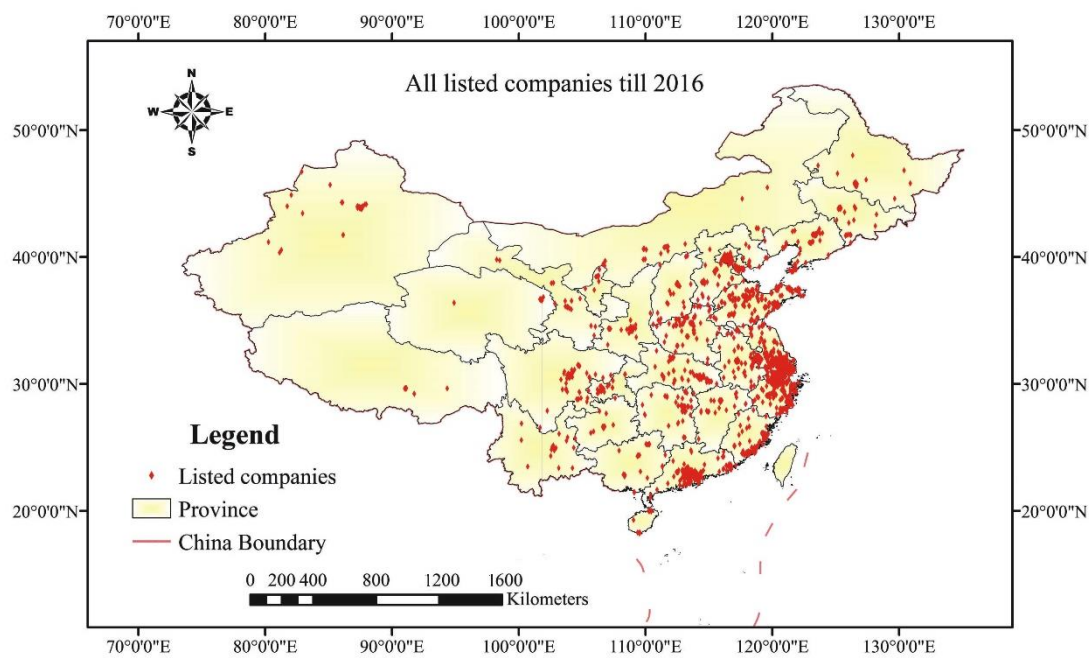
The preceding summary shows that extensive research has been conducted on the development of the distribution and location patterns of company headquarters and the complex interaction rules behind them. A large number of these studies have taken MNCs (multinational corporations) as the research object, and only a small number have focused on temporal tracking and geographic analysis of the headquarters of listed firms belonging to different industries in China. Few researchers have investigated the rules of spatial association and frequent patterns of headquarters in different industries. In light of these opportunities for contribution, the present study has the potential to broaden research on the spatial distribution of headquarters.

### 3. Data and Methodology

#### 3.1. Data

Information (e.g., the date of establishment, office address, and financials) on all firms listed prior to the end of 2016 in both domestic and overseas stock exchanges was obtained from Wind, a high-quality, large-scale financial engineering and financial data warehouse in China. A total of 4667 publicly listed companies with offices in China were included in this research, as shown in Figure 1. The WGS84 coordinates of all listed firms were extracted from the head office address using the API interface of the Baidu Map open platform.

This study adopted the classification standards used by Wind, which include 11 first-level industries and 24 second-level industries; these refer to the Global Industries Classification Standard (GICS) industry system but with slight adjustments according to actual Chinese conditions.



**Figure 1.** Distribution of listed companies in China.

### 3.2. Spatial Autocorrelation Model

#### 3.2.1. Global Moran's I Statistic

The Global Moran's I statistic was used to explore the degree of clustering or dispersion of the headquarters of listed firms throughout China and measure spatial autocorrelation based on feature location and attribute values, with prefecture-level cities as the assessment units [42–44]. This can be expressed by Formulas (1) and (2):

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} z_i z_j}{\sum_{i=1}^n z_i^2} \quad (1)$$

where

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{i,j} \quad (2)$$

where  $z_i$  is the deviation of an attribute for city  $i$  from its mean ( $x_i - \bar{x}$ ),  $w_{i,j}$  is the spatial weight between cities  $i$  and  $j$ ,  $n$  is equal to the total number of cities, and  $S_0$  is the aggregate of all spatial weights.

The Global Moran's I statistic evaluates whether the pattern expressed is clustered, dispersed, or random, using both the  $z$ -score and  $p$ -value to evaluate the significance of the index;  $p$ -values are numerical approximations of the area under the curve for a known distribution limited by the test statistic. The  $z_I$ -score for the statistic is computed as per Formula (3):

$$z_I = \frac{I - \frac{1}{1-n}}{\sqrt{E[I^2] - E[I]^2}} \quad (3)$$

The range of values for the Global Moran's I index is  $-1$  to  $1$ . A positive coefficient indicates that contiguous cities have a similar number of listed firms, and a higher value suggests a stronger association. By contrast, the lower the  $z_I$  value, the stronger the negative correlation. When Global Moran's I approaches zero, it indicates a random distribution.

### 3.2.2. Getis-Ord $G_i^*$ Statistic

Performing better than Local Moran's Index in detecting the center and range of headquarter gathering area, the Getis-Ord  $G_i^*$  statistic was used to identify cold and hot spots. The locations of clusters of high or low values were identified by calculating their  $Z_G$  scores. The formula by which the statistic is calculated is as follows [45,46]:

$$G = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} x_i x_j}{\sum_{i=1}^n \sum_{j=1}^n x_i x_j} \quad (4)$$

where  $x_i$  and  $x_j$  are attribute values for cities  $i$  and  $j$ ,  $w_{i,j}$  is the spatial weight between features  $i$  and  $j$ ,  $n$  is the number of prefecture-level cities, and  $i \neq j$ .

The  $Z_G$ -score for the statistic is computed as

$$Z_G = \frac{G - E[G]}{\sqrt{V[G]}}, \quad (5)$$

where

$$E[G] = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j}}{n(n-1)} \text{ and} \quad (6)$$

$$V[G] = E[G^2] - E[G]^2 \quad (7)$$

The Getis-Ord  $G_i^*$  statistic of each element is its  $Z_G$  score. For positive  $Z_G$  scores with obvious statistical significance, a higher  $Z_G$  score implies a closer grouping of high values. For negative  $Z_G$  scores with obvious statistical significance, a lower  $Z_G$  score indicates a closer grouping of low values. The Jenks natural breaks method was used in this study to divide the  $Z_G$  scores into four grades: cold spot, secondary cold spot, secondary hot spot, and hot spot.

### 3.3. Kernel Density Estimation

Kernel density estimation was used in this study to describe the agglomeration of firms [47]. If one industry has a larger frequency at a specific distance, it indicates that this industry has a higher degree of agglomeration at this distance. The kernel estimator used in this paper is shown in Formula (8), which was developed by Silverman [48]:

$$K(d) = \frac{1}{n(n-1)h} \sum_{i=1}^{n-1} \sum_{j=i+1}^n f\left(\frac{d - d_{i,j}}{h}\right), \quad (8)$$

where  $n$  is the number of observations,  $d_{i,j}$  is the distance between the  $i$ th and the  $j$ th observation, and  $h$  is the smoothing parameter that is significant, whose formula is as follows:

$$H = 0.9An^{-1/5} \quad (9)$$

where  $A$  is min (standard deviation of distance between headquarters, four-minute distance between headquarters).

### 3.4. Apriori Algorithm

The objective of this work was to integrate GIS and association mining techniques to determine the distribution relationships among headquarters classified in different sectors within a certain distance. Such information can be valuable for decision-making. If the frequency of occurrence of industries  $X$  and  $Y$  is high, we can say they are strongly associated with each other. The association rules are defined as follows:

$A = \{a_1, a_2, \dots, a_m\}$  denotes the itemset, in which each item represents a specific literal.  $D$  represents a set of industries distributed in a database in which each transaction  $I$  stands for an itemset ( $I \subseteq A$ ).

An association rule is denoted by  $X \rightarrow Y$ . The rule  $X \rightarrow Y$  holds in  $D$  according to two measure standards: *support* and *confidence*.

*Support* (denoted as  $Sup(X, D)$ ) represents the number of times a rule shows up in a database. The higher the value of *Support*, the more important the  $D$  is.

*Confidence* (denoted as  $Conf(X \rightarrow Y)$ ) represents the conditional probability of  $Y$  given  $X$ . This indicates that if one set of industry distribution includes  $X$ , the chance of occurrence of  $Y$  is relatively high. Thus, it is employed to evaluate the level of confidence regarding the association rule  $X \rightarrow Y$ .

