Ann Reg Sci (2014) 53:671–701
DOI 10.1007/s00168-014-0638-9

ORIGINAL PAPER

Agglomeration economies and firm R&D efforts: an analysis of China's electronics and telecommunications industries

Peng Zhang · Canfei He · Yifei Sun

Received: 18 March 2013 / Accepted: 2 September 2014 / Published online: 17 September 2014

Abstract China is on the way toward an innovation-oriented economy as well as a manufacturing powerhouse. R&D investments play a central role in improving China's industrial competitiveness. This study conducts empirical analysis to test the role of agglomeration economies in R&D efforts using plant-level data of the electronic and telecommunication equipment manufacturing industry in 2007. Statistical results suggest that both localization and urbanization economies significantly affect firm R&D. Downstream sectors are the major driving force of business R&D. Downstream firms are more likely to generate externalities than upstream ones. Upstream and midstream agglomerations even generate negative externalities due to preemption.

JEL Classification D22 · L63 · O32 · R12

1 Introduction

Recently, scholars from geography and economics have increasingly recognized the importance of agglomeration economies in innovation. Theories of the geography of innovation center on the classic conceptions of agglomeration economies. One debate is about the relative importance of localization and urbanization economies (Rosenthal

P. Zhang · C. He (✉)
College of Urban and Environmental Sciences, Peking University,
Beijing 100871, People's Republic of China
e-mail: hecanfei@urban.pku.edu.cn

P. Zhang
e-mail: pz247@cam.ac.uk

Y. Sun
Department of Geography, California State University Northridge,
Northridge, CA 91330-8249, USA
e-mail: Yifei.sun@csun.edu

and Strange 2004; Agrawal et al. 2010). Localization economies imply that firms benefit from clustering with other firms in the same sector (Glaeser et al. 1992) while urbanization economies mean that diversity can better promote innovation (Feldman 1999). In the Western literature, urbanization economies are believed to be more conducive to innovative activities (Feldman and Audretsch 1999; Henderson 1997), while evidence from China supports the localization thesis (Zhang and Li 2007).

Existing studies have placed more emphasis on the outputs of innovation (Autant-Bernard and LeSage 2011). Little research has been conducted in terms of the relationship between agglomeration economies and firm R&D inputs. The two exceptions proposed by Lee (2009) and Suarez and Walrod (1997) report a negative effect of clustering on firm R&D intensity. The effect is mediated by technological competence, absorptive capability, R&D collaboration and global networks (Lee 2009). However, they suffer from a few weaknesses. First, clustering is based on firm's self-identification about whether or not it is located in a cluster. Second, it only focuses on the clustering of firms in the same industry. The asymmetric influences of urbanization economies, especially those based on the value chain, are not investigated in the literature.

Using plant-level data of the electronic and telecommunication equipment manufacturing industry in China, this study will test the importance of agglomeration economies in R&D intensity. We measure the different aspects of agglomeration economies from industrial specialization to diversification. We then examine the agglomeration of firms in different segments in the value chain and their effects on firm R&D. We divide the whole value chain into upstream, midstream and downstream sectors based on the Standard Industrial Classification (SIC) codes in China (GB/T 4754-2002). Upstream firms include electronic components manufacturing companies, while midstream firms consist of electronic device manufacturing firms and downstream producers indicate equipment manufacturing and assembling. We also highlight the influence of institutional environments in China on firm R&D efforts. In particular, we consider the implication of weak protection of intellectual property rights and China's position in the low end of the global value chain.

Statistical results show that both localization and urbanization economies exert positive influences on firm R&D intensity in China. Downstream sectors in the value chain are the major driving force of business R&D. Clustering of downstream firms exerts positive effects, but significant effects can only be detected within downstream sectors. Upstream and midstream agglomerations even generate negative externalities due to preemption.

This paper is structured as follows. The second part develops the analytical framework, followed by the introduction of data sources. This paper then describes and analyzes the extent of agglomeration and R&D efforts in the industry. Statistical analysis and robustness checks will be presented afterward. The paper concludes with a summary of empirical findings.

2 Agglomeration economies and R&D efforts in a transitional economy: analytical framework and research hypotheses

We develop an analytical framework to explore the role of firm heterogeneity, agglomeration economies and their influence on R&D efforts in China. Specifically, we consider the influence of China's inadequate protection of intellectual property rights (IPR), its position in the low-end part of global value chains and government intervention.

2.1 Business R&D and agglomeration economies: localization or urbanization economies?

Two types of agglomeration economies have been identified, i.e., localization and urbanization economies. Localization economies refer to spillovers among concentrated firms in the same industry, while urbanization economies are benefits to firms among diverse industries. The spillovers occur through a few channels: demonstration, competition, business linkages between suppliers and providers, and labor mobility. Despite the positive influence of externalities on innovation output, the relationship between externalities and business R&D is not straightforward.

Localization economies work in the following ways. First, the demonstration effect of leading firms is inductive to other firms' innovation efforts. Once a leading firm in a concentration adopts some new technologies, other firms may follow. Second, the concentration of firms within the same industries will enhance the competition and consequently enforce firms to invest more in firm R&D to survive the competition.

However, a clustering of firms in the same industry could have negative impacts on firms' R&D efforts because of the free-riding problems associated with spillovers among firms in local clusters (Beal and Gimeno 2001). For example, labor mobility may reduce the motives to conduct R&D, since internally generated knowledge may leak to competitors. Locating in a cluster of firms in the same industry can also have a lock-in effect where firms overlook opportunities in the broad market (Pouder and John 1996). Failure to recognize the technological and market opportunities beyond the local clusters will lead to smaller R&D efforts among firms in the clusters. The relationship between localization economies and firm R&D is unclear in China.

In contrast, urbanization economies rely on the complementary relationship among firms in different industries and other factors such as universities and producer services in the urban environment. As such competitive effects are less significant, spillovers through interactions among different factors are more likely to promote firm R&D. Such R&D will help firms enhance their absorptive capabilities and capture spillovers through urbanization economies. As such, we hypothesize that urbanization economies have positive effects on firms' R&D while the impact of localization economies on firms' R&D is uncertain.

2.2 Further discussions on urbanization economies: which part of the value chain drives business R&D

Urbanization economies sometimes rely on business linkages across the value chain. In the Western context, upstream ends are found to be intensive in R&D activities due to more value-added (Mudambi 2008). In China, the intellectual property right (IPR) regulation is not well enforced. Insufficient IPR protection results in two possible consequences. First, upstream sectors lack strong incentives to conduct R&D. This is confirmed by Gao and Fu (1996) in a survey of managers, which found that unclear IPR is among the most important barriers to undertaking innovation activities. The geographical clustering of related firms increases the chance of technology leakage under weak IPR protection.

Compared with downstream sectors, upstream firms are more vulnerable to uncertainties inherent to the innovation process (Tellis and Peter 1996). As a consequence, the rational choice for upstream firms is to follow the steps of other partners rather than be the first to make innovations (Zhang et al. 2007). Evidence shows that China's manufacturing firms have increased their funding of R&D by buying or contracting for research from outside rather than making R&D efforts in-house (White 2000). Thus, with clusters of firms in the same/related industry, upstream firms are concerned with knowledge leakages because of the weak IPR protection regime. We expect that the agglomeration of upstream sectors will not have significant effects on firms' R&D intensity in China.

Second, downstream sectors, however, might make use of second-mover advantages through imitation and become more active in R&D efforts under weak IPR protection. It is easier for downstream firms, intrinsically less R&D intensive than their upstream counterparts, to get involved in imitation process, which is encouraged by insufficient IPR protection (DiMaggio and Powell 1983). As a result, agglomeration of downstream sectors might generate positive externalities by stimulating other firms' incentives of imitation. Major changes have taken place in the distribution of R&D activities in China with reforms creating strong incentives for applied research to be more responsive to downstream manufacturers (Liu and White 2001). This has strengthened the spillover effects from downstream firms. Thus, we expect that agglomeration of downstream sectors will have significant positive effects on firms' R&D intensity in China.

2.3 Heterogeneous urbanization economies along the value chain

Urbanization economies across the value chain can influence business R&D, but heterogeneous effects may exist in different sectors (Battisti and Stoneman 2003). We expect different effects of the agglomeration of upstream firms on upstream and downstream firms' R&D efforts. Spillover effects from downstream firms can also differ across the value chain.

Industrial agglomeration along the value chain can bring about many benefits due to imperfect competition (Krugman and Venables 1995). Venables (1996) argued that the production expansion in downstream firms increases the output of upstream sectors by enlarging demand of intermediate goods, contributing to the agglomeration of vertically linked industries. New breakthroughs, often originated in downstream firms, will lead to the upgrading of equipment and products in these sectors. This requires its upstream counterparts to conduct more R&D to catch up with technological improvements. This will favor the growth of other downstream industries and accelerate the overall technological progress along the value chain (Sun 2010).

China's upstream firms are more technologically advanced than downstream firms. As such, they are more likely to face the pressure from their upstream counterparts.

