Papers in Evolutionary Economic Geography

15.24

How to jump further? Path dependent and path breaking in an uneven industry space

Shengjun Zhu, Canfei He & Yi Zhou

Utrecht University
Urban & Regional research centre Utrecht

How to jump further? Path dependent and path breaking in an uneven industry space

Shengjun Zhu[a], **Canfei He**[b,*], and **Yi Zhou**[c]

[a] Department of Geography, College of Science, Swansea University, Singleton Park, Swansea, SA2 8PP, U.K.
Email: s.zhu@swansea.ac.uk

[b] College of Urban and Environmental Sciences, Peking University, Beijing, 100871, China
Email: hecanfei@urban.pku.edu.cn
[*] Corresponding author. Tel.: +86 10 6275 7075

[c] College of Urban and Environmental Sciences, Peking University, Beijing, 100871, China

Abstract: By using the proximity product index, recent studies have argued that regional diversification emerged as a path-dependent process, as regions often branch into industries that are related to preexisting industrial structure. It is also claimed that developed countries that start from the core, dense areas in the uneven industry space have more opportunities to jump to new related industries and therefore have more opportunities to sustain economic growth than do developing countries that jump from peripheral, deserted areas. In this paper, we differentiate two types of regional diversification—path-dependent and path-breaking—and ask questions from a different angle: can developing countries/regions jump further in the industry space to break path-dependent development trajectories and more importantly to catch up with developed ones? Based on China's export data, this paper shows that regions can jump further by investing in extra-regional linkages and internal innovation. Not only do these two sets of factors promote regions' jumping capability, but they also contribute to regions' capability of maintaining a comparative advantage in technologically distant and less related industries. In addition, different extra-regional linkage and internal innovation factors have affected regional diversification to different extents, and these effects also vary across regions and industries. Empirically, this research seeks to find a more promising future for developing countries/regions. Theoretically, our research testifies some key findings of theoretical works in evolutionary economic geography by using a quantitative framework. In addition, this paper includes some economic and institutional factors that have been left out in previous studies.

Keywords: path-dependent, path-breaking, industrial relatedness, proximity index, transition industries

Acknowledgements
This work was supported by the Natural Science Foundation of China under Grant No. 41425001 and No. 41271130.

1. Introduction

It is argued that technological relatedness not only pushes forward the growth of existing industries through agglomeration externalities derived from related variety, but is also responsible for the formation of new growth paths (Neffke, Henning, and Boschma 2011). On the one hand, Frenken, Van Oort and Verburg (2007) and Boschma and Iammarino (2009) have divided Jacobs's externalities into related and unrelated variety, and argued that Jacobs's externalities does not necessarily result in knowledge spillover, instead, it only takes place effectively when complementarities and technological relatedness exist among industrial sectors in terms of shared competences. In the literature on relatedness in the context of spatial externalities and regional growth, it is increasingly acknowledged that related variety has the potential to provide more learning opportunities for local industries, generate more inter-sectoral knowledge spillovers, and finally facilitate regional economic development (Boschma and Iammarino 2009; Boschma, Minondo, and Navarro 2012, 2013; Frenken, Van Oort, and Verburg 2007; Neffke, Henning, and Boschma 2011). On the other hand, in addition to driving growth of existing industries in the short term, relatedness also plays a key role in new path creation in the long term. Recent studies show that new growth paths do not emerge from scratch, but evolve out of preexisting regional industrial structures, because the set of competences and assets that region possesses determines what new paths and new industries this region is able to develop (Boschma, Minondo, and Navarro 2012; Hidalgo et al. 2007; Neffke, Henning, and Boschma 2011). Here, the set of competences and assets consists of a broad range of productive inputs, including both tangible assets such as physical and human capital and intangible assets like institutions and norms. It reflects a region's capabilities to develop new paths and new industries. If a region already has most of the capabilities that a certain new industry requires, it will be easy for this region to jump onto this new path. In contrast, if not, the barrier to develop this industry could be too high for this region to overcome (Boschma, Minondo, and Navarro 2013). In a word, regions tend to diversify into new industries that are related with preexisting regional industrial structure, and relatedness among industries affect the ways in which regions create new paths over time.