The Apriori algorithm is used to find out all transaction rules that have certain user-specified minimum *support* (*Minsup*) and *confidence* (*Minconf*). First, a large itemset is detected whose *support* is larger than *Minsup*; second, association rules are generated using the large itemset. Such rules must satisfy the conditions  $Sup(X \cup Y, D) \geq Minsup$  and  $Conf(X \rightarrow Y) \geq Minconf$ .

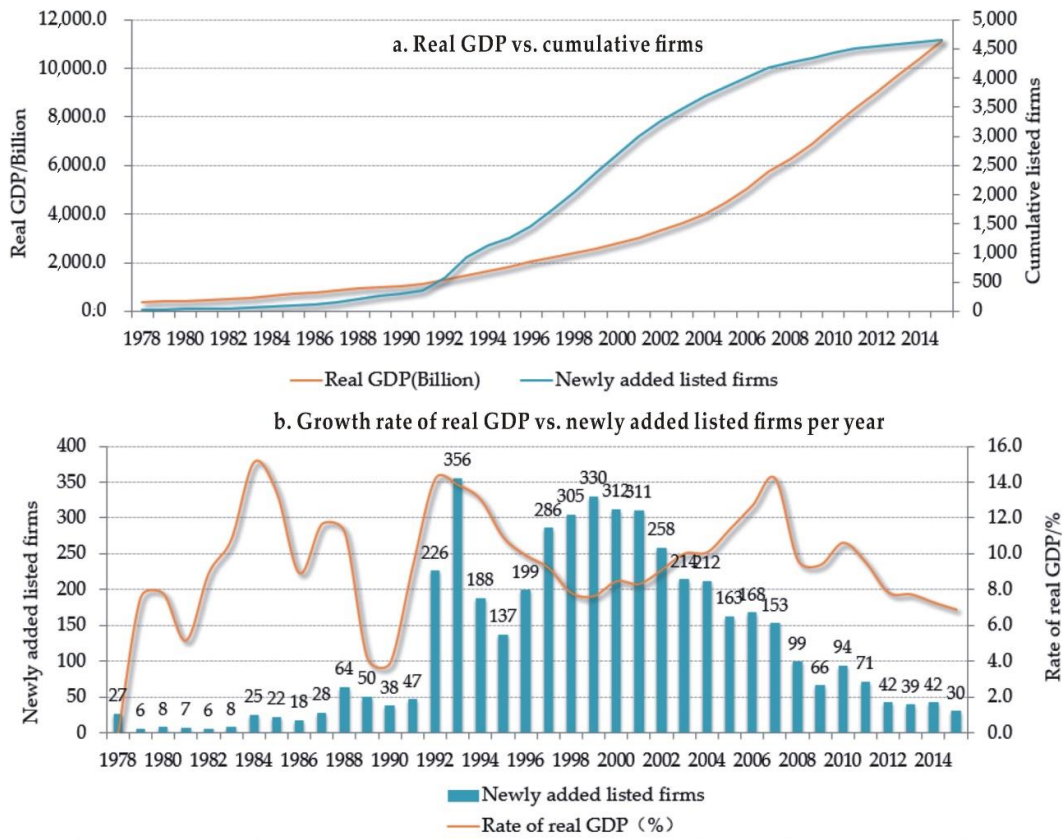
The *lift* judgment is used to judge possible biases incurred when using these measure standards. It is defined as  $Lift = Conf(X \rightarrow Y) / Sup(Y)$  [49].

## 4. Results and Analysis

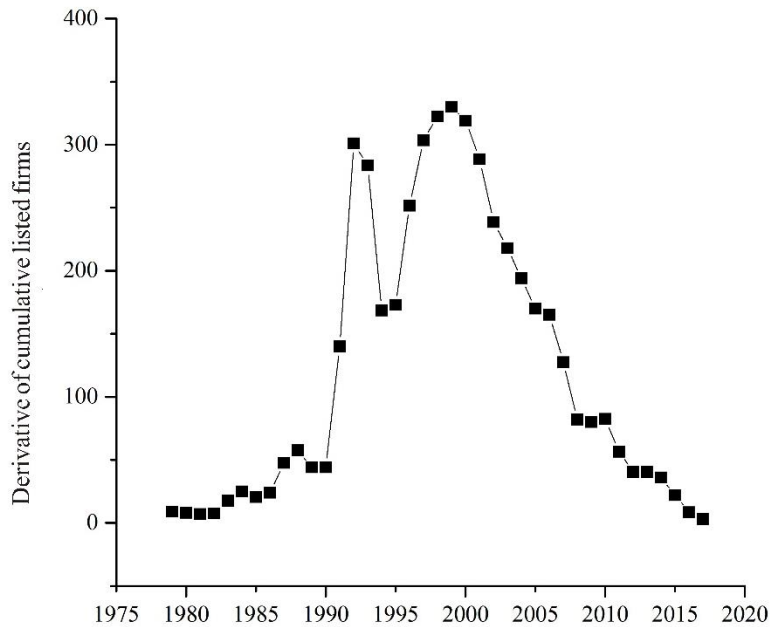
### 4.1. Temporal Growth of Listed Firms in China

China has exhibited dramatic economic growth over the past few decades. As the engines of this growth, there has been a simultaneous and comparable rise in the number of Chinese companies listed overseas or domestically (Figure 2). In 2001, revenues from publicly listed companies accounted for 14% of China's GDP [50]. By 2007, that figure had doubled, accounting for 28% of the GDP. In 2016, the total operating revenues of listed companies accounted for 44% of China's GDP. The listing activity of Chinese firms began in the late 1980s via reverse takeovers in the Hong Kong Stock market. From 1992 to the end of 2016, nearly 4700 Chinese firms became listed on a variety of stock markets.

The study period can be broadly divided into four stages on the basis of the derivative of the cumulative number of listed companies established each year (Figure 3). Because the date of their founding precedes their being listed, such firms will be referred to below as "potentially listed firms". The first period (Period I) was before 1991, during which time both annual increases in the number and cumulative number of potentially listed firms were small (Figure 2). The second period (Period II) was from 1992 to 2000, a period that witnessed extraordinary growth in the number of potentially listed firms (totaling 2337). This phenomenon was promoted by Xiaoping Deng's "South Tour Speech" in 1992. The derivative value of cumulative listed firms peaked in 2000 (Figure 3). Affected by the global financial crisis initiated by the subprime mortgage crisis of 2007–2008 in the US, there was a slowing in the growth in the number of potentially listed firms from 2001 to 2007 (Period III). Lastly, in Period IV, from 2008 to 2016, the number of new potentially listed companies dropped sharply to the level of Period I, with a decreasing trend year by year. To sum up, the annual number of new potentially listed companies has been a barometer of China's economy.



**Figure 2.** (a) The annual real GDP and the number of cumulative firms from 1978 to 2016; and (b) trends in the annual GDP growth rate and new listed firms from 1978 to 2016.



**Figure 3.** The derivative of cumulative potentially listed firms, 1978–2016.

### 4.2. Spatial Analysis in Different Periods

#### 4.2.1. Global Moran’s I Statistic

There are significant regional differences in the distribution of cities hosting the headquarters of listed firms. Global spatial autocorrelation analysis was performed on the locations of headquarters in the four periods (Table 1). The z-values ranged from 4.84 to 7.74, well over 2.58, indicating that the headquarters showed definite clustering at 99% significance (Figure 4). The Global Moran index, ranging from 0.05 to 0.09, suggests a weak positive spatial autocorrelation among the headquarters. Overall, the degree of agglomeration in headquarters locations experienced an “increase-decrease-increase” fluctuation.

Table 1. Output for Global Moran’s I statistic during different periods.