Spillover effects among upstream firms might not be very clear because firms facing fierce competition can be either more or less likely to conduct R&D. Spillovers from upstream to downstream sectors might be more significant. This process only happens if the suppliers can appropriate the benefits from innovation (Battisti and Stoneman 2003). Most downstream firms in China are not very innovative despite their integration into global value chains (Fan 2011). These firms at the low-value-added segments of the global value chain do not have sufficient capabilities to generate adequate knowledge spillovers to upstream factories. Downstream firms controlling product sales are strongly motivated to learn from other firms and conduct R&D in order to improve their product quality and increase their chance to be integrated into global value chains (Zhang et al. 2007).

We argue that the agglomeration of firms in the same or related industries will provide more opportunities for such downstream firms to learn from others. This includes reducing the costs of final products, improving product quality and strengthening new product development across functional teams. As a result, the clustering of downstream firms represents more learning opportunities for downstream counterparts.

However, the spillover effects from downstream to upstream firms might not be significant due to their limited cooperation with upstream suppliers in new product development. Less product similarity and larger technology gaps make interactions between upstream and downstream firms weaker than those within downstream sectors. The agglomeration of downstream firms may generate positive externalities to upstream counterparts. As a consequence, we expect that the agglomeration of downstream firms will have positive impacts on downstream firms' R&D efforts, while the clustering of upstream firms will have no impacts on upstream firms' R&D efforts. Spillovers between upstream and downstream firms are expected to be limited.

3 Data sources and model specification

3.1 Data sources

We rely on the Annual Survey of Industrial Firms (ASIF) from 2004 to 2007 in China to capture firms' R&D efforts.[1] Those are the most recent firm-level data we can obtain. ASIF is conducted by the State Statistics Bureau of China and covers all Chinese industrial state-owned enterprises and non-state-owned enterprises with annual sales of 5 million Yuan or more. The dataset provides detailed information on enterprises' location, capital structure, total employees, total shipments, exported shipments, intermediate inputs, among others. We mainly rely on data in 2007 to measure R&D inputs because information about R&D efforts is only available in 2007. To avoid possible reverse causality problems, data in 2004–2006 are also included in the analysis.

We select the electronic and telecommunication equipment manufacturing (ETEM) industry[2] for two reasons. First, it is among the five manufacturing industries that are

[1] Enterprises in this database make up over 90 % of the national manufacturing output value and each reports annual business income of more than five million. Thus, they are good representatives of China's national sample.

[2] The two-digit code for this manufacturing in database is 40.

categorized as high-tech industries in China[3] and plays a leading role in the development of China's high-tech industries. For example, in 2008, intramural R&D expenditure in ETEM industry was 4,029.384 million Yuan, which represented over 60 % of the R&D spending among the five high-tech industries.[4] Second, it is reported as the most competitive industry with the highest capacity of business innovation by Chinese Academy of Science. The number of inventive patents owned by ETEM industry was 15,418 in 2008, which was about 65 % of the total number of inventive patents in China's high-tech industries (See footnote 4). Technologically competitive firms in this industry include Huawei, Zhongxin and Lenovo (Fan 2006).

We measure R&D efforts as R&D intensity (that is, the ratio of R&D expenditure to total business income). In analogy to the literature on international trade, both localization and urbanization economies can influence R&D intensity through either intensive margin or extensive margin (Stoneman and Battisti 2010). Agglomeration increases firms' return to innovative activities by acting on the intensive margin while it affects the extensive margin by stimulating new entries and exits in business R&D. We discuss the effects of agglomeration on both intensive and extensive margins when interpreting the empirical results.[5]

3.2 Measurements of localization and urbanization economies

Agglomeration is measured by a few proxies. In accordance with Henderson (2003), firm density in a city can be used as a proxy for the localization effect. We use variable 'firm density at 4-digit level', which refers to the firm density in the same 4-digit subsector within a city, to capture the localization effect. However, firm density might not be appropriate to capture urbanization economies because it may simply reflect concentration outside certain industries rather than urban diversity. Thus, we refer to He and Pan (2010) to construct some indices for urbanization economies.

For urbanization economies from different subsectors within the same industry, we use the location quotient to capture the importance of industrial agglomeration. The location quotient is measured as the ratio of the gross industrial output of subsector s in city j to the gross industrial output of the same subsector in the whole country n, defined as follows:

$$LQ_{sj} = \frac{OUT_{sj}/OUT_j}{OUT_{sn}/OUT_n}$$

where OUT_{sj} is the gross output of subsector s in city j; OUT_j represents the gross industrial output in city j; OUT_{sn} is the gross output of the same subsector at the national level; and OUT_n is the gross industrial output at the national level. If the

[3] The other four industries are: pharmaceutical manufacturing industry, aerospace manufacturing industry, computer and office equipment industry and medical equipment and instrument manufacturing industry.

[4] Data come from *China's High-tech Industry Statistical Yearbook 2008*.

[5] Our empirical research can only present a net effect of these two features because R&D expense in 2007 is the only information on business R&D in our dataset and it is hard to identify the change in composition of firms conducting R&D with cross-sectional data.

location quotient LQ_{sj} is greater than 1, then city j has a relatively high level of concentration of subsector s.

Two variables are considered for a detailed division of urbanization economies in different subsectors within the same industry. Variable 'location quotient at 4 digit level' is to measure the location quotient of firms belonging to different subsectors but still inside the ETEM industry. 'location quotient at 2 digit level' refers to the location quotient of all subsectors within ETEM industry.

He and Pan (2010) also suggest that industrial diversity is introduced as a proxy for urbanization economies from all industries other than ETEM industry in the city. It is the inverse of a normalized Herfindhal index of industrial concentration, defined as follows:

$$DIV_j = \frac{1/\sum_{s' \neq s}^{S} \left(\frac{OUT_{s'j}}{OUT_j - OUT_{sj}} \right)^2}{1/\sum_{s' \neq s}^{S} \left(\frac{OUT_{s'n}}{OUT_n - OUT_{sn}} \right)^2}$$

where s refers to the ETEM industry; $OUT_{s'j}$ is the gross industrial output of industries other than the one that is studied in city j; and OUT_n and OUT_{sn} are the gross industrial output at the national level and the gross output of ETEM industry in China. Higher industrial diversity indicates urbanization economies among different industries.

Variable 'industrial diversity' is set up for industrial diversity of firms outside the ETEM industry. The coefficient of the first variable "firm density at 4-digit level" might be either positive or negative since the effects of localization economies on R&D efforts are not clear while the coefficients of the remaining variables in this category will be positive due to the potential urbanization economies.

Further disaggregation is used to test the heterogeneous effects of clustering along the value chain. We separate the effects of agglomeration economies from upstream, midstream and downstream counterparts and set up three variables to calculate the clustering of firms along the value chain: variable 'location quotient upstream' for location quotient of upstream producers, 'location quotient midstream' for location quotient of midstream firms and 'location quotient downstream' for local quotient of downstream plants. We will run the models separately for firms in upstream, midstream and downstream sectors to test the interactions among firms in different segments of the value chain. Only 'location quotient downstream' will have positive impacts on overall R&D intensity. Also, we expect the agglomeration of Chinese upstream firms (*location quotient upstream*) will be negative, while the concentration of downstream firms (*location quotient downstream*) will have positive impacts on R&D intensities for downstream sectors.

3.3 Model specifications

In order to examine the impacts of agglomeration economies on R&D intensity, we adopt the Tobit model rather than the OLS regressions since there are many zero values in the dependent variable (*R&D intensity*). Ignoring these zero values can result in a downward bias in estimates (Autant-Bernard and LeSage 2011).

With cross-sectional analysis, the models could be criticized for potential reverse causality and endogeneity problems. Reverse causality arises from the fact that agglomeration economies and spillovers reinforce each other. The geographical proximity of firms accelerates knowledge spillovers and innovation through externalities. Firms tend to locate in clusters to capitalize on the knowledge stock in neighboring firms (Koo 2005). Endogeneity problem arises when unobserved characteristics have impacts on firm innovation. In response to the possible reverse causality, we introduce the lag term of possible endogenous variables in regressions. As multi-phase lag effects are proven not significant based on evidence from China (Zhang et al. 2007), we first adapt the possible endogenous variables in the previous year as the lag term in our major empirical results. We use the 2007 data to measure R&D intensity while base the calculation of agglomeration on the 2006 data.

In the first robustness check, we introduce the lag values of endogenous variables in the previous two and three years (agglomeration in 2005 and 2004) to further solve the problem of possible reverse causality. We make the following improvements to minimize the impacts of endogeneity problem on the consistency of coefficients by introducing the matching method. We construct a matched sample and use conditional logit identification in accordance with Agrawal et al. (2010) to control for the underlying distribution of R&D intensity and other unobservable confounders.

To control for the pooling aspect of model implementation, we introduce industry fixed effects in all regressions. Industry fixed effects over different 4-digit subindustries are introduced, with reference to relevant studies (Autant-Bernard and LeSage 2011). Furthermore, to avoid heteroscedasticity of the disturbance terms, we rely on HSK-robust models to guarantee the consistency of estimates by assigning more weights to observations with larger variances. Furthermore, all variables are standardized to make the coefficients more comparable.