By emphasizing industrial relatedness, Hidalgo and Hausmann (2008) have directed our attention away from a one-dimensional, GDP-centered view of development where a region's development is measured by its upward movement on a GDP ladder (or ramp) regardless of the products and services the region produces, towards a two-dimensional network view of development. In the "network" metaphor, regions are thought as monkeys jumping among trees (i.e., industries) in a big, heterogeneous forest (i.e., a two-dimensional, uneven industry space/network where different industries are related with each other to different extents). Monkeys can only jump onto trees within certain distance (i.e., relatedness between two industries). In other words, regions are more likely to "jump" (or diversify) into industries that are closely related to their existing industries. Development of regions is thus shaped by technological relatedness among industries. In addition, the monkey's position in the forest also affects its jumping trajectories. It would be easier for monkeys in the dense areas to jump to neighboring trees than their counterparts in more deserted areas in the heterogeneous forest. Likewise, Hidalgo et al. (2007) have argued that developed countries that start from the core, dense areas in the uneven industry space/network have more opportunities to jump to new related industries and therefore have more opportunities to sustain economic growth than do developing countries that jump from peripheral, deserted areas. Boschma et al. (2013) and Neffke et al. (2011) further pointed out that this process of related diversification and path-dependent economic development is more phenomenal at the regional level.

Even though recent studies do not exclude the possibility that developing countries/regions can enter the core, dense areas in the uneven industry space/network, the future for them is not quite promising. In some extreme cases, developing countries/regions may only be able to wander around in the peripheral areas of the industry space, and have no capabilities to jump into the core areas (see (Hidalgo et al. 2007) for an example where the divergence between developing countries and developed countries persists forever and it is impossible for developing countries such as Chile to catch up with developed ones no matter how much time the former have). Overall, the development prospect for developing countries/regions is dimmed, if not hopeless, due not only to their peripheral starting point, but also to the fact that rich countries are capable to jump further than poor ones (see (Boschma and Capone 2014b) for an example where richer countries in Europe are more capable to diversify into less related industries). Based on recent studies, this research seeks to shed new light on the evolution of regional industrial structure and the process of regional diversification characterized by path dependence, by asking questions from a different angle: can developing countries/regions transcend the "confinement" of technological relatedness to catch up with developed ones? If yes, the developing countries/regions have to jump further in the industry space. A follow-up question is: how to jump further in the industry space/network? In this paper, we use a developing country that has huge regional disparity—China—as an example to explore whether developing regions/countries can jump further and catch up, and through what ways. The next section will present an analytical framework, and develop hypotheses. The third section introduces the data, variables and specifications for empirical analysis. After presenting some descriptive analyses in the fourth section, section five discusses the empirical results. The last section concludes the paper by summarizing main findings, and points out the theoretical and empirical significances of our central question.

2. Path-dependent and path-breaking

Before examining regional economic development and industrial restructuring through the lens of technological relatedness, it is useful to make some clarifications of the notion path dependence, which is arguably the core concept of the regional diversification or "regional branching" model (Boschma and Capone 2014b; Boschma, Minondo, and Navarro 2012, 2013; Neffke, Henning, and Boschma 2011). In this paper, we differentiate two types of new growth path creation. The first one takes place when regions rely on the technological relatedness among industries (distance between trees) in order to jump into new industries. This type of new path creation is path-dependent, since regional diversification is determined not only by region's position in the industry space and the density in the vicinity, but also by the relatedness between industries. As regions jump according to technological relatedness, technologies evolve over time through cycles of long periods of incremental innovations, which enhance and institutionalize an existing productive structure. This path-dependent process means that there is some degree of cohesion in the industrial structure of a region (Neffke, Henning, and Boschma 2011). When this type of new path creation dominates, it is extremely difficult for developing countries/regions to catch up, since each region follows its own industrial trajectory (Rigby and Essletzbichler 1997).