Distance Method	p-Value	z-Score	Variance	Moran Index	Period
Euclidean	0.000000	7.235523	0.000146	0.0846	I
Euclidean	0.000001	4.992633	0.000146	0.0575	II
Euclidean	0.000000	7.743269	0.000146	0.0907	III
Euclidean	0.000001	4.840777	0.000118	0.0498	IV
Euclidean	0.000000	6.143428	0.000145	0.0711	All

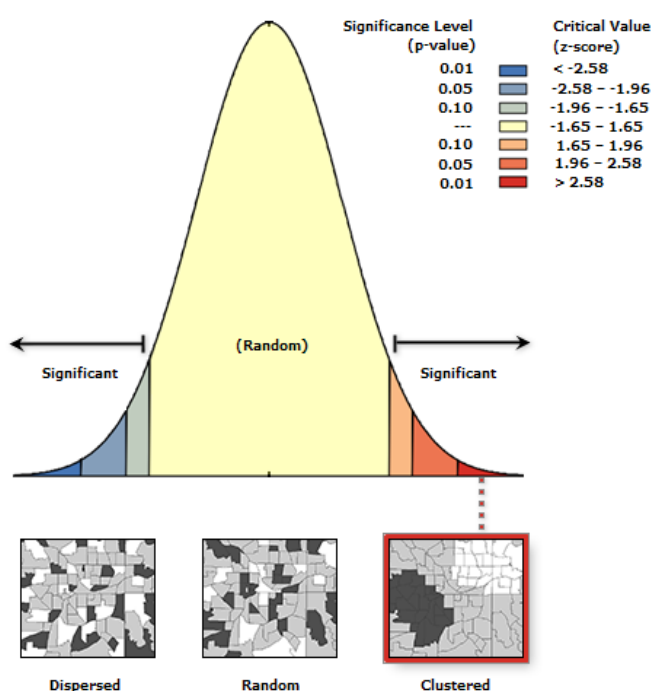


Figure 4. A spatial autocorrelation report for Global Moran’s I statistic.

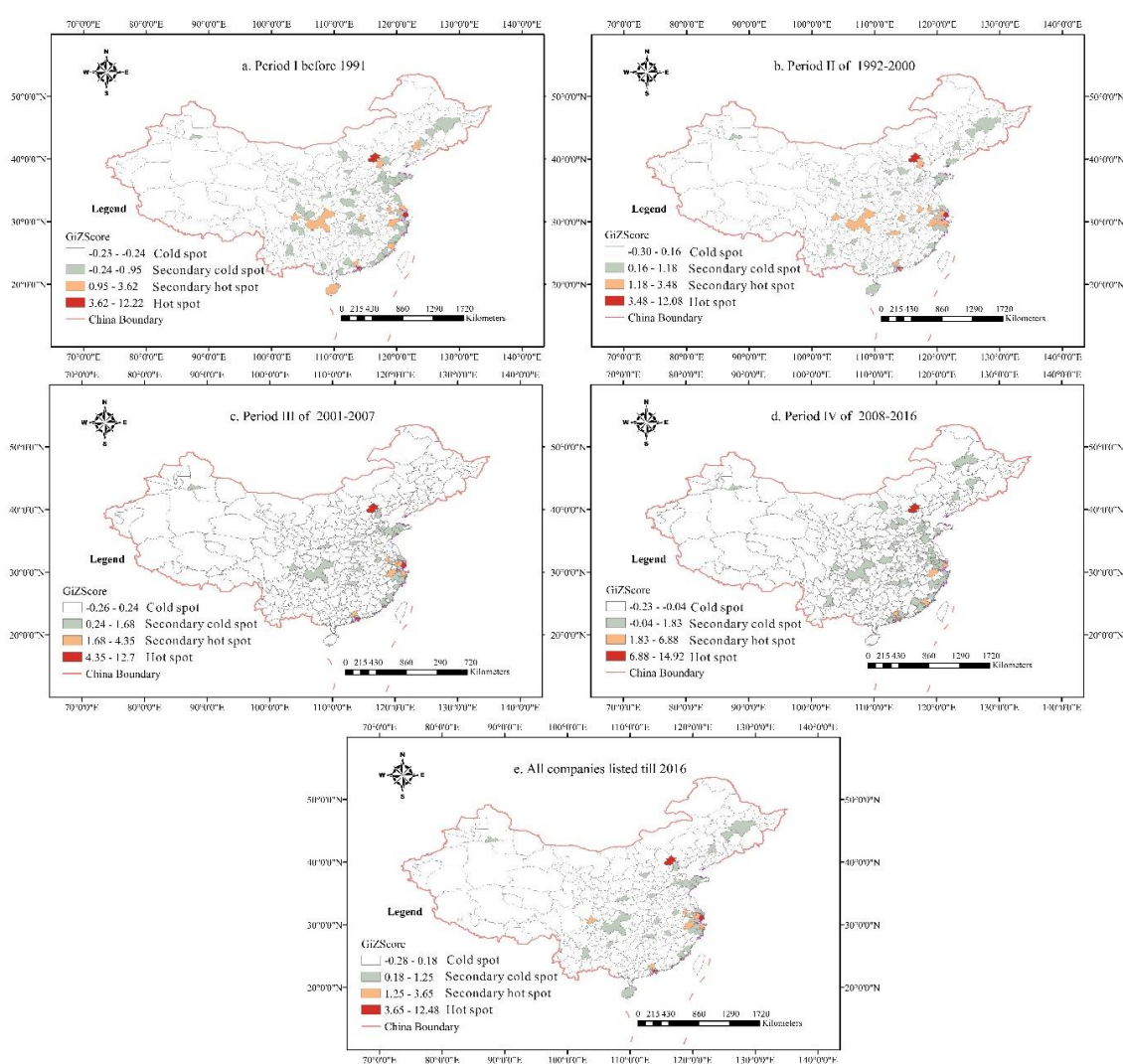
#### 4.2.2. Getis-Ord $G_i^*$ Statistic in Different Periods

The Global Moran index provides only a rough description of the distribution patterns observed. The Getis-Ord  $G_i^*$  statistic was used in this study to detect more concrete hot spots: areas with concentrations of events that are higher or lower than expected. Figure 5 illustrates the hot and cold spots for the headquarters locations identified, and Table 2 lists all hot spots and secondary cold spots in the five periods. Beijing, Shanghai, and Shenzhen were hot spots in all time nodes. Prior to 2000, Tianjin, Guangzhou, and several cities along the Yangtze River were secondary hot spots. Subsequently, the secondary hot spots mainly clustered around the Yangtze River Delta. The secondary cold spots were mainly distributed along the coastal areas and central region of China. These three types of



clustering only accounted for 12–16% of all of the cities. Both market orientation and government guidance contribute to the accumulation patterns of headquarters [1,18,29,51,52]. The implementation of the “Unbalanced Development Strategy” largely resulted in this pronounced inhomogeneity of headquarters among cities and districts.

Overall, listed company headquarters in China tend to agglomerate in a few metropolitan areas with easy access to the producer service industry, high-quality infrastructure, and economies of agglomeration. Beijing overwhelmingly outstrips the other cities in terms of the resource advantages it offers to company headquarters. The sound basic infrastructure, advanced business facilities, strong momentum of scientific and technological innovation, and continuous optimization of the innovation environment foster and attract large numbers of listed firms to establish their headquarters there [29]. The first-rate services environment, excellent business facilities, and openness of the port city of Shanghai make it another optimal location for listed firm headquarters. Shenzhen, meanwhile, offers a great degree of openness, a sound environment for creativity and entrepreneurship, and favorable policies, and these advantages have likewise attracted a large number of listed firms to locate their headquarters there. This phenomenon has verified Marshall’s external economic theory [53].



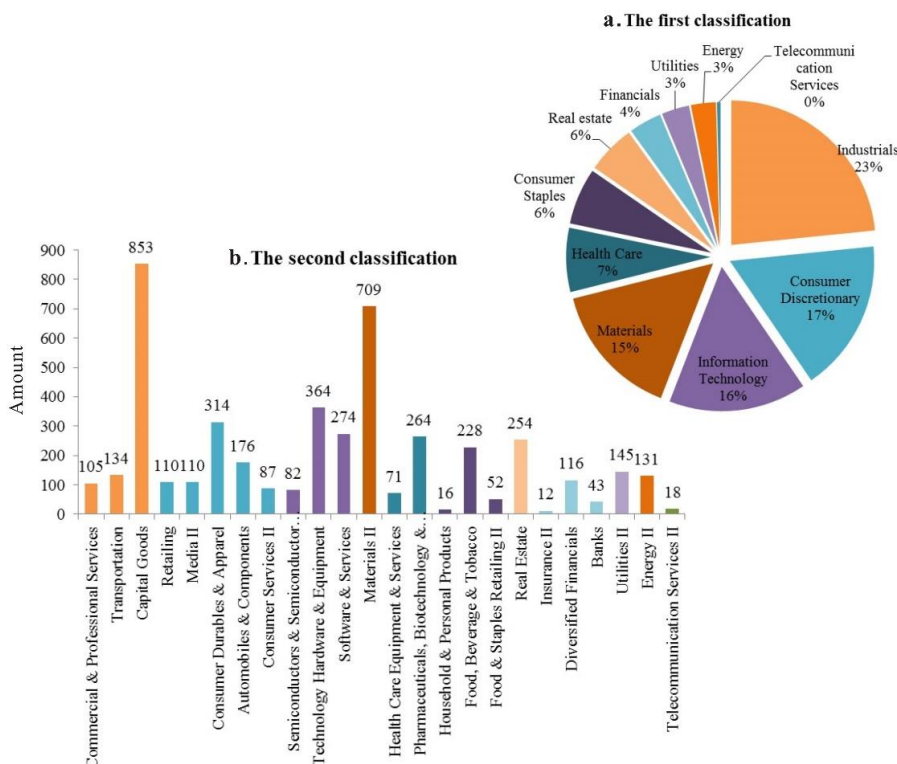
**Figure 5.** Getis-Ord  $G_i^*$  statistics for the headquarters of listed firms in different periods: (a) period I before 1991; (b) period II from 1992 to 2000; (c) period III from 2001 to 2007; (d) period IV from 2008 to 2016; and (e) all listed firm till 2016.