In addition, we include a number of controls for firm features and urban environment. Firm characteristics include firm size, ownership, export and age, which have been shown to affect firm-level R&D efforts. First, firm size matters for R&D inputs but it is still uncertain whether large or small firms are more willing to be engaged in R&D efforts (Feldman 1999; Koo 2005). We introduce a variable ‘firm size’ to measure firm size and only include small- and medium-sized firms to avoid multicollinearity issue.[6] This variable based on statistical classification might reflect more comprehensive information about firm size than measurements according to the number of employees or sales. Second, ownership is an important factor in the Chinese context. We consider five categories of ownership in the models: state-owned, collectively owned, privately owned, firms funded by Hong Kong, Macao and Taiwan and foreign investments.[7] Exports are captured by export intensity (the ratio between export and total output). In other countries, exports may have positive impacts on firm R&D, though in China we expect exports will have negative impacts on firm R&D. Last, we consider the variable ‘duration’ (gap between year 2007 and the year

[6] Our database has already categorized existing firms into three parts and assigned value 1 for large companies, 2 for mid-scale firms and 3 for small setups.

[7] This has already been categorized in our database.

of establishment) to capture the relationship between time and R&D intensity. Most studies have shown that young firms are more R&D intensive and innovative.

Our proxies for city characteristics, mainly used to avoid omitted variable bias, include *municipality*, *capital* and *three-line frontier*. *Municipality* refers to whether a city is one of the four provincial-level cities (Beijing, Shanghai, Tianjin and Chongqing). *Capital* refers to whether a city is a provincial capital. *Three-line frontier* refers to whether a city was once the front of "Third-Frontier Construction" during the 1960s and 1970s in China (see Part 4 for more details).

We further include a few city-level variables as controls: the overall industrial R&D intensity, the average wage and the labor pool. The average industrial R&D intensity (*average research*) is included to capture the overall urban technological sophistication. It is expected that firms in more technologically advanced cities will report higher R&D intensity. Following Bartel and Lichtenberg (1987), we include the average wage in the ETEM industry (*average human capital*) to proxy for human capital. Finally, we include the total industrial employment (*labor pool*) to control the scale of urban economy.

The Tobit model is defined as follows:

$$research_{ij} = \boldsymbol{\eta}_i + \beta_0 + \beta_1 \boldsymbol{A}_{ij} + \beta_2 \boldsymbol{X}_i + \beta_3 \boldsymbol{Z}_j + \beta_4 \boldsymbol{D}_{ij} + \boldsymbol{\varepsilon}_{ij}$$

where i represents observations and j is the corresponding city. $research_{ij}$ indicates R&D intensity in 2007 for firm i in city j. Industry fixed effect is captured by η_i. Matrix A_{ij} includes proxies for agglomeration economies. X_i refers to firm characteristics, including size, ownership, duration and market structure. Z_j includes *average human capital*, *average city research* and *labor pool*. D_{ij} is a set of dummy variables proxy for innovative environments. $\varepsilon_i j$ is the stochastic error term.

The Tobit model for the agglomeration of the value chain is as follows:

$$research_{ij} = \boldsymbol{\eta}_i + \beta_0 + \beta_1 \boldsymbol{A\_chain}_{ij} + \beta_2 \boldsymbol{X}_i + \beta_3 \boldsymbol{Z}_j + \beta_4 \boldsymbol{D}_{ij} + \boldsymbol{\varepsilon}_{ij}$$

where variable A_chain_{ij} indicates location quotient of firms in different segments of the value chain. We test whether it is upstream or midstream or downstream sector that drives China's business R&D intensity.

Tests of heterogeneous effects of agglomeration across segments of the value chain consist of three sets of models conditional on the value chain:

$$research\_chain_{ij} = \boldsymbol{\eta}_i + \beta_0 + \beta_1 \boldsymbol{A\_chain}_{ij} + \beta_2 X_i + \beta_3 \boldsymbol{Z}_j + \beta_4 \boldsymbol{D}_{ij} + \boldsymbol{\varepsilon}_{ij}$$

where variables $research\_chain$ and A_chain_{ij} are defined to separately measure R&D intensities and the agglomeration of firms in the upstream, midstream and downstream sections along the value chain. Industry fixed effects are still needed here because R&D efforts might be far from homogenous even within each segment. Similarly, we consider the lag term of location quotient and use the 2006 data to measure industrial agglomeration.

4 Spatial pattern of ETEM industries and R&D efforts in China

4.1 Spatial pattern of China's ETEM industries

Since 1980, the economic reform has spurred the development of China's ETEM industry. From the 1990s, the rapid growth of information and computer industries has played the leading role in China's ETEM industries (Meng and Li 2001). Global leading firms that set their operations in China in the 1980s and 1990s include Cisco, Ericsson, Lucent Technologies, Motorola, Nokia, Nortel Networks and Siemens. Domestic firms, led by Huawei, ZTE, DTT and GDT, also developed quickly after the 1980s (Fan 2006).

Industrial agglomeration has appeared as a result of the industrial growth. Beijing, Shanghai, Jiangsu and Guangdong contributed 72 % of the national manufacturing output and 80 % of exports in 2007 (Zhou et al. 2011). In Fig. 1, we measure agglomeration by the density of firms in ETEM industry in year 2007 (the number of firms divided by the area of a prefectural level city). China's ETEM industry clusters are located in regions such as Beijing, Shanghai–Suzhou and Shenzhen–Dongguan, in pursuit of large-scale economies and strong industrial linkages (He and Zhu 2008).

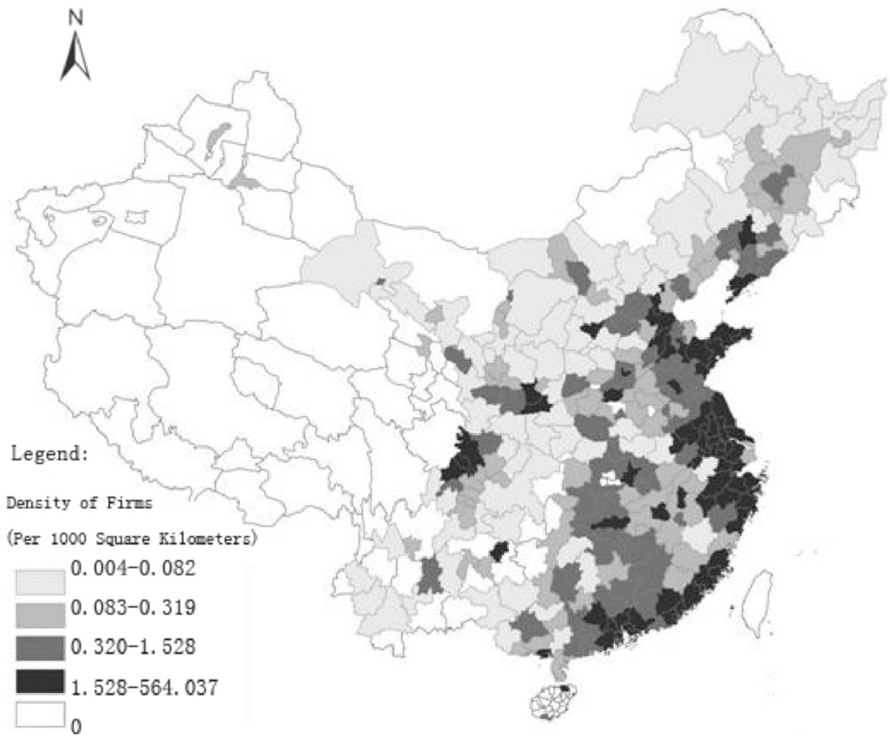

Fig. 1 Spatial distribution of electronic and telecommunication equipment manufacturing (ETEM) industry in China (2007). *Data sources*: Annual Survey of Industrial Firms (ASIF)

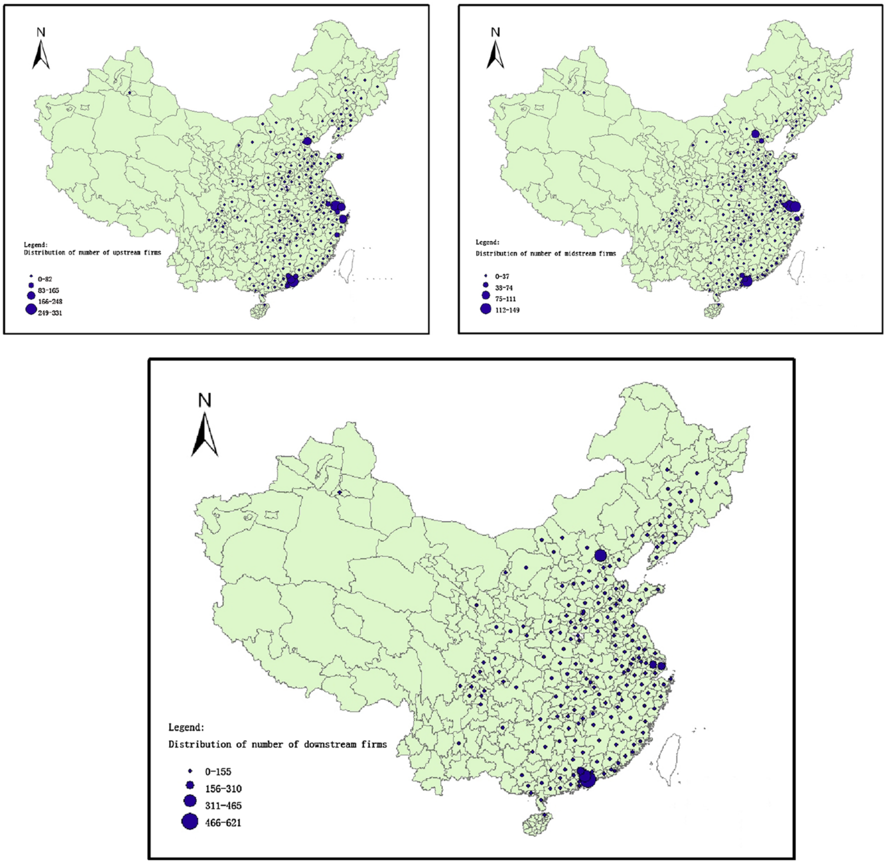

Fig. 2 Spatial distribution of upstream, midstream and downstream electronic and telecommunication equipment manufacturing (ETEM) industry (2007). *Data sources*: Annual Survey of Industrial Firms (ASIF)