However, regional industrial trajectory is constantly redirected or redefined through the process of creative destruction (Schumpeter 1947). New paths can be created by technological discontinuities in which regions replace old, inferior technologies/industries with new, radically superior technologies/industries (Baum 1996). In this case, regions, particularly those in the peripheral areas of the industry space, rely less on technological relatedness, and jump directly to less related or even unrelated industries. While the first type is path dependent because new industries are created based on tangible and intangible assets embodied in existing industrial structure, the second type of path creation is more path-breaking driven by technological discontinuities and radical innovations that allow regions to jump further. The path-dependent regional diversification model has already been perfectly testified by recent empirical, quantitative studies based on export and plant-level data in developed countries (Boschma, Minondo, and Navarro 2012, 2013; Neffke, Henning, and Boschma 2011). In this research, based on export data in a developing country, we seek to focus on the second type of path creation and examine how developing countries/regions can jump further in a more path-breaking way, so that the "confinement" of technological relatedness can be transcended. Two sets of factors that have the potential to enable developing countries/regions to jump further are identified based on evolutionary economic geography (EEG) theory: extra-regional linkages and internal innovation.

2.1. Two trajectories of path dependence, openness of region and extra-regional linkages

The standard canonical path dependence model portrays the regional industrial evolution as a four-phase development: (1) *path creation*, where historical accidents initiate a new path and have significant long-run effects on the technological, industrial and institutional structure of a region; (2) *path development*, where emergence and development of local increasing returns and externalities assists the development of the path; (3) *path rigidification*, characterized by increasing rigidification of knowledge, networks and structures of firms; (4) *path de-locking*, where an exogenous shock disrupts or dislodges the regional economy resulting in an atrophy or industrial restructuring (Arthur 1989, 1994; David 1985; Martin 2010; Martin and Sunley 2006).

After criticizing the standard canonical path dependence model's overemphasis on continuity and stability, Martin (2010) has suggested a second type of trajectory which is more open and allows for constant endogenous change. His model diverges from the canonical one in the third step, and proposes a new phase three where local industry is able to adapt and mutate constantly, which prevents it from being trapped in a stable, inflexible and rigid state that only can be destabilized by an external shock. This idea has been further developed by Martin and Sunley (2011) when they employ a modified regional adaptive cycle model and introduce much more diverse trajectories of regional economic evolution. The gist of their argument for our present purposes is that apart from being stuck in a state of fixity and rigidification and waiting for an unpredictable external shock to set it free, a regional economy can evolve along another trajectory where firms in the region are able to innovate more or less continuously and the industrial structure constantly mutates and adapts.

This second trajectory can be realized by keeping the regional economy relatively open (Hassink 2005). The openness of a region is partly supported by the diverse overlaps between organizations and institutions inside and outside the region, and the subsequent information, technology and knowledge exchange across regional borders (Sydow, Lerch, and Staber 2010). The kind of understanding and learning that derives from participation in various kinds of links with others beyond the region is referred to as 'pipelines' or extra-regional linkages (Bathelt, Malmberg, and Maskell 2004; Boschma and Iammarino 2009). Extra-regional linkages which connect actors inside and outside the region may be important in enabling firms in the region to avert tendencies towards path dependence in the evolution of regional economy, thus enabling the region to remain innovative and competitive (Bathelt and Li 2013; Bathelt, Malmberg, and Maskell 2004; Sturgeon, Van Biesebroeck, and Gereffi 2008; Sydow, Lerch, and Staber 2010). The idea that nurturing connections with distant actors may help prevent systematic industrial rigidification is also supported by Maskell and Malmberg (2007) who have argued that, from a micro-level perspective, localized learning and knowledge development often lead to overreliance on localized routines and over-embeddedness in existing structure, what Maskell and Malmberg called 'spatial myopia'. The potentially devastating long-run effects of spatial myopia may be avoided as long as some firms in the region actively invest in establishing linkages to extra-regional knowledge pools with dissimilar routines or institutional patterns. Rigidification and path dependence is therefore alleviated by rejuvenation processes where externally connected local firms are able to keep importing fresh knowledge and state-of-the-art technology.