**Table 2.** Summary of hot spot/cold spot clusters of headquarters in different periods.

Types of Clusters	Intensity of Clusters	Periods	Number of Cities	Percentage
Hot spots	Primary	Period I	3	0.76%
		Period II	3	0.76%
		Period III	3	0.76%
		Period IV	1	0.25%
		All	3	0.76%
	Secondary	Period I	13	3.29%
		Period II	12	3.04%
		Period III	8	2.04%
		Period IV	6	1.52%
		All	8	2.04%
Cold spots	Secondary	Period I	47	11.9%
		Period II	32	8.1%
		Period III	27	6.84%
		Period IV	49	12.4%
		All	33	8.35%

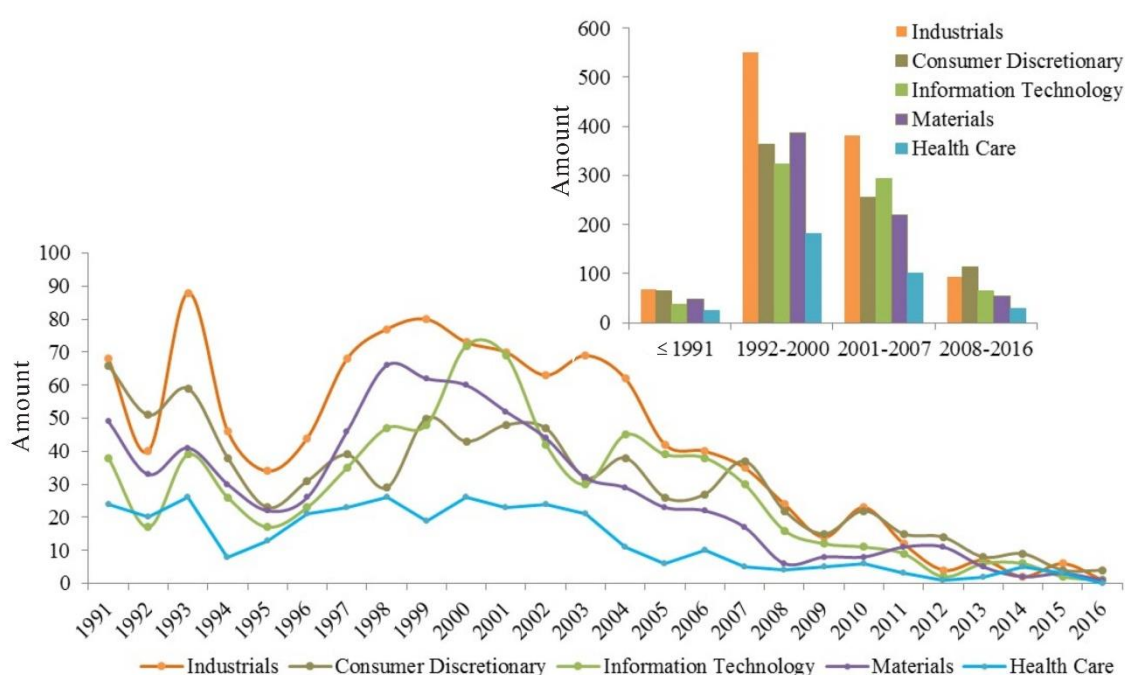
### 4.3. Spatial Patterns in Different Sectors

The number of headquarters in different sectors reflects the importance of each type of industry to the domestic economy. Figure 6 shows the proportion of different industries in the first and second levels, respectively. At the first level, the listed firms belonging to industrials, consumer discretionary, information technology, materials, and health care are recognized as pillars of Chinese industry, accounting for 78% of the total listed firms. Capital goods, materials II, technology hardware and equipment, consumer durables and apparel, software and services, pharmaceuticals, biotechnology & life, and real estate are the principal industries at the sub-industry level, making up 63% of the total listed firms.



**Figure 6.** Proportion of different industries in the first and second levels of the classification system.

Line and bar graphs are plotted in Figure 7 to further describe and understand the characteristics of the development of publicly listed firms belonging to the top five first-level industries during 1992–2016. In general, the amount of potentially listed firms in the top five sectors showed a continuous decline from Period II to Period IV. The core period for the establishment of listed companies in industrials and materials was in the earlier stage (Period I). This provided a foundation for China’s early economic development, promoting economic growth through heavy industry and infrastructure construction. Large numbers of potentially listed firms in information technology established in Periods II and III represents the vigorous development of the information technology industry. Improvements in living standards and an increase in consumption have increased spending on consumer discretionary items. Thus, after 2008, the largest proportion of newly established listed firms belonged to consumer discretionary. This is an omen of the advent of a new consumption era, suggesting that China’s economic development has entered a new period.



**Figure 7.** The development of publicly listed firms belonging to the top five first-level industries, 1992–2016.

#### 4.3.1. Global Moran’s I Statistics in the Top Five Sectors

Table 3 shows the results for Global Moran’s I for the top five sectors. The z-values are above 2.58 and are significant at the 1% level, which demonstrates that there is strong spatial clustering among the locations of headquarters in the different sectors. All values of the Global Moran’s index are larger than zero, suggesting that the spatial autocorrelation of the headquarters in different sectors are positive. There is a remarkable difference in z-values and the Global Moran index among these sectors. The spatial autocorrelation of materials and consumer discretionary industries is larger than that of health care and information technology, indicating a stronger spillover effect on surrounding cities. Meanwhile, health care and information technology tend to gather in a few regions.

Table 3. Output for Moran’s I statistic in the top five sectors.

Distance Method	p-Value	z-Score	Variance	Moran Index	Industry
Euclidean	0.000000	7.79356	0.000169	0.0988	Industrial
Euclidean	0.000000	14.3722	0.000186	0.1936	Materials
Euclidean	0.000077	3.95275	0.000141	0.0444	Information Technology
Euclidean	0.000000	8.12152	0.000165	0.1016	Consumer Discretionary
Euclidean	0.002054	3.08225	0.000164	0.0369	Health Care

### 4.3.2. Getis-Ord Gi\* Statistics in the Top Five Sectors

For deeper insight into the clustering pattern of headquarters in these five sectors, Figure 8 shows the hot and cold spot maps. Table 4 provides a summary of cities that are hot and secondary cold spots for headquarters. We can conclude the following: the primary hot spots for listed firms in the industrials sector are Beijing, Shanghai, and Shenzhen; Beijing and Shenzhen are the most-favored destinations for the headquarters of firms in information technology; the core areas for consumer discretionary and health care firms are Beijing and Shanghai; and the materials sector has more hot spots than the other sectors. Besides Beijing and Shenzhen, such types of firms prefer to locate their headquarters in the Yangtze River Delta, in cities such as Shanghai, Wuxi, Jiaying, and Suzhou.