High densities of firms also appear in certain inland provinces and the Northeast China, including Hubei, Sichuan, Shannxi and Liaoning, which can be seen as evidence of "path dependence." With many state-owned enterprises in the electronic industry in the prereform period, cities in the Northeast China benefit substantially from the long-standing government supports. Sichuan, Shannxi and Hubei, the forefront of the "Third-Frontier Construction," have a tradition of advanced development of military industries. Large investments in human capital and scientific research during the "Third-Frontier Construction" period equipped these regions well with necessary capital and technicians for innovation during the transition from military industries to civilian-oriented production.

Spatial distribution of upstream, midstream and downstream segments in China's ETEM industries can be seen in Fig. 2. Still, Beijing, Shanghai and Guangdong have the highest level of agglomeration of firms in all three segments. However, Shanghai and Jiangsu have the largest number of midstream firms while downstream sectors concentrate in Guangdong (Meng and Li 2001). Moreover, some cities in West China also

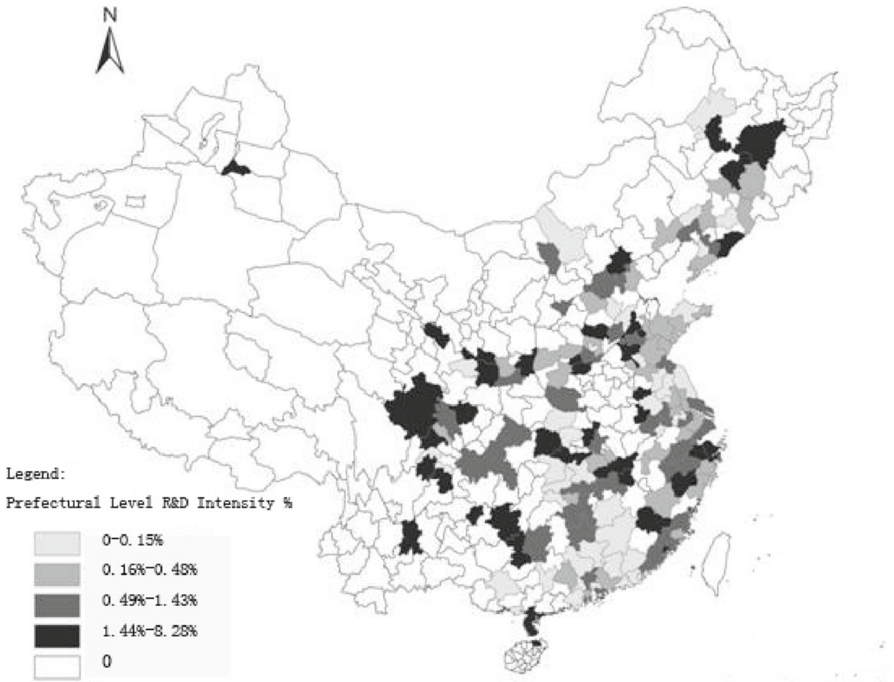

Fig. 3 Distribution of R&D intensity among cities in electronic and telecommunication equipment manufacturing (ETEM) industry in China (2007). *Data sources*: Annual Survey of Industrial Firms (ASIF). *Note* R&D intensity is the ratio of R&D expenditure to total business income for each firm

have their comparative advantages in ETEM industries. For example, Xi'an is strong in the photoelectron. Chengdu and Chongqing have the concentration of software and manufacturing, respectively (Yang 2006).

4.2 Spatial pattern of R&D efforts

The spatial pattern of R&D intensity in ETEM industries is shown in Fig. 3. Provinces with the highest average R&D intensity include Beijing and Gansu, Xinjiang, Guizhou, Yunnan and Heilongjiang. In contrast, coastal provinces such as Jiangsu, Zhejiang and Guangdong, with high density of firms, are paradoxically corresponding to lower R&D intensities. Beijing tops all regions in the R&D intensity owing to its accumulation of scientific resources and technicians. High R&D intensity in inland provinces is associated with the larger proportion of state-owned enterprises. Their R&D activities are strongly encouraged and supported by the governments. Their R&D intensities can also be traced back to the 'Third-Frontier Construction' period when rapid technological progress took place in the military factories (Yang 2006).

Lower R&D intensities in cities in the Yangtze River Delta and Pearl River Delta are related to the high concentration of foreign invested firms and their positions in the low end of global value chains. Many firms in the two major concentrations adopt the

OEM strategy, which discourage firms from investing in R&D activities. Moreover, R&D intensities in provincial capitals are relatively high in the West China, which can be explained by advanced universities and institutions in these provinces (Ponds and Oort 2010).

5 Agglomeration economies and firm R&D efforts

5.1 Localization economies, urbanization economies and R&D efforts

Table 1 reports regression results about the effects of localization and urbanization economies on firm R&D efforts. Localization economies represented by *firm density at 4-digit level* have positive effects on firm R&D efforts. The coefficient is 1.118 and significant at 0.01 level. This means spillovers from firms in the same subsector outride competition among firms in China's ETEM industries. The two proxies for urbanization economies, concentration of firms in the other ETEM industries (*location quotient at 4-digit level*) and concentration of firms in the whole ETEM industry (*location quotient at 2-digit level*), also show significant and positive impacts on R&D intensity. However, the concentration of firms in non-ETEM industries (*industrial diversity*) has no significant effects. This shows that although both localization and urbanization economies exert positive impacts on business R&D in China, influences of urbanization economies decline with the decrease in industrial similarities.

Firm heterogeneity also affects R&D inputs. Large firms tend to be more R&D intensive as both the "middle" and "small" variables show significant and negative impacts on firm R&D. Larger firms have more resources available for R&D and receive strong governmental support (He and Zhu 2007), which can add to their R&D budgets. Sources of innovative efforts in small firms are the close cooperations within clusters (Koo 2005), which is not the common case in China where mutual learning among small firms and spillovers are not that significant. Neither the vitality for innovation nor the spirit of adventure is sufficient for new Chinese firms to be actively involved in R&D efforts in the early phases. Mature firms with routinized production methods, professional technicians and adequate funding are likely to be more R&D intensive.

Foreign firms are less R&D intensive than domestic firms. More export-oriented firms are less R&D intensive. The results also show that older firms are more R&D intensive than young firms. Among the city-related variables, provincial capitals and cities' overall industrial R&D intensities have a positive and significant impact on firm R&D.

5.2 Urbanization economies, value chain and R&D efforts

Further investigation of urbanization economies reveals that upstream, midstream and downstream clusters play different roles in driving firms' R&D efforts (Table 2). Agglomeration economies of downstream firms are effective, shown by the significant corresponding coefficient. Upstream and midstream firms, however, generate negative externalities (with coefficients of −0.177 and −0.491). This is in accordance with current literature indicating that there are strong incentives for research in downstream

Table 1 Agglomeration economies measured at the sector level and R&D intensity in the electronics sector

Variables	[1] R&D intensity	[2] R&D intensity	[3] R&D intensity	[4] R&D intensity
Agglomeration economies				
Firm density at 4-digit level	1.118***			
	(0.270)			
Location quotient at 4-digit level		0.842***		
		(0.242)		
Location quotient at 2-digit level			0.807***	
			(0.242)	
Industrial diversity				0.0826
				(0.154)
Firm characteristics				
Duration	0.326***	0.345***	0.344***	0.350***
	(0.0982)	(0.0982)	(0.0982)	(0.0978)
Export	−0.128	−0.132	−0.131	−0.108
	(0.134)	(0.133)	(0.133)	(0.133)
HK and Taiwan	−2.738***	−2.935***	−2.934***	−2.829***
	(0.418)	(0.425)	(0.425)	(0.418)
Foreign	−2.843***	−3.072***	−3.074***	−3.045***
	(0.422)	(0.423)	(0.423)	(0.422)
Middle size	−1.718***	−1.621***	−1.623***	−1.641***
	(0.454)	(0.450)	(0.450)	(0.452)
Small size	−5.415***	−5.187***	−5.190***	−5.215***
	(0.618)	(0.608)	(0.608)	(0.610)
Nature of cities				
Municipality	−0.560	−0.460	−0.459	−0.543
	(0.463)	(0.460)	(0.460)	(0.472)
Capital	0.629*	0.881***	0.873**	0.671*
	(0.342)	(0.341)	(0.341)	(0.376)
Three-line frontier	−0.848**	−0.759*	−0.770*	−0.884**
	(0.404)	(0.399)	(0.399)	(0.395)
Average city research	2.759***	2.753***	2.756***	2.829***
	(0.298)	(0.300)	(0.300)	(0.303)
Average human capital	−0.0210	−0.0246**	−0.0243**	−0.0164*
	(0.0168)	(0.0116)	(0.0115)	(0.00934)
Labor pool	−0.604**	−0.209	−0.181	0.462***
	(0.278)	(0.232)	(0.232)	(0.137)
Industry FE	Yes	Yes	Yes	Yes
Constant	2.925***	2.914***	2.903***	2.687***
	(0.938)	(0.867)	(0.866)	(0.857)