By deliberately investing in building extra-regional linkages to distant communities, regions may be able to increase the variety of knowledge, resources, and capabilities available to them and escape the potential limits stemming from their peripheral starting point in the industry space and the confinement of relatedness. Since extra-regional linkages bring in fresh know-how and technologies that are less related to region's existing productive structure, it may enable regions to jump further.

2.2. Purposeful actions of actors and internal innovation

Not only has the canonical path dependence model overlooked Martin's second trajectory of cluster evolution, but it also rarely pays attention to the role of individual agency in affecting path-dependent processes (Arthur 1989, 1994; David 1985; Henning, Stam, and Wenting 2013; Martin 2010). Once a regional economy enters into the third phase in the traditional path dependence model, it is assumed that it will remain or be trapped in path-dependent development until it is disturbed or liberated by some unpredictable and unexpected exogenous shock (Martin 2010; Martin and Sunley 2011). This focus on exogenous impacts has meant that the role of individual agency in re-working or disrupting forms of path dependency and creating new pathways is much less well developed over against the much stronger analytical focus on the role of external shocks in dislodging stable, inflexible or rigid state of regions. To counter this lacuna, Martin and Sunley (2011) have suggested that, while the path dependence of firms, networks, and structures in a region may limit the vitality and adaptability of the region resulting in an atrophy, it may also encourage or enable a reorganization of resources and greater opportunities for surviving firms or it may force purposeful action by individual actors in a region (firms, industry associations and governments) who are deliberately trying to de-lock themselves. In this way, the entire region can dislodge itself from an old path and create a new one (Sydow, Lerch, and Staber 2010). As a result, there are reasons to consider path-dependent process of regional economic evolution as possibly being shaped by purposeful actions of actors.

In this paper, we build on these insights to evaluate the relative roles of purposeful actions of actors within the region in shaping path-dependent processes as they inflect regional economic evolution through internal innovation. Here by purposeful actions of actors, we are referring not only to the corporate strategies and innovation implemented by enterprises, but also to various supporting facilities and services, financial, political and technological aids provided by regional governments. In the Chinese case, in particular, it is vital that the evolution of regional economy should be examined in ways that recognize the strategic intents of both regional governments and enterprises to innovate and adopt new technologies/products. Such internal innovation may also generate radical technological shifts, and enable regions to break old pathways and jump to less related industries.

3. Data and research design

3.1. Proximity between industries and industry space in China

To investigate whether regions are able to break old pathways and jump to less related industries, we need a relatedness indicator to measure the technological relatedness between new and existing industries. This indicator also reflects the distance between new industries and existing industrial structure, thus allowing us to observe whether regional industrial structure changes towards related or relatively unrelated industries. Various indicators of relatedness have been developed and used in recent studies (Boschma, Minondo, and Navarro 2012; Bryce and Winter 2009; Frenken, Van Oort, and Verburg 2007; Hidalgo et al. 2007; Neffke, Henning, and Boschma 2011; Porter 2003; Teece et al. 1994). Boschma et al. (2012) have pointed out that the *ex-post* relatedness indicator developed by Hidalgo et al. (2007) based on proximity product index can better capture the essence of technological relatedness than can the conventional *ex ante* measure of related and unrelated variety (Boschma and Iammarino 2009; Frenken, Van Oort, and Verburg 2007) and the cluster-based *ex-post* indicator of industrial relatedness formulated by Porter (2003). This product-proximity-based approach is similar to the co-occurrence-based analysis method developed by Neffke et al. (2011) to measure the industrial relatedness by assessing whether two industries are often found together in one and the same economic entity.