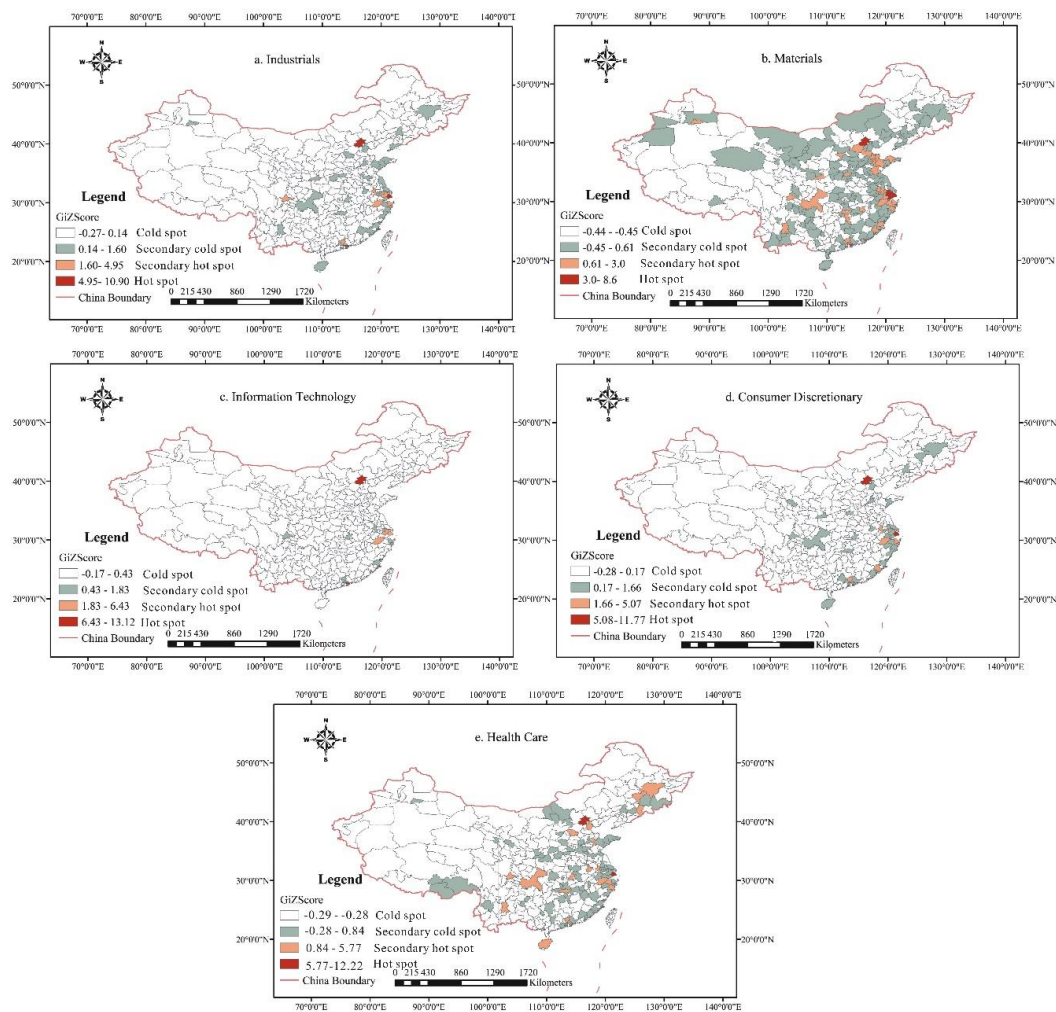


Figure 8. Getis-Ord Gi\* statistics for the headquarters of listed firms in different sectors : (a) industrials; (b) materials; (c), information technology; (d), consumer discretionary; and (e),health care.

**Table 4.** Summary of hot spot/cold spot clusters of headquarters in the top five sectors.

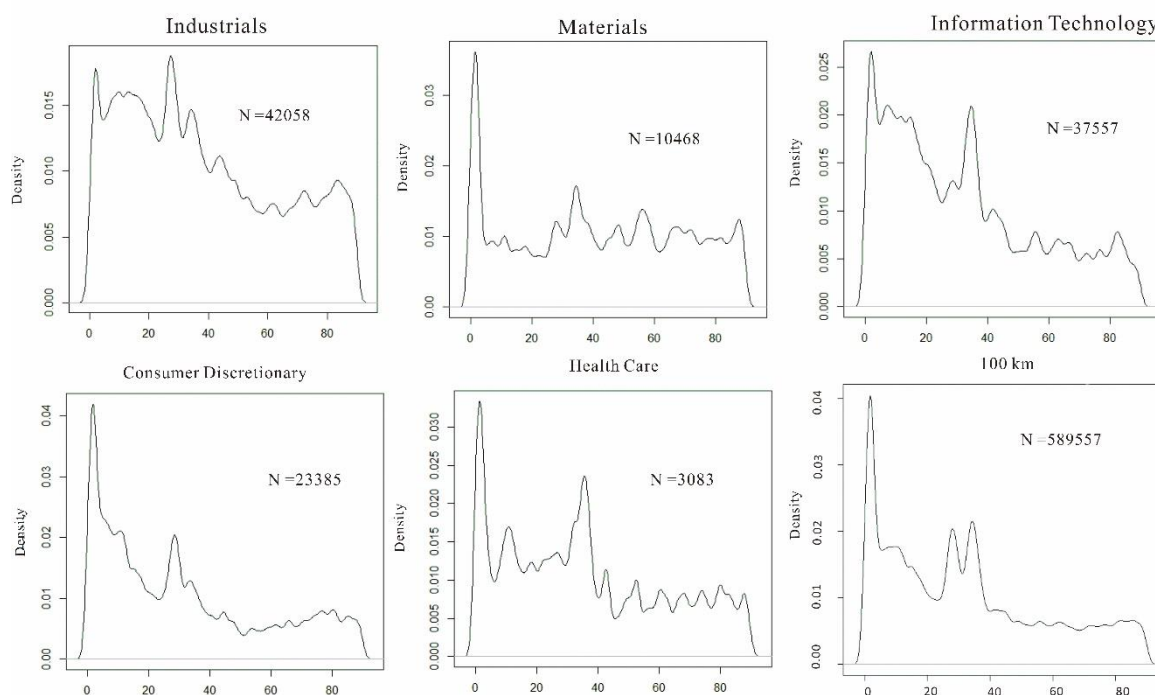
Types of Clusters	Intensity of Clusters	Industry	Number of Cities	Percentage
Hot spots	Primary	Industrials	3	0.76%
		Materials	6	1.52%
		Information Technology	2	0.51%
		Consumer Discretionary	2	0.51%
		Health Care	2	0.51%
	Secondary	Industrials	7	1.77%
		Materials	32	8.1%
		Information Technology	3	0.76%
		Consumer Discretionary	8	2.03%
		Health Care	20	5.06%
Cold spots	Secondary	Industrials	35	8.86%
		Materials	134	33.92%
		Information Technology	10	2.53%
		Consumer Discretionary	35	8.86%
		Health Care	77	19.5%

The proportion of hot spots and secondary cold spots for information technology (3.8%), industrials (11.4%), and consumer discretionary (11.4%) are extremely low, suggesting a higher degree of agglomeration than in materials and health care. All these distribution patterns are the consequence of historical factors, specific features of industries, and government actions. Giving priority to industrials, the early founding of the industrials sector was mainly located in the central areas of major cities, producing a high degree of agglomeration. Being close to consumer markets and scientific and technological talent, companies in consumer discretionary and information technology have a stronger internal impetus to gather in a few megacities with strong consumer demand and a large quantity of high-end talent. Materials and health care have a more scattered distribution. This is due to these firms preferring to locate their headquarters in regions with an integrated industrial foundation or with raw materials.

#### 4.3.3. Kernel Density Estimation with Different Radius

Figure 9 plots the kernel density distribution of headquarters classified in the top five industries and the total headquarters within 100 km. The Gaussian variant of the kernel function was used, and the bandwidth was set to 1. As shown in the kernel density estimation curve, the maximum kernel density value for the total headquarters appears in the 0–5 km interval, which indicates that headquarters tend to agglomerate in an area with a 5 km radius to obtain the benefits of a cluster economy. The second and third peak values are approximately 28 and 35 km. These two peaks may be the general distance between two separate agglomeration areas. The kernel density value falls with distance, suggesting a diminishing spatial agglomeration.

Headquarters categorized into different industries show different characteristics. With the exception of industrials, all other industries show a maximum value within a 5 km radius. In contrast to the other industries, the peak kernel density value of industrials is approximately 30 km, and the overall kernel density values tend to be low for this sector. The headquarters of consumer discretionary firms have the highest level of agglomeration within a 5 km radius, with materials headquarters ranking second. The high mean kernel density value within 20 km implies information technology has a higher degree of agglomeration than the other sectors. All of these results are consistent with the conclusion in Section 4.3.2.



**Figure 9.** Kernel density estimation curves of the total headquarters and headquarters belonging to the top five industries.

#### 4.4. Spatial Association Analysis of Industry Sectors

##### 4.4.1. Rules of Spatial Association

The spatial association rules between the headquarters of different sectors were mined on the basis of ArcGIS 10.2 (Esri, California, USA, 2014) and the Apriori algorithm in the R language. For each distance, the support threshold was set as 0.01, and the confidence threshold was set as 0.5. The association rules range from several hundred to tens of thousands, with different maxlen of frequent itemsets. The rules were sorted in descending order by lift. In this paper, only the five rules with the highest degree of lift at 5, 10, 30, and 50 km thresholds are shown (Table 5). Smaller distance thresholds have a lower degree of support and confidence and a higher degree of lift. Meanwhile, at longer distances, the degree of confidence is larger, and the degree of lift is lower.