Table 1 continued

Variables	[1] R&D intensity	[2] R&D intensity	[3] R&D intensity	[4] R&D intensity
Sigma	8.801***	8.794***	8.794***	8.792***
	(0.833)	(0.833)	(0.832)	(0.832)
Observations	11,045	11,059	11,059	11,081

Tobit regressions with robust standard errors
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 2 Agglomeration economies measured at the value chain and R&D intensity in the electronics sector

Variables	[1] R&D intensity	[2] R&D intensity	[3] R&D intensity
Agglomeration economies			
Location quotient upstream	−0.177		
	(0.160)		
Location quotient midstream		−0.491**	
		(0.245)	
Location quotient downstream			1.011***
			(0.222)
Firm characteristics			
Duration	0.345***	0.341***	0.344***
	(0.0981)	(0.0980)	(0.0982)
Export	−0.0898	−0.0913	−0.119
	(0.133)	(0.133)	(0.133)
HK and Taiwan	−2.786***	−2.741***	−2.907***
	(0.423)	(0.419)	(0.422)
Foreign	−3.007***	−2.955***	−3.028***
	(0.424)	(0.424)	(0.421)
Middle size	−1.626***	−1.676***	−1.598***
	(0.452)	(0.454)	(0.450)
Small size	−5.224***	−5.316***	−5.200***
	(0.610)	(0.614)	(0.609)
Nature of cities			
Municipality	−0.506	−0.136	−0.481
	(0.468)	(0.473)	(0.459)
Capital	0.631*	0.718**	0.831**
	(0.346)	(0.336)	(0.339)
Three-line frontier	−0.820**	−0.767*	−0.804**
	(0.401)	(0.397)	(0.397)
Average city research	2.814***	2.787***	2.675***
	(0.300)	(0.299)	(0.299)

Table 2 continued

Variables	[1] R&D intensity	[2] R&D intensity	[3] R&D intensity
Average human capital	−0.0147	−0.0139	−0.0261**
	(0.00896)	(0.00879)	(0.0120)
Labor pool	0.516***	0.743***	−0.352
	(0.150)	(0.209)	(0.221)
Industry fixed effects	Yes	Yes	Yes
Constant	2.601***	2.570***	2.946***
	(0.844)	(0.844)	(0.871)
Sigma	8.799***	8.797***	8.794***
	(0.832)	(0.833)	(0.833)
Observations	11,059	11,059	11,059

Tobit regressions with robust standard errors
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

segments (Liu and White 2001). There is another reason for the fact that motivation for R&D efforts comes mainly from downstream sectors. The peculiar pattern of division of labor for R&D efforts is inherited from the centrally planning system, in which they had already been involved in many downstream activities (Gu and Lundvall 2006).

It is interesting to note that the agglomerations of upstream and midstream segments have negative externalities on R&D efforts even if value-added is becoming increasingly concentrated at these sectors (Mudambi 2008). Activities at the upper segments of the value chain are not R&D intensive, confirming the second-mover strategy implemented by upstream and midstream firms in China. This is more closely associated with imitation rather than creating new technology (Liu and White 2001). This reflects the perception among firm managers that buying research from outside is more cost effective than developing new technology in-house (White 2000). Finally, concerned about IPR, technologically advanced firms in the upper end of value chain do not have active interactions with other firms in the same city and have not developed the capabilities to coordinate internal R&D activities and external opportunities afforded by the clustering of other firms.

5.3 Input/output linkages: heterogeneous effects across value chain

Table 3 shows statistical results on different sectors along the value chain. The research hypotheses are verified because heterogeneous effects of agglomeration exist across the value chain. Downstream firms generate significant and positive externalities not only to downstream counterparts but also to midstream sections with similar magnitudes. Conversely, the agglomeration of upper end firms does not significantly influence R&D intensity in the lower end of the value chain. The upstream and midstream sectors even generate negative externalities to their counterparts in the lower end (with coefficients of −1.244 and −1.257).

Table 3 Heterogeneous effects of agglomeration economies across the value chain and R&D intensity in the electronics sector

Variables	[1] Upstream	[2] Upstream	[3] Upstream	[4] Midstream	[5] Midstream	[6] Midstream	[7] Downstream	[8] Downstream	[9] Downstream
Agglomeration economies									
Location quotient upstream	−0.0647			−1.244**			−0.383		
	(0.154)			(0.570)			(0.263)		
Location quotient midstream		−0.0838			−0.642			−1.257***	
		(0.163)			(0.613)			(0.338)	
Location quotient downstream			0.229			1.503**			1.408***
			(0.235)			(0.721)			(0.359)
Firm characteristics									
Duration	0.250**	0.251**	0.255**	0.310	0.293	0.303	0.394***	0.377***	0.379***
	(0.103)	(0.103)	(0.103)	(0.262)	(0.263)	(0.263)	(0.140)	(0.140)	(0.141)
Export	0.142	0.136	0.122	−0.305	−0.371	−0.423	−0.767***	−0.760***	−0.792***
	(0.144)	(0.142)	(0.143)	(0.387)	(0.383)	(0.385)	(0.219)	(0.218)	(0.219)
HK and Taiwan	−1.790***	−1.795***	−1.846***	−2.950***	−3.038***	−3.186***	−3.092***	−2.998***	−3.243***
	(0.396)	(0.392)	(0.387)	(0.977)	(0.975)	(0.978)	(0.675)	(0.667)	(0.679)
Foreign	−1.942***	−1.938***	−1.947***	−2.703***	−2.838***	−2.827***	−2.906***	−2.767***	−2.986***
	(0.387)	(0.385)	(0.384)	(1.025)	(1.018)	(0.993)	(0.599)	(0.592)	(0.601)
Middle size	−1.241**	−1.232**	−1.239**	−2.975**	−3.477***	−3.258**	−2.034***	−2.092***	−1.946***
	(0.504)	(0.504)	(0.504)	(1.279)	(1.293)	(1.289)	(0.660)	(0.660)	(0.649)
Small size	−4.764***	−4.758***	−4.755***	−5.294***	−5.803***	−5.519***	−5.269***	−5.398***	−5.146***
	(0.701)	(0.702)	(0.703)	(1.446)	(1.472)	(1.456)	(0.892)	(0.899)	(0.876)
Nature of cities									
Municipality	−0.944*	−0.868	−0.929*	−1.224	−0.216	−0.484	0.288	1.207*	0.488
	(0.543)	(0.551)	(0.536)	(1.467)	(1.369)	(1.363)	(0.664)	(0.681)	(0.675)

Table 3 continued

Variables	[1] Upstream	[2] Upstream	[3] Upstream	[4] Midstream	[5] Midstream	[6] Midstream	[7] Downstream	[8] Downstream	[9] Downstream
Capital	0.00350	0.0478	0.0800	−0.220	0.263	0.440	1.197**	1.350***	1.606***
	(0.416)	(0.409)	(0.413)	(1.053)	(1.102)	(1.109)	(0.484)	(0.462)	(0.479)
Three-line frontier	−0.141	−0.135	−0.189	−0.938	−0.561	−0.534	−0.694	−0.611	−0.522
	(0.434)	(0.434)	(0.424)	(1.180)	(1.164)	(1.177)	(0.585)	(0.578)	(0.582)
Average city research	1.898***	1.900***	1.880***	4.195***	4.231***	4.081***	2.677***	2.609***	2.508***
	(0.371)	(0.369)	(0.378)	(0.987)	(1.007)	(1.049)	(0.347)	(0.342)	(0.335)
Average human capital	0.00743	0.00689	0.00243	−0.0405	−0.0525	−0.0860*	−0.0204	−0.0168	−0.0403
	(0.0145)	(0.0142)	(0.0145)	(0.0354)	(0.0459)	(0.0514)	(0.0155)	(0.0139)	(0.0260)
Labor pool	−0.0885	−0.0527	−0.273	1.277**	1.132*	−0.346	0.953***	1.510***	−0.303
	(0.167)	(0.198)	(0.232)	(0.497)	(0.592)	(0.754)	(0.250)	(0.312)	(0.353)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	−0.682	−0.695	−0.542	−0.426	0.403	1.179	−0.395	−0.609	−0.0320
	(0.633)	(0.635)	(0.634)	(1.636)	(1.732)	(1.879)	(0.861)	(0.843)	(1.025)
Sigma	5.264***	5.263***	5.262***	10.91***	10.91***	10.91***	9.471***	9.466***	9.467***
	(0.594)	(0.594)	(0.596)	(1.834)	(1.836)	(1.843)	(1.237)	(1.238)	(1.237)
Observations	4,675	4,675	4,675	1,650	1,650	1,650	4,734	4,734	4,734

Tobit regressions with robust standard errors
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Downstream firms are more likely to exert knowledge outflows because they bet on their market power and possess strong capabilities to integrate technological progress along the value chain. However, this only happens within downstream sections while knowledge flows from downstream to upstream and midstream firms are limited. This is because producers are more active in absorbing knowledge inflows when they are located in the same part of the value chain due to tight contacts within integrated business. As firms, consumers, suppliers and partners are highly connected with each other (Gordon and McCann 2005), frequent exchanges are conducive to the improvement of product quality and the comprehension of market trend. Firms within the same subsector are often the most likely to be exposed to such contacts and transfers of ideas. Therefore, they have the potential to experience fast technological progress.