Since a variety of factors may affect the relatedness between two industries, including similarities in the combination of productive factors, the use of technologies, the characteristics of customers, required institutions and social norms, Hidalgo et al. (2007) have adopted an *ex-post* approach to measure proximity between two industries. Two industries are considered to be related with each other if regions tend to have revealed comparative advantage (RCA) in both. A region has a comparative advantage in an industry when the share of this industry in the region's total exports is larger than the share of this industry in the national total[1]. The proximity (ϕ) between industry i and industry j at year t can be calculated as:

[1] In this research, we follow Neffke et al. (2011) and calculate proximity indicators based on data in one country. First, China is a geographically large country with huge regional disparity, which means

$$\phi_{i,j} = min\{P(RCA_{c,i} > 1 | RCA_{c,j} > 1), P(RCA_{c,j} > 1 | RCA_{c,i} > 1)\} \quad (1)$$

Where,

$$RCA_{c,i} = \frac{Export\ Value_{c,i} / \sum_i Export\ Value_{c,i}}{(\sum_c Export\ Value_{c,i} / \sum_{c,i} Export\ Value_{c,i})} \quad (2)$$

$RCA_{c,i}$ is the revealed comparative advantage of industry *i* in city *c*. City *c* is considered as having a comparative advantage in industry *i*, if $RCA_{c,i}$ is above 1. The proximity between industry *i* and industry *j* is the minimum between the conditional probability of having a comparative advantage in industry *i*, given that the city *c* has a comparative advantage in industry *j* (i.e., $P(RCA_{c,i} > 1 | RCA_{c,j} > 1)$), and the conditional probability of having a comparative advantage in in industry *j*, given a revealed comparative advantage in industry *i* (i.e, $P(RCA_{c,j} > 1 | RCA_{c,i} > 1)$). The rationale behind this proximity indicator is that if two industries are related with each other, they probably demand similar institutions, infrastructure, factor inputs, capabilities and technology, and are likely to be produced together.

The proximity indicator between industries is computed using Chinese customs data during 2001-2013. This dataset reports imports and exports by 6-digit product, providing information on export/import value and quantity, location and ownership of exporting/importing firms, destination of exports, origin of imports, and transportation mode. In this research, we focus on 4-digit level industries (1,080 industries in the dataset[2]). The geographical unit of analysis is China's prefecture-level cities. A matrix of proximity indicators among all 4-digit industries can be estimated. This 1,080*1,080 matrix therefore defines the industry space. In this paper, we adopt a more dynamic understanding of technological relatedness between industries, in which industry space is conceptualized as a network structure that changes over time (Hidalgo and Hausmann 2008; Hidalgo et al. 2007; Neffke, Henning, and Boschma 2011). As a result, proximity indicator is calculated for each year of the 2001-2013 time period. Table 1 shows the Pearson's correlation coefficient between the industry spaces generated with data from different years. Industry space appears to be stable, but links do change over time. Correlation between two adjacent years is around 0.7-0.8, higher than that between two distant years. The further two years are from each other, the lower the correlation between these two years is. The correlation between 2013 and 2001 is the smallest (0.42), indicating China's industrial structure has undergone a dramatic transformation in this time period. Nonetheless, China's industrial structure has also been stabilizing, as the correlation between proximity indicators in two adjacent years has increased from 0.71 (between 2001 and 2002) to 0.84 (between 2012 and 2013).

Table 1 Correlation between proximity indicators between industries in different years

	2013	2012	2011	2010	2009	2008	2007	2006	2005	2004	2003	2002	2001

calculation based on data in such a big economy should be sufficient. Second, calculating proximity indicators based on Chinese export data rather than world trade data may allow us to better control China's unique, nation-wide political and economic environments.

[2] We focus on the secondary industry and therefore exclude data on agriculture.