**Table 5.** The top five rules by degree of lift at different distance thresholds.

Radius	Rules	Support	Confidence	Lift	Count
5 km	{Insurance II} ⇒ {Banks}	0.066	0.754	4.883	52
	{Household ... Personal Products} ⇒ {Banks}	0.013	0.667	4.320	10
	{Telecommunication Services II} ⇒ {Banks}	0.043	0.596	3.865	34
	{Telecommunication Services II} ⇒ {Consumer Services II}	0.057	0.789	3.706	45
	{Telecommunication Services II} ⇒ {Health Care Equipment Services}	0.052	0.719	3.460	41
10 km	{Insurance II} ⇒ {Consumer Services II}	0.135	0.986	1.629	616
	{Insurance II} ⇒ {Banks}	0.132	0.963	1.562	602
	{Telecommunication Services II} ⇒ {Food ... Staples Retailing II}	0.192	0.969	1.547	875
	{Telecommunication Services II} ⇒ {Consumer Services II}	0.183	0.922	1.525	833
	{Telecommunication Services II} ⇒ {Health Care Equipment Services}	0.197	0.993	1.522	897

Table 5. Cont.

Radius	Rules	Support	Confidence	Lift	Count
30 km	{Insurance II} ⇒ {Consumer Services II}	0.161	0.970	1.748	723
	{Telecommunication Services II} ⇒ {Health Care Equipment Services}	0.243	0.978	1.668	1087
	{Telecommunication Services II} ⇒ {Consumer Services II}	0.228	0.919	1.655	1022
	{Telecommunication Services II} ⇒ {Food Staples Retailing II}	0.235	0.947	1.654	1053
	{Insurance II} ⇒ {Banks}	0.153	0.922	1.652	687
	50 km	{Insurance II} ⇒ {Consumer Services II}	0.135	0.986	1.629
{Insurance II} ⇒ {Banks}		0.132	0.963	1.562	602
{Telecommunication Services II} ⇒ {Food Staples Retailing II}		0192	0969	1547	875
{Telecommunication Services II} ⇒ {Consumer Services II}		0183	0922	1525	833
{Telecommunication Services II} ⇒ {Health Care Equipment Services}		0197	0993	1522	897

The degree of confidence and degree of lift of “{Insurance II} ⇒ {Banks}” at the 5 km threshold are high. The business co-operation model between banks and insurance contributes to this aggregation phenomena. “{Insurance II} ⇒ {Consumer Services II}” is a set of strong association rules at distance thresholds of 10, 30, and 50 km. The top five association rules at all distance thresholds are mainly in the financial, services, and retailing industries, suggesting that these industries have close connections with other industries.

4.4.2. Frequent Itemsets

To illustrate the findings in a more intuitive manner, Figure 10 shows the graphs of the two and three frequent itemsets generated at each threshold. The shortened forms of the names of the industries used in Figure 10 are shown in Table 6. The size of each circle in Figure 10 represents the confidence value, and the color indicates the lift value. At the 5 km threshold (Figure 10a), telecommunication services II, banks, and insurance II take the central positions in both the two and three frequent itemsets, connecting other industries. In the two frequent items at both the 10 and 30 km thresholds (Figure 10b,c), headquarters are located around telecommunication services II, banks, insurance II, and food & staples retailing II. At the 10 km threshold of the three frequent items, there are two clear separate association systems. As displayed in Figure 10c,d, for the three frequent itemsets at the 30 and 50 km thresholds, headquarters tend to cluster around insurance II and consumer service II.

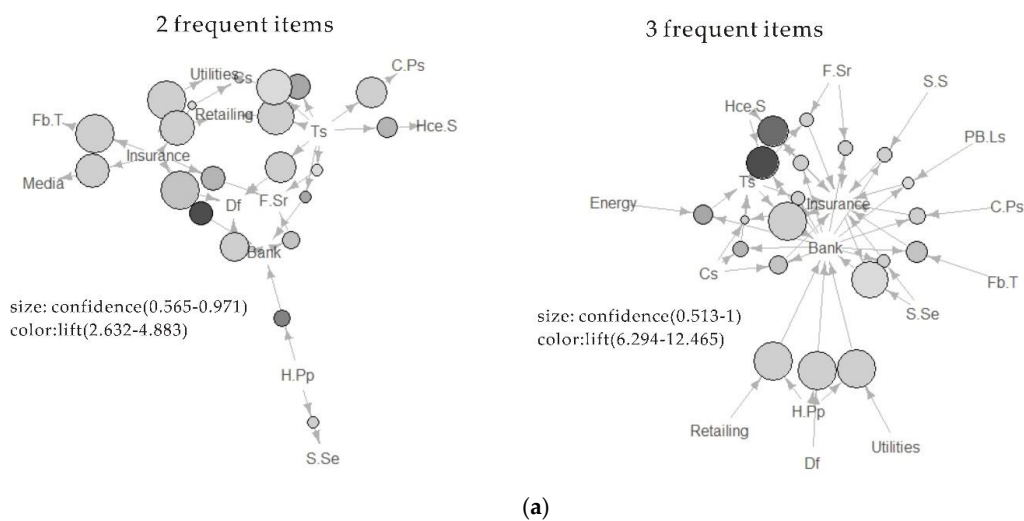
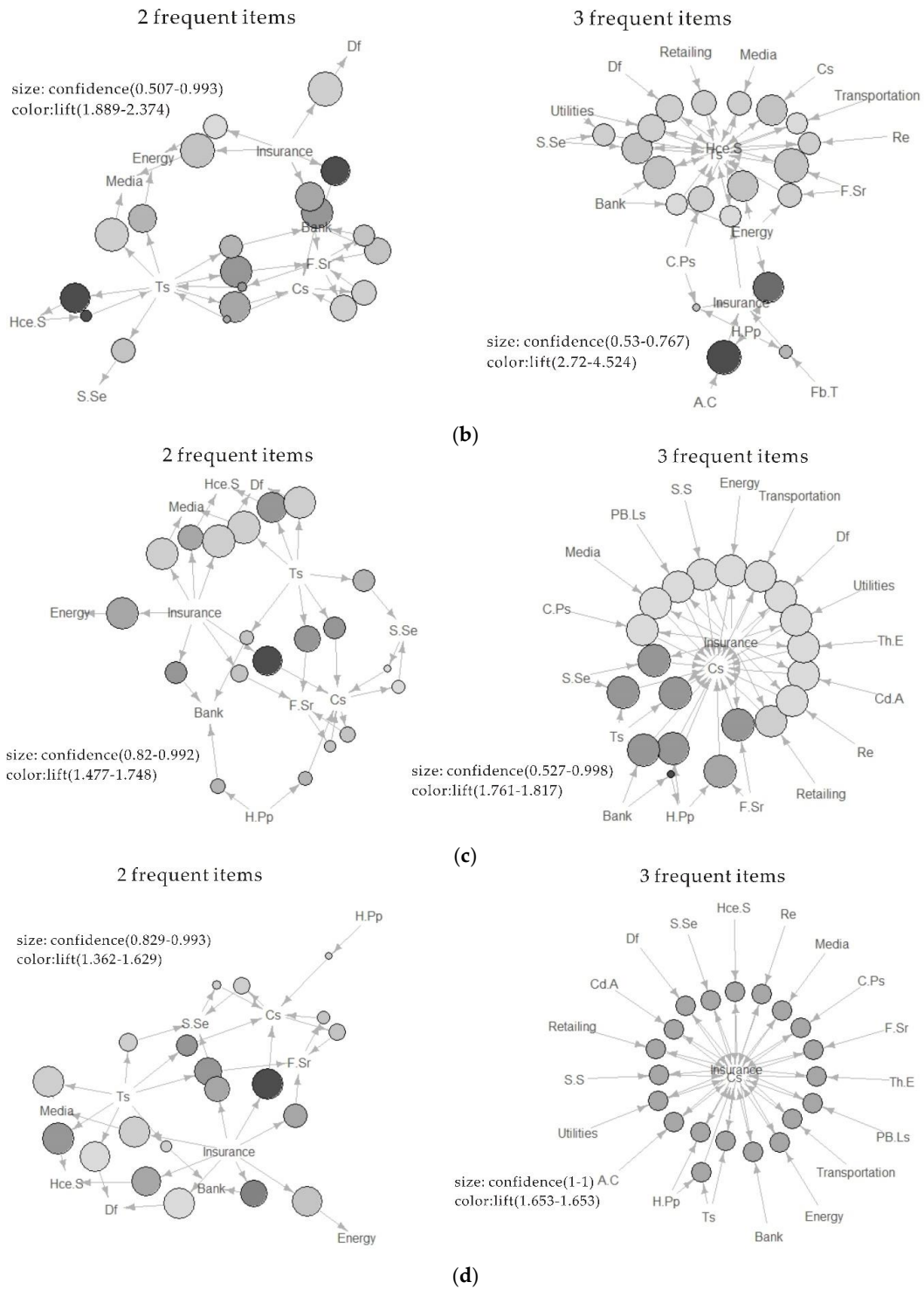


Figure 10. Cont.