Similarly interesting, upstream firm R&D efforts are affected by the clustering of upstream, midstream or downstream counterparts. This is consistent with the statement that China's economy still cannot embark upon technological upgrading (Gu and Lundvall 2006). For technological upgrading, vertically related firms need to coordinate their activities by exchanging information, including the quantities available and the precise technical characteristics of the products to be exchanged (Gu and Lundvall 2006). However, upstream firms in China just focus on providing raw materials or components. Their competitiveness does not lie in their R&D efforts. They are the least innovative part of Chinese economies.

It is surprising that the agglomerations of upstream and midstream sectors discourage R&D efforts of firms in the downstream part. One possible explanation is that firms in the upper end of the value chain follow the strategy of preemption when faced with competition from downstream sectors. If two firms decide how much to invest in R&D in the next period, then the innovator following a preemption strategy may overinvest to deter the other firms from investing.[8] This can well deter inter-firm diffusion of knowledge (Stoneman and Battisti 2010). This crowd-out effect is more significant when upstream/midstream firms concentrate within a certain area because agglomeration will lead to more competition and more incentives for preemption and deterrence.

6 Robustness check for possible reverse causality and endogeneity problems

6.1 Possible reverse causality

Since one-year lag might not be enough to mitigate possible reverse causality, we introduce multi-phase lag terms as the first robustness check. In response to this, we measure agglomeration economies using the data in 2005 and 2004 and run the same regressions. Significance and magnitude of coefficients on proxies for agglomeration (*firm density at 4-digit level, location quotient at 4-digit level, location quotient at 2-digit level* and *industrial diversity*) vary little between Tables 1 and 4. Comparing the results between Tables 2 and 5 reveals that downstream firms still generate positive and significant spillovers with similar magnitudes.

[8] This does not contradict the fact that upstream sectors are the least innovative. Upstream firms may invest less if there are no strategies of preemption.

Table 4 Agglomeration economies measured at the sector level and R&D intensity in the electronics sector: year 2005 and year 2004

Variables	[1] R&D intensity 05	[2] R&D intensity 05	[3] R&D intensity 05	[4] R&D intensity 05	[5] R&D intensity 04	[6] R&D intensity 04	[7] R&D intensity 04	[8] R&D intensity 04
Agglomeration economies								
Firm density at 4 digit level	1.067***				1.269***			
	(0.256)				(0.277)			
Location quotient at 4 digit level		0.745***				0.639***		
		(0.225)				(0.216)		
Location quotient at 2 digit level			0.693***				0.588***	
			(0.222)				(0.213)	
Industrial diversity				0.184				0.208
				(0.166)				(0.165)
Firm characteristics								
Duration	0.320***	0.339***	0.339***	0.350***	0.317***	0.337***	0.337***	0.351***
	(0.0982)	(0.0981)	(0.0981)	(0.0979)	(0.0982)	(0.0982)	(0.0982)	(0.0981)
Export	−0.127	−0.126	−0.124	−0.107	−0.110	−0.133	−0.131	−0.104
	(0.134)	(0.133)	(0.133)	(0.133)	(0.133)	(0.133)	(0.133)	(0.133)
HK and Taiwan	−2.762***	−2.956***	−2.952***	−2.828***	−2.741***	−2.946***	−2.941***	−2.824***
	(0.418)	(0.426)	(0.426)	(0.418)	(0.418)	(0.426)	(0.426)	(0.418)
Foreign	−2.845***	−3.097***	−3.095***	−3.053***	−2.823***	−3.094***	−3.091***	−3.048***
	(0.422)	(0.425)	(0.425)	(0.423)	(0.420)	(0.425)	(0.425)	(0.422)
Middle size	−1.720***	−1.625***	−1.626***	−1.631***	−1.682***	−1.637***	−1.637***	−1.630***
	(0.454)	(0.450)	(0.450)	(0.452)	(0.454)	(0.450)	(0.450)	(0.452)

Table 4 continued

Variables	[1] R&D intensity 05	[2] R&D intensity 05	[3] R&D intensity 05	[4] R&D intensity 05	[5] R&D intensity 04	[6] R&D intensity 04	[7] R&D intensity 04	[8] R&D intensity 04
Small size	−5.407***	−5.179***	−5.181***	−5.206***	−5.378***	−5.186***	−5.187***	−5.209***
	(0.617)	(0.607)	(0.607)	(0.610)	(0.616)	(0.608)	(0.608)	(0.610)
Nature of cities								
Municipality	−0.591	−0.472	−0.469	−0.717	−0.626	−0.433	−0.432	−0.747
	(0.464)	(0.459)	(0.459)	(0.488)	(0.464)	(0.459)	(0.459)	(0.491)
Capital	0.633*	0.848**	0.835**	0.577	0.592*	0.797**	0.779**	0.511
	(0.342)	(0.341)	(0.340)	(0.384)	(0.342)	(0.340)	(0.339)	(0.404)
Three-line frontier	−0.846**	−0.729*	−0.747*	−0.946**	−0.803**	−0.720*	−0.731*	−0.995**
	(0.404)	(0.399)	(0.399)	(0.399)	(0.404)	(0.401)	(0.401)	(0.403)
Average city research	2.756***	2.773***	2.777***	2.824***	2.716***	2.797***	2.800***	2.835***
	(0.298)	(0.300)	(0.300)	(0.302)	(0.298)	(0.300)	(0.300)	(0.300)
Average human capital	−0.0230	−0.0231**	−0.0227**	−0.0174*	−0.0261	−0.0228**	−0.0224**	−0.0181*
	(0.0170)	(0.0112)	(0.0111)	(0.00952)	(0.0170)	(0.0112)	(0.0110)	(0.00969)
Labor pool	−0.530**	−0.124	−0.0823	0.472***	−0.730**	−0.0330	0.00659	0.463***
	(0.263)	(0.217)	(0.215)	(0.137)	(0.285)	(0.209)	(0.207)	(0.137)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	3.033***	2.924***	2.908***	2.748***	3.118***	2.941***	2.923***	2.797***
	(0.940)	(0.864)	(0.863)	(0.861)	(0.940)	(0.864)	(0.863)	(0.866)
Sigma	8.801***	8.794***	8.794***	8.791***	8.802***	8.794***	8.795***	8.793***
	(0.834)	(0.833)	(0.833)	(0.832)	(0.834)	(0.833)	(0.833)	(0.832)
Observations	11,025	11,038	11,038	11,081	11,004	11,016	11,016	11,076

Tobit regressions with robust standard errors
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 5 Agglomeration economies measured at the value chain and R&D intensity in the electronics sector: year 2005 and year 2004

Variables	[1] R&D intensity 05	[2] R&D intensity 05	[3] R&D intensity 05	[4] R&D intensity 04	[5] R&D intensity 04	[6] R&D intensity 04
Agglomeration economies						
Location quotient upstream	−0.182			−0.211		
	(0.154)			(0.146)		
Location quotient midstream		−0.310			−0.245	
		(0.218)			(0.204)	
Location quotient downstream			0.873***			0.799***
			(0.209)			(0.202)
Firm characteristics						
Duration	0.341***	0.340***	0.339***	0.339***	0.340***	0.337***
	(0.0980)	(0.0980)	(0.0981)	(0.0981)	(0.0982)	(0.0982)
Export	−0.0858	−0.0948	−0.117	−0.0881	−0.100	−0.128
	(0.133)	(0.133)	(0.133)	(0.133)	(0.133)	(0.133)
HK and Taiwan	−2.793***	−2.781***	−2.940***	−2.780***	−2.794***	−2.941***
	(0.423)	(0.421)	(0.424)	(0.423)	(0.421)	(0.424)
Foreign	−3.017***	−3.003***	−3.066***	−3.011***	−3.023***	−3.073***
	(0.425)	(0.425)	(0.423)	(0.424)	(0.424)	(0.423)
Middle size	−1.629***	−1.670***	−1.608***	−1.637***	−1.654***	−1.625***
	(0.452)	(0.453)	(0.450)	(0.452)	(0.452)	(0.450)
Small size	−5.229***	−5.275***	−5.186***	−5.241***	−5.245***	−5.193***
	(0.610)	(0.612)	(0.608)	(0.611)	(0.611)	(0.608)
Nature of cities						
Municipality	−0.497	−0.251	−0.487	−0.519	−0.294	−0.434
	(0.466)	(0.468)	(0.459)	(0.468)	(0.462)	(0.459)