2013	1												
2012	0.84	1											
2011	0.75	0.82	1										
2010	0.72	0.78	0.84	1									
2009	0.71	0.75	0.79	0.83	1								
2008	0.62	0.65	0.70	0.73	0.78	1							
2007	0.58	0.60	0.65	0.67	0.71	0.78	1						
2006	0.55	0.58	0.62	0.64	0.67	0.71	0.77	1					
2005	0.53	0.55	0.59	0.61	0.64	0.67	0.71	0.78	1				
2004	0.50	0.52	0.55	0.56	0.60	0.63	0.66	0.71	0.77	1			
2003	0.47	0.49	0.52	0.53	0.57	0.59	0.62	0.66	0.70	0.75	1		
2002	0.45	0.47	0.50	0.51	0.54	0.56	0.59	0.62	0.65	0.68	0.75	1	
2001	0.42	0.44	0.47	0.48	0.50	0.53	0.55	0.57	0.60	0.62	0.66	0.71	1

3.2. Model Specification

To measure the distance between new industries and region's existing productive structure, we calculate the density indicator, developed by Hidalgo et al. (2007). It has been argued that the regional diversification is path-dependent, because if a new industry is related to a number of industries that the region already has a comparative advantage in, the density of this new industry for this region is high, and the probability for this region to jump into this new industry will also be high. Density indicator is therefore measured as follows:

$$d_{i,c,t} = \frac{\sum_j x_{j,c,t} \phi_{i,j,t}}{\sum_j \phi_{i,j,t}} \quad (3)$$

where $x_{j,c,t}$ takes the value of 1 if city *c* has a comparative advantage in industry *j* at year *t*, and zero otherwise. Density around a new industry will be high if a region has a comparative advantage in most of the industries related to the focal one.

The following econometric equation is estimated:

$$x_{i,c,t2} = \alpha_0 + \alpha_1 x_{i,c,t1} + \alpha_2 d_{i,c,t1} + \beta_1 x_{i,c,t1} EXT_{i,c,t1} + \beta_2 (1 - x_{i,c,t1}) EXT_{i,c,t1} + \beta_3 x_{i,c,t1} INT_{i,c,t1} + \beta_4 (1 - x_{i,c,t1}) INT_{i,c,t1} + \gamma_1 x_{i,c,t1} d_{i,c,t1} EXT_{i,c,t1} + \gamma_2 (1 - x_{i,c,t1}) d_{i,c,t1} EXT_{i,c,t1} + \gamma_3 x_{i,c,t1} d_{i,c,t1} INT_{i,c,t1} + \gamma_4 (1 - x_{i,c,t1}) d_{i,c,t1} INT_{i,c,t1} + \delta X + \varepsilon_{i,c,t1} \quad (4)$$

where *t2>t1*, and *EXT* and *INT* represent extra-regional linkage and internal innovation variables respectively. *X* is a vector of city-year and industry-year dummy variables, which is added to control any time-varying city or industry characteristics. Dependent variable takes the value of 1 if city *c* has a comparative advantage in industry *i* at year *t2* and zero otherwise. Following Hausmann and Klinger (2007) and Boschma et al. (2013), we distinguish the role that independent variables play in developing a comparative advantage in new industries at the city level from their contribution in keeping a comparative advantage in current industries. β_1, β_3, γ_1 and γ_3 capture the impact of specific independent variables and the interaction terms between independent variables in keeping a comparative advantage in industry *i*, while β_2, β_4, γ_2 and γ_4 capture the impact of specific independent variables and the interaction terms between independent variables in developing a comparative advantage in new industries.

High level of $d_{i,c,t}$ indicates that the distance between industry *i* and city *c*'s existing industrial structure is small, the positive effect of density thus means that regions jump into related industries and regional economic development is path-dependent. Our hypothesis is that two sets of variables (*EXT* and *INT*) may change regions' jumping capability, reducing (or strengthening) regions' reliance on relatedness among industries while jumping. Therefore, the impact of density should vary across regions as the latter are different from one another in terms of extra-regional linkages and internal innovation. To test our hypothesis, we follow Boschma and Capone (2014a) and include the interaction terms between density indicator and *EXT*/*INT* variables in Equation (4). A positive and significant sign of the interaction term indicates specific *EXT*/*INT* variable enhance regions' reliance on density while jumping, whereas a negative and significant sign means a weaker effect of density (Boschma and Capone 2014a, 2014b). In the latter case, regions' certain characteristics in terms of extra-regional linkage and internal innovation reduce the confinement of relatedness, enable regions to jump further to less related industries, and finally allow regions to become path-breaking. A non-significant sign is also possible, suggesting the impact of density does not vary across regions because regions have different levels of extra-regional linkages or internal innovation.