**Figure 10.** (a) Two and three frequent items at the 5 km threshold; (b) Two and three frequent items at the 10 km threshold; (c) Two and three frequent items at the 30 km threshold; and (d) Two and three frequent items at the 50 km threshold.



**Table 6.** Shortened forms of the names of the 24 second-level industries.

Full Name	For Short	Full Name	For Short	Full Name	For Short
Banks	Bank	Utilities II	Utilities	Retailing	Retailing
Diversified Financials	Df	Media II	Media	Transportation	Transportation
Food & Staples	F.Sr	Software and Services	S.S	Telecommunication	Ts
Retailing II	Re	Consumer Services II	Cs	Services II	Materials
Real Estate	Cg	Energy II	Energy	Insurance II	Insurance
Capital Goods	H.Pp	Food, Beverage and Tobacco	Fb.T	Health Care Equipment and Services	Hce.S
Household and Personal Products	Cd.A	Automobiles and Components	A.C	Semiconductors and Semiconductor Equipment	S.Se
Consumer Durables and Apparel	PB.Ls	Commercial and Professional Services	C.Ps	Technology Hardware and Equipment	Th.E
Pharmaceuticals, Biotechnology and Life Sciences					

Generally, insurance II and consumer service II showed strong correlations with other industries at all distance thresholds and for all frequent items. They show a dispersive spatial distribution pattern and have no fixed aggregation model with any particular industry, being scattered around the headquarters of other sectors. Such a layout enables them to conveniently provide services for clients.

## 5. Conclusions and Policy Implications

The headquarters of listed firms are the commanding centers of an economy, and their locations can have a major influence on urban structure and local economic growth. The spatial distribution of listed firms reflects the patterns of economic development and the spatial distribution of resources in a district. Based on the spatial autocorrelation and spatial association analysis detailed in this paper, we can conclude the following distribution characteristics of the headquarters of listed firms in China:

- (1) The headquarters of listed firms in China agglomerate around megacities, especially Beijing, Shanghai, and Shenzhen, and are concentrated in certain regions along the Pacific coast, notably the Pearl River Delta region in the south, the Yangtze River Delta region in the southeast, and the Bohai Rim region in the northeast.
- (2) Headquarters belonging to firms in different sectors show different clustering patterns. Headquarters of listed companies classified as consumer discretionary and information technology display a higher degree of agglomeration than those in materials and health care. This enables them to obtain clustering benefits or take advantage of technology diffusion and knowledge spillover. Being close to raw materials or located in industrial parks with specific industrial facilities, the headquarters of materials firms show more dispersed location patterns than other sectors.
- (3) The service industries, especially insurance II and consumer service II, show strong correlations with other industries at all distance thresholds. This is so they can provide professional services to clients conveniently and quickly.

Headquarters are an important force leading the innovation and development of industry. And the “headquarters economy” can not only enhance China’s industry to the high range of the global value chain, but also prompt the development of modern service industry. This study provides preliminary grounds for understanding the geography of the headquarters of listed firms in China, revealing how headquarters classified into different sectors connect to each other. There exist significant differences in the development capacity of “headquarters economy” among major cities in China. Occupying more resources in the national resource allocation signifies the advantage of “headquarters economy” in the eastern region obvious. Overall, the Matthew Effect in headquarters locating is prominent, which indicates the task of promoting regional balanced development is still arduous.

Along with the constraints of market orientation, government guidance has played a significant role in shaping the accumulation patterns of headquarters. Given the importance of listed firms' headquarters in promoting economic growth and shaping urban structure, the results of this study can help identify interconnections between cities and can provide vital reference materials for firms deciding on future locations and for planners reshaping urban networks and constructing city group: (1) The development directions of different urban function zones for industrial undertakings should be clear, so that regional industrial division, joint development, and industrial upgrading are promoted. Development models such as "enclave economy", "joint development zone", and "service outsourcing" can be implemented in the Beijing-Tianjin-Hebei region to adjust the regional spatial economic structure. The Yangtze River Delta should make full use of the advantages of its developed industrial system to update its position in the global value chain, whereas the Pearl River Delta region should continue to deepen reform and develop a world-class manufacturing industry and advanced service industries. (2) Building a multi-industry distribution architecture incorporating leading listed firms together with small, medium-small, and microenterprises should be encouraged. This can help form a regional division and cooperation system where continuous innovation and development can be achieved through the correlation and complementarity of multiple industries. (3) High-performing listed companies are crucial for advancing innovation and development. The leading technological innovation enterprises should receive encouragement since they ultimately drive the industrialization of new advanced technologies. In addition, setting up their innovation achievement industrialization base in other cities can help promote the local economic growth. (4) Finally, the government should offer unique services to attract headquarters settled according to the local characteristics. The practical problems of professional talents should be solved effectively (e.g., settling down and children's education), in order to garner more outstanding people. The service of government in the fields of trade and business, finance, commerce, and logistics should be improved and perfected, providing headquarters more support for the development.

**Author Contributions:** Conceptualization, L.Z.; Methodology, L.Z. and H.Y.; Software, L.Z.; Validation, L.Z., H.Z. and H.Y.; Formal Analysis, L.Z.; Resources, L.Z.; Data Curation, L.Z.; Writing-Original Draft Preparation, L.Z.; Writing-Review & Editing, H.Z. and H.Y.; Visualization, H.Y.; Supervision, H.Z.; Project Administration, H.Z.

**Funding:** This research was supported by "Philosophy and Social Science Planning Project of Beijing" [15JDJGB073], and L.Z. was hosted by the School of Economy of Peking University through the Burson Scholarship.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Gao, B.; Chan, W.K.; Li, H. On the increasing inequality in size distribution of China's listed companies. *China Econ. Rev.* **2015**, *36*, 25–41. [[CrossRef](#)]
2. Cho, Y.; Honorati, M. Entrepreneurship programs in developing countries: A meta regression analysis. *Labour Econ.* **2014**, *28*, 110–130. [[CrossRef](#)]
3. Wang, C.; Madsen, J.B.; Steiner, B. Industry diversity, competition and firm relatedness: The impact on employment before and after the 2008 global financial crisis. *Reg. Stud.* **2016**, *51*, 1801–1814. [[CrossRef](#)]
4. Giovanni, J.D.; Levchenko, A.A.; Mejean, I. *Firms, Destinations, and Aggregate Fluctuations*; National Bureau of Economic Research: Cambridge, MA, USA, 2014.
5. Acemoglu, D.; Carvalho, V.M.; Ozdaglar, A.; Tahbaz-Salehi, A. The network origins of aggregate fluctuations. *Econometrica.* **2012**, *80*, 1977–2016. [[CrossRef](#)]
6. Mai, Y.; Meng, L.; Ye, Z. Regional variation in the capital structure adjustment speed of listed firms: Evidence from China. *Econ. Model.* **2017**, *64*, 288–294. [[CrossRef](#)]
7. Chen, G.; Firth, M.; Xu, L. Does the type of ownership control matter? Evidence from China's listed companies. *J. Bank. Financ.* **2009**, *33*, 171–181. [[CrossRef](#)]
8. Pan, F.; Brooker, D. Going global? Examining the geography of Chinese firms' overseas listings on international stock exchanges. *Geoforum.* **2014**, *52*, 1–11. [[CrossRef](#)]