Table 5 continued

Variables	[1] R&D intensity 05	[2] R&D intensity 05	[3] R&D Intensity 05	[4] R&D intensity 04	[5] R&D intensity 04	[6] R&D intensity 04
Capital	0.615*	0.730**	0.813**	0.574*	0.743**	0.756**
	(0.345)	(0.337)	(0.339)	(0.343)	(0.340)	(0.338)
Three-line frontier	−0.811**	−0.795**	−0.779*	−0.814**	−0.803**	−0.741*
	(0.402)	(0.400)	(0.398)	(0.403)	(0.402)	(0.399)
Average city research	2.811***	2.800***	2.719***	2.815***	2.810***	2.764***
	(0.300)	(0.300)	(0.299)	(0.299)	(0.299)	(0.299)
Average human capital	−0.0155*	−0.0147	−0.0243**	−0.0165*	−0.0156*	−0.0244**
	(0.00908)	(0.00898)	(0.0115)	(0.00927)	(0.00917)	(0.0117)
Labor pool	0.509***	0.638***	−0.233	0.503***	0.611***	−0.168
	(0.147)	(0.198)	(0.209)	(0.143)	(0.199)	(0.203)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.673***	2.638***	2.953***	2.739***	2.672***	2.986***
	(0.846)	(0.846)	(0.867)	(0.848)	(0.849)	(0.869)
Sigma	8.800***	8.798***	8.794***	8.800***	8.798***	8.794***
	(0.832)	(0.832)	(0.833)	(0.832)	(0.832)	(0.833)
Observations	11,038	11,038	11,038	11,016	11,016	11,016

Tobit regressions with robust standard errors
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

However, negative effects of midstream sectors are no longer significant when the 2005 and 2004 data are explored. This is also the case in the regressions on heterogeneous effects of agglomeration economies and R&D intensity across value chain (To save space, we did not show the results). This might not be contradictory to the major result. Strategies of preemption performed by midstream segments have side effects on their own profits so that they are not sustainable. As a result, we might not be able to detect long-term negative externalities of preemption. Furthermore, spillovers from downstream to midstream sectors are also not significant when agglomeration is calculated with the data in 2004 and 2005. This difference shows that technological upgrading in this way might be an instantaneous process and is also not sustainable. In general, the major results are still robust with multi-phase lag terms.

6.2 Possible endogeneity problem

Agrawal et al. (2010) provide another way to deal with potential endogeneity issues by introducing matching methods and conditional logit regressions based on cross-sectional data. The new method takes into account important distributional differences between R&D activities in and outside areas of agglomeration.

We use the propensity score matching method. Firms locating in cities with high levels of agglomeration economies[9] are classified as the high group (treatment group), while others are defined as the low group (control group). This method has the following steps. First, we choose observable characteristics on which two groups could be matched, which may include firm size, ownership and products. Second, we run logistic regressions on these multiple observable characteristics and obtain the predicted probability. Third, we match the two groups by choosing observations with the highest predicted probability in the treatment group for each firm in the control group. Thus, we can pick up the "best" possible control sample with closest characteristics to the treatment group. Finally, we run conditional logit regressions on the matched sample. The two firms in each pair differ in the levels of agglomeration economies but are similar in other characteristics. Thus, we can capture the net effect of agglomeration by running regressions on the matched sample.

Following Agrawal et al. (2010), we consider the following observable characteristics. (1) Control sample should be originated in the same city as the corresponding treatment sample; (2) control and treatment groups should come from the same 4-digit subsector; (3) of these, identify control firms having the same type of ownership as the treatment counterparts; and (4) ensure that matched samples have similar size in terms of the number of employees.

The results of matching methods are reported in Table 6[10], including mean values across treatment and control groups, difference of mean values, corresponding t statistics and *p* values. Note that the two groups do not have substantial overlap as the number of observations (which indicates the number of matched samples) is

[9] They are defined as cities whose density or location quotient or industrial diversity is higher than the median value.

[10] We allow multiple choices. That is why observations of treatment and control groups might be different.

Table 6 R&D intensity across different levels of agglomeration economies: matched sample

	Observations			R&D intensity			Difference	*t* statistics
	Control	Treatment	Overall	Control	Treatment	Overall		
Agglomeration economies								
Firm density at 4 digit level	564	561	1,125	0.0083	0.0084	0.0084	−0.0001	−0.0395
Location quotient at 4 digit level	1,500	1,486	2,986	0.0071	0.0159	0.0115	−0.0088	−4.6852
Industrial diversity	1,136	1,132	2,268	0.0074	0.0103	0.0009	−0.0029	−1.6317
Value chain								
Location quotient: upstream overall	1,172	1,144	2,316	0.0106	0.0090	0.0098	0.0016	0.8161
Location quotient: up to up	469	433	902	0.0048	0.0048	0.0048	0.0001	−0.054
Location quotient: up to mid	184	173	357	0.0176	0.0090	0.0134	0.0086	1.0921
Location quotient: up to down	519	538	1,057	0.0134	0.0124	0.0129	0.0010	0.3288
Location quotient: midstream overall	1,144	1,126	2,270	0.0087	0.0107	0.0097	−0.0019	−1.4308
Location quotient: mid to up	493	372	865	0.0052	0.0050	0.0051	0.0001	0.1219
Location quotient: mid to mid	176	175	351	0.0118	0.0084	0.0101	0.0034	1.0371
Location quotient: mid to down	475	579	1054	0.0113	0.0149	0.0133	−0.0036	−1.4744
Location quotient: downstream overall	1,127	1,108	2,235	0.0068	0.0115	0.0092	−0.0047	−4.1422
Location quotient: down to up	505	393	898	0.0030	0.0067	0.0046	−0.0036	−3.4842
Location quotient: down to mid	151	156	307	0.0103	0.0138	0.0121	−0.0035	−1.0725
Location quotient: down to down	471	559	1,030	0.0097	0.0143	0.0122	−0.0046	−2.2185

Data source: Annual Survey of Industrial Firms (ASIF)
"Difference" refers to the difference in R&D intensity between the control and treatment groups

not large. However, this does not generate significant bias with the propensity score matching method because we only focus on matched samples in the conditional logit regressions.

Comparing the treatment (high agglomeration) and control groups (low agglomeration), we find significant spillover effects from urbanization economies and agglomeration of downstream sectors, with p values on 'location quotient at 4-digit level', 'industrial diversity' and 'location quotient downstream' being 0.0000, 0.0541 and 0.0000, respectively. This is in accordance with the main result. There is no significant evidence of localization economies as the p value of 'firm density at 4-digit level' is 0.4842. But positive externalities from agglomeration still override free-riding problems here because R&D intensity is higher among treatment group. In addition, we find strong evidence of positive spillovers from the midstream to the downstream segments and from the downstream to the upstream sectors, which will, however, disappear in more rigorous analysis later on.

We further estimate the conditional logit models for the probability of R&D efforts to confirm the results of the matched pair analysis and to control for other confounders. The dependent variable is a dummy on whether there is R&D expense within a firm in year 2007. Independent variables include measurements of agglomeration economies and control variables on firm and city characteristics. Coefficients and particularly marginal effects of agglomeration economies are reported in Table 7. We find significant spillovers from agglomeration of downstream firms, with the coefficient of 0.433 and the marginal effect of 0.0152. The coefficient on externalities from midstream clusters is −0.473 with a marginal effect of −0.0189, which also confirms our previous results.

7 Summary

This study contributes to the literature by examining the relationship between agglomeration economies and firm R&D intensity using data of electronic and telecommunication industry in China.

Both localization and urbanization economies have significant impacts on firm R&D efforts. Such findings are contrary to what was revealed in a handful of other studies (Lee 2009; Suarez and Walrod 1997). In addition, downstream sectors are the major driving force of business R&D in ETEM industry. Spillovers exist along the value chain. Downstream firms are more likely to generate spillover effects than upstream ones, but the spillover is more significant within downstream segments. Upstream and midstream agglomeration may even have negative externalities due to preemption.

The different findings between this study and existing literature (Lee 2009; Suarez and Walrod 1997) call for more research along this line. In this study, we further explore the impacts of agglomeration economies from the concentration of firms that are in different segments of the value chain. Agglomeration economies are not symmetric, and such asymmetric relationships are related to the transitional nature of China's economy, namely its weak IPR protection regime and its position in the low-value-added part of the global value chain.