9. Goyal, P.; Rahman, Z.; Kazmi, A.A. Corporate sustainability performance and firm performance research literature review and future research agenda. *Manag. Decis.* **2013**, *51*, 361–379. [[CrossRef](#)]
10. Zahid, M.; Ghazali, Z.; Rahman, H.U. Corporate sustainability practices among Malaysian REITs and property listed companies. *Int. J. Econ. Financ. Issues* **2016**, *6*, 688–693. [[CrossRef](#)]
11. Csomós, G. The command and control centers of the United States (2006/2012): An analysis of industry sectors influencing the position of cities. *Geoforum.* **2013**, *50*, 241–251. [[CrossRef](#)]
12. Lyons, D.I.; Salmon, S. World cities in a world-system: World cities, multinational corporations, and urban hierarchy: The case of the United States. In *World Cities in A World-System*; Knox, P.L., Taylor, P.J., Eds.; Cambridge University Press: Cambridge, MA, USA, 1995; pp. 98–114.
13. Card, D.; Hallock, K.F.; Moretti, E. The geography of giving the effect of corporate headquarters on local charities. *J. Public Econ.* **2010**, *94*, 222–234. [[CrossRef](#)]
14. Davis, J.C.; Henderson, J.V. The agglomeration of headquarters. *Reg. Sci. Urban Econ.* **2008**, *38*, 445–460. [[CrossRef](#)]
15. Pan, F.; Xia, Y. Location and agglomeration of headquarters of publicly listed firms within China's urban system. *Urban Geogr.* **2014**, *35*, 757–779. [[CrossRef](#)]
16. Rice, M.D.; Lyons, D.I. Geographies of corporate decision-making and control development, applications, and future directions in headquarters location research. *Geogr. Compass* **2010**, *4*, 320–334. [[CrossRef](#)]
17. Rice, M.D.; Lyons, D.I.; O'Hagan, S.B. Fast-growing firms as elements of change in Canada's headquarters city system. *Urban Geogr.* **2014**, *36*, 844–863. [[CrossRef](#)]
18. Lain, K. Differentiated markets: Shanghai, Beijing and Hong Kong in China's financial centre network. *Urban Stud.* **2011**, *49*, 1275–1296.
19. Jiang, Y.; Shen, J. Measuring the urban competitiveness of Chinese cities in 2000. *Cities* **2010**, *27*, 307–314. [[CrossRef](#)]
20. Jan, N. Mumbai as a global city: Theoretical essay. In *International Handbook of Globalization and World Cities*; Derudder, B., Hoyler, M., Taylor, P.J., Witlox, F., Eds.; Edward Elgar Publishing: Cheltenham, UK, 2012; pp. 447–454.
21. Ellison, G.; Glaeser, E.L.; Kerr, W.R. *What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns*; National Bureau of Economic Research: Cambridge, MA, USA, 2007; Volume 100, pp. 1195–1213.
22. Wang, J.; Pan, F.; Guo, J. The spatial pattern of the headquarters of listed enterprises in Shanghai. *Geogr. Res.* **2015**, *34*, 1920–1932. (in Chinese).
23. Goodwin, W. The management center in the United States. *Geogr. Rev.* **1965**, *55*, 1–16. [[CrossRef](#)]
24. Semple, R.K.; Phipps, A.G. The spatial evolution of corporate headquarters within an urban system. *Urban Geogr.* **1982**, *3*, 258–279. [[CrossRef](#)]
25. Semple, R.K.; Green, M.B.; Martz, D.J.F. Perspectives on corporate headquarters relocation in the United States. *Urban Geogr.* **2013**, *6*, 370–391. [[CrossRef](#)]
26. Taylor, M.J.; Thrift, N. Large corporations and concentrations of capital in Australia. *Econ. Geogr.* **1980**, *56*, 261–280. [[CrossRef](#)]
27. Bel, G.; Fageda, X. Getting there fast: Globalization, intercontinental flights and location of headquarters. *J. Econ. Geogr.* **2008**, *8*, 471–495. [[CrossRef](#)]
28. Naughton, B. China's economic policy today: The new state activism. *Eurasian Geogr. Econ.* **2011**, *52*, 313–329. [[CrossRef](#)]
29. Pan, F.; Guo, J.; Zhang, H.; Liang, J. Building a "headquarters economy": The geography of headquarters within Beijing and its implications for urban restructuring. *Cities* **2015**, *42*, 1–12. [[CrossRef](#)]
30. Tonts, M.; Taylor, M. Corporate location, concentration and performance: Large company headquarters in the Australian urban system. *Urban Stud.* **2010**, *47*, 2641–2664. [[CrossRef](#)]
31. Céline, B.-O.; Rachel, G. Changes in the intrametropolitan location of producer services in le-de-France (1978–1997): Do information technologies promote a more dispersed spatial pattern? *Urban Geogr.* **2013**, *25*, 550–578.
32. Tonts, M.; Taylor, M. The shifting geography of corporate headquarters in Australia: A longitudinal analysis. *Reg. Stud.* **2013**, *47*, 1507–1522. [[CrossRef](#)]
33. Birkinshaw, J.; Braunerhjelm, P.; Holm, U.; Terjesen, S. Why do some multinational corporations relocate their headquarters overseas? *Strateg. Manag. J.* **2006**, *27*, 681–700. [[CrossRef](#)]

34. Birkinshaw, J.; Ambos, T.C.; Bouquet, C. Boundary spanning activities of corporate HQ executives insights from a longitudinal study. *J. Manag. Stud.* **2017**, *54*, 422–454. [[CrossRef](#)]
35. Meyer, K.E.; Benito, G.R.G. Where do MNEs locate their headquarters? At home! *Glob. Strateg. J.* **2016**, *6*, 149–159. [[CrossRef](#)]
36. Pan, F.; Xia, Y.; Liu, Z. The relocation of headquarters of public listed firms in China: A regional perspective study. *Acta Geogr. Sinica.* **2013**, *68*, 449–463. (In Chinese)
37. Noyelle, T.J.; Stanback, T.M. *The Economic Transformation of American Cities*; Rowman & Allanheld: Totowa, NJ, USA, 1984.
38. Rice, M.D. Regional and sectoral growth in Canada's emerging economy. *Can. J. Reg. Sci.* **2004**, *27*, 237–254.
39. Rice, M.D.; Pooler, J.A. Subsidiary headquarters: The urban geography of the “second tier” of corporate decision-making activity in North America. *Urban Geogr.* **2013**, *30*, 289–311. [[CrossRef](#)]
40. Li, J.; Jiang, F.; Shen, J. Institutional distance and the quality of the headquarters–subsidiary relationship: The moderating role of the institutionalization of headquarters' practices in subsidiaries. *Int. Bus. Rev.* **2016**, *25*, 589–603. [[CrossRef](#)]
41. Xu, D.; Shenkar, O. Institutional distance and the multinational enterprise. *Acad. Manag. Rev.* **2002**, *27*, 608–618. [[CrossRef](#)]
42. Huang, Q.; Chand, S. Spatial spillovers of regional wages: Evidence from Chinese provinces. *China Econ. Rev.* **2015**, *32*, 97–109. [[CrossRef](#)]
43. Wang, X.; Du, L. Carbon emission performance of China's power industry: Regional disparity and spatial analysis. *J. Ind. Ecol.* **2017**, *21*, 1323–1332. [[CrossRef](#)]
44. Fallah Ghalhari, G.A.; Dadashi Roudbari, A.A.; Asadi, M. Identifying the spatial and temporal distribution characteristics of precipitation in Iran. *Arab. J. Geosci.* **2016**, *9*, 595. [[CrossRef](#)]
45. Wubuli, A.; Xue, F.; Jiang, D.; Yao, X.; Upur, H.; Wushouer, Q. Socio-demographic predictors and distribution of pulmonary tuberculosis (TB) in Xinjiang, China: A spatial analysis. *PLoS ONE* **2015**, *10*, e0144010. [[CrossRef](#)] [[PubMed](#)]
46. Peng, J.; Chen, X.; Liu, Y.; Lü, H.; Hu, X. Spatial identification of multifunctional landscapes and associated influencing factors in the Beijing-Tianjin-Hebei region, China. *Appl. Geogr.* **2016**, *74*, 170–181. [[CrossRef](#)]
47. Jiang, G.; Ma, W.; Qu, Y.; Zhang, R.; Zhou, D. How does sprawl differ across urban built-up land types in China? A spatial-temporal analysis of the Beijing metropolitan area using granted land parcel data. *Cities* **2016**, *58*, 1–9. [[CrossRef](#)]
48. Dehnad, K. *Density Estimation for Statistics and Data Analysis by Bernard Silverman*; Chapman & Hall/CRC: London, UK, 1998; Volume 39, pp. 296–297.
49. Liao, S.-H.; Chen, Y.-J.; Deng, M.-Y. Mining customer knowledge for tourism new product development and customer relationship management. *Expert Syst. Appl.* **2010**, *37*, 4212–4223. [[CrossRef](#)]
50. Bryson, A.; Forth, J.; Zhou, M. Same or different the CEO labour market in China's public listed companies. *Econ. J.* **2014**, *124*, 90–108. [[CrossRef](#)]
51. Ma, X.; Delios, A.; Lau, C. Beijing or Shanghai the strategic location choice of large mnes' host-country headquarters in China. *J. Int. Bus. Stud.* **2013**, *44*, 953–961. [[CrossRef](#)]
52. Karreman, B.; Knaap, B.V.D. The geography of equity listing and financial centre competition in mainland china and Hong Kong. *J. Econ. Geogr.* **2012**, *12*, 899–922. [[CrossRef](#)]
53. Potter, A.; Watts, H.D. Evolutionary agglomeration theory: Increasing returns, diminishing returns, and the industry life cycle. *J. Econ. Geogr.* **2010**, *11*, 417–455. [[CrossRef](#)]