Table 7 Agglomeration economies across the value chain and R&D intensity in the electronics sector: conditional logit regressions

Variables	[1] Coefficient	[2] Marginal effect	[3] Coefficient	[4] Marginal effect	[5] Coefficient	[6] Marginal effect	[7] Coefficient	[8] Marginal effect
Agglomeration economies								
Industrial diversity	0.231	0.0326						
	(0.203)							
Location quotient upstream			−0.0100	−0.000505				
			(0.205)					
Location quotient midstream					−0.473*	−0.0189		
					(0.268)			
Location quotient downstream							0.433*	0.0152
							(0.243)	
Firm characteristics								
Duration	0.209**		0.0202		0.0970		0.0802	
	(0.102)		(0.0885)		(0.0905)		(0.0907)	
Export	−0.102		−0.234**		0.156		0.144	
	(0.115)		(0.110)		(0.124)		(0.112)	
HK and Taiwan	−1.066***		−0.664**		−0.500		−1.031**	
	(0.297)		(0.263)		(0.382)		(0.419)	
Foreign	−0.759***		−0.296		−0.772		−1.028*	
	(0.268)		(0.356)		(0.524)		(0.583)	
Middle size	−0.276		−0.236		−0.914		−1.971*	
	(0.494)		(0.448)		(0.898)		(1.099)	
Small size	−1.690***		−1.823***		−2.517***		−3.125***	
	(0.498)		(0.457)		(0.913)		(1.100)	

Table 7 continued

Variables	[1] Coefficient	[2] Marginal effect	[3] Coefficient	[4] Marginal effect	[5] Coefficient	[6] Marginal effect	[7] Coefficient	[8] Marginal effect
Nature of cities								
Municipality	−0.243		−0.232		0.628		−0.506	
	(0.393)		(0.302)		(0.383)		(0.319)	
Capital	−0.318		−0.110		0.409		0.548**	
	(0.268)		(0.289)		(0.291)		(0.255)	
Three-line frontier	−0.352		−0.295		−0.549		−0.920***	
	(0.368)		(0.334)		(0.363)		(0.338)	
Average city research	0.589***		0.683***		0.747***		0.738***	
	(0.169)		(0.133)		(0.160)		(0.147)	
Average human capital	0.0182		−0.00901		−0.0195		0.000260	
	(0.0115)		(0.0117)		(0.0144)		(0.00422)	
Labor pool	−0.143		−0.0132		0.472***		−0.107	
	(0.148)		(0.123)		(0.164)		(0.124)	
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	856	856	916	916	870	870	890	890

We calculate the coefficients of all independent variables and the marginal effects of agglomeration economies (different measurements of location quotient and industrial diversity) from conditional logit models. We only report the results of industrial diversity and measurements of location quotient of upstream, midstream and downstream firms to save place. All the variables have been standardized. *Data source*: Annual Survey of Industrial Firms (ASIF)

This study has theoretical implications. We suggest that leading theories about localization and urbanization economies require further revision when applied to transitional economies. In this paper, we deal with agglomeration economies under imperfect competition and our focus on firm characteristics can inspire further investigation of firm heterogeneity in the field of economic geography.

Acknowledgments The authors would like to acknowledge funding from the Natural Science Foundation of China (Nos. 41271130 and 41071075) and the constructive comments and suggestions of the anonymous referees and the editor. The errors remain those of the authors.

Appendix

See Table 8.

Table 8 Summary of dependent and independent variables

Variables	Features	Expected sign	Mean	SD
Dependent (year 2007)				
R&D intensity	Continuous; many 0 values		0.0096	0.0417573
Independent: agglomeration economies (year 2006)				
Firm density at 4 digit level	Continuous	Uncertain	0.0021	0.0029
Location quotient at 4 digit level	Continuous	+	2.0729	1.9009
Location quotient at 2 digit level	Continuous	+	2.0741	1.8830
Industrial diversity	Continuous	+	0.3426	0.1649
Location quotient upstream	Continuous	Depends on different types	2.0745	1.8461
Location quotient midstream	Continuous	Depends on different types	1.8071	1.7059
Location quotient downstream	Continuous	Depends on different types	2.1185	2.2023
Average city research	Continuous	+	0.0095	0.0120
Average human capital	Continuous	+	28.6844	14.8505
Labor pool	Continuous	+	321,069.1	403,077.9
Independent: firm characteristics (year 2007)				
Duration	Continuous	+	7.5648	6.8310
Export	Continuous	−	0.3249	0.4198
HK and Taiwan		−	0.2153	0.4110
Foreign		−	0.2617	0.4396
Middle	Dummy	−	0.2198	0.4141
Small	Dummy	−	0.7454	0.4357

Table 8 continued

Variables	Features	Expected sign	Mean	SD
Independent: nature of cities (year 2007)				
Municipality	Dummy	+	0.1474	0.3545
Capital	Dummy	+	0.1416	0.3486
Three-line frontier	Dummy	+	0.0867	0.2815

Data source: Annual Survey of Industrial Firms (ASIF)

References

Agrawal A, Cockburn I, Rosell C (2010) Not invented here? Innovation in company towns. J Urban Econ 67:78–89

Autant-Bernard C, LeSage JP (2011) Quantifying knowledge spillovers using spatial econometric models. J Reg Sci 51:471–496

Bartel A, Lichtenberg F (1987) The comparative advantage of educated worker in implementing new technologies. Rev Econ Stat 69:1–11

Battisti G, Stoneman P (2003) Interfirm and intrafirm effects in the diffusion of new process technologies. Res Policy 32:1641–1655

Beal BD, Gimeno J (2001) Geographic agglomeration, knowledge spillovers, and competitive evolution. INSEAD working papers (2001–26)

DiMaggio Paul J, Powell W (1983) The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. Am Sociol Rev 48:903–941

Fan P (2006) Catching up through developing innovation capability: evidence from China's telecom-equipment industry. Technovation 26:359–368

Fan P (2011) Innovation, globalization, and catch-up of latecomers: cases of Chinese telecom firms. Environ Plan A 43:830–849

Feldman M, Audretsch DB (1999) Innovation in cities: science-based diversity, specialization and localized competition. Eur Econ Rev 43:409–429

Feldman MP (1999) The new economics of innovation, spillovers and agglomeration: a review of empirical studies. Econ Innov New Technol 8:5–25

Gao J, Fu J (1996) The key problems of technological innovation in business firms. Sci Technol Int 1:24–33

Glaeser EL, Kallal HD, Scheinkman JA, Shleifer A (1992) Growth in cities. J Polit Econ 100:1126–1152

Gordon IR, McCann P (2005) Innovation, agglomeration, and regional development. J Econ Geogr 5:523–543

Gu S, Lundvall B (2006) China's innovation system and the move toward harmonious growth and endogenous innovation. Innov Manag Policy Pract 8:1–26

He C, Pan F (2010) Economic transition, dynamic externalities and city-industry growth in China. Urban Stud 47:121–144

He C, Zhu S (2007) Economic transition and regional industrial restructuring in China: structural convergence or divergence. Post Communist Econ 19:321–346

He C, Zhu S (2008) Industrial agglomeration of manufacturing industries in China: a comparison of Jiangsu and Anhui Province. Sci Geogr Sin 28:715–721

Henderson V (1997) Externalities and industrial development. J Urban Econ 42:449–470

Henderson V (2003) Marshall's scale economies. J Urban Econ 53:1–28

Koo J (2005) Agglomeration and spillovers in a simultaneous framework. Annu Reg Sci 39:35–47

Krugman PR, Venables AJ (1995) Globalization and the inequality of nations. Q J Econ 110:857–880

Lee C (2009) Do firms in clusters invest in R&D more intensively? Theory and evidence from multi-country data. Res Policy 38:1159–1171

Liu X, White S (2001) Comparing innovation systems: a framework and application to China's transitional context. Res Policy 30:1091–1114

Meng QX, Li MZ (2001) New economy and ICT development in China. WIDER Discussion Papers 2001/76, World Institute for Development Economics (UNUWIDER), New York
Mudambi R (2008) Location, control and innovation in knowledge-intensive industries. J Econ Geogr 8:699–725
Ponds R, van Oort F, Frenken K (2010) Innovation, spillovers and university-industry collaboration: an extended knowledge production function approach. J Econ Geogr 10:231–255
Pouder R, John C (1996) Hot spots and blind spots: geographical clusters of firms and innovation. Acad Manag Rev 21:1192–1225
Rosenthal SS, Strange WC (2004) Evidence on the nature and sources of agglomeration economies. Handbook Reg Urban Econ 4:2119–2171
Stoneman P, Battisti G (2010) The diffusion of new technology. Handbook Econ Innov 2:733–760
Suarez V, Walrod W (1997) Operational strategy, R&D and intra-metropolitan clustering in a polycentric structure: the advanced electronics industries of the Los Angeles Basin. Urban Stud 34:1343–1380
Sun Y (2010) What matters for industrial innovation in China: R&D, technology transfer or spillover impacts from foreign investment? Int J Bus Syst Res 4:621–647
Tellis GJ, Peter NG (1996) First to market, first to fail? Real causes of enduring market leadership. MIT Sloan Manag Rev 37:65–75
Venables AJ (1996) Equilibrium locations of vertically linked industries. Int Econ Rev 37:341–359
White S (2000) Competition, capabilities and the make, buy or ally decisions of Chinese state-owned firms. Acad Manag J 43:324–341
Yang FR (2006) The research of development of information industry in west China. Ph.D. thesis, Department of Economics, Sichuan University (in Chinese)
Zhang J, Liu Z, Zheng J (2007) Industrial chain positioning, divide, agglomeration and innovation: an empirical study based on questionnaire of manufacturing firms in Jiangsu Province. China Ind Econ 7:47–55 (in Chinese)
Zhang X, Li L (2007) Manufactory industry agglomeration, knowledge spillover and regional innovation performance. Quant Econ Res 8:35–43 (in Chinese)
Zhou Y, Sun YF, Wei YH, Lin CS (2011) De-centering 'spatial fix'-patterns of territorialization and regional technological dynamism of ICT hubs in China. J Econ Geogr 11:119–150