3.3. Variables

First, we seek to test the effects of extra-regional linkages that breathe new life into a region through a diversified set of import sectors and through foreign direct investment (FDI). On the one hand, imports expand the set of inputs available in the economy and thus increase regions' productivity (Amighini and Sanfilippo 2014). The rising availability of inputs may encourage the creation of new domestic varieties (Goldberg et al. 2010). Imports can also provide more sophisticated inputs that enable regions to upgrade their production and export. On the other hand, more importantly, there is a certain degree of new knowledge embedded in imported products, which could translate into new learning opportunities involved in the use of new products (Dollar 1992; Schiff and Wang 2006). As a result, $IMPORT_{i,c,t}$ is included, measured as the import value of city *c* in industry *i* and year *t*.

It is acknowledged that FDI plays a critical role in promoting regional economic development through a variety of channels, such as the formation of forward and backward linkages, the existence of competitive and demonstration effects, the possibility for domestic firms to recruit more experienced and skilled workers that are released from foreign-owned firms, and finally through the knowledge spillover effects between domestic and foreign-owned firms (Görg and Greenaway 2004; Lall and Narula 2004; Poncet and Starosta de Waldemar 2013; Zhu and Fu 2013). Foreign-owned firms are important for regional economy, as they not only contribute to productivity increase in existing industries, but, more importantly for our present purpose, they also bring new knowledge and ideas, that may enable regions to break their old paths and jump into less related industries (Amighini and Sanfilippo 2014). Variable $FDI_{i,c,t}$ is calculated as the share of exports of foreign-owned firms in industry *i*, city *c* and year *t*.

Not only can new knowledge and ideas that enable regions to jump further be acquired through access, transfer and assimilation of external knowledge, but regions can also create new knowledge through indigenous innovation. R&D, human and physical capital, and government supports are all important factors that contribute to indigenous knowledge creation (Fu and Gong 2011; Romer 1990). At the regional level, these factors determine the region's capability. Regions with different types of R&D, human and physical capital, and government policies are likely to have different jumping capabilities and this will affect the process of regional diversification as well as region's reliance on relatedness among industries (Boschma and Capone 2014b). Human capital ($HCAP_{c,t}$) is measured as the number of college students per ten thousand people in city *c* and year *t*. We use the length of highway over land area in city *c* and year *t* as a proxy of physical capital and infrastructure ($PCAP_{c,t}$). Data on these two indicators are derived from China City Statistical Yearbook. $R\&D_{i,c,t}$, is calculated as the R&D investment by enterprises in industry *i*, city *c* and year *t*. For this indicator, we use data from China's Annual Survey and Industrial Firms (ASIF). The problem is that this dataset uses SIC industrial classification system while Chinese customs data is compiled in HS industrial classification system. We concord R&D indicators for all 4-digit SIC industries (525 industries) at the city level to 4-digit HS industries (1,080 industries)[3]. Finally, government supports in different industries is computed as the difference between export tax rebates and tax rates in industry *i* and year *t* ($REBATE_{i,t}$). Data for this indicator is taken from the Chinese customs data on export tax rebates and tax rates.

4. The relationship between density and the emergence of new industries in Chinese cities

We divide the time period 2001-2013 into two stages: 2001-2006 and 2007-2013 for two reasons. First, the HS industrial classification system used by Chinese customs dataset was slightly adjusted in 2007. Second, these two stages are both around five years, which are long enough to allow new industries to emerge (Boschma, Minondo, and Navarro 2013). In addition, it takes at least around five years for density to have a strong impact (Boschma and Capone 2014b). Such a division also allows us to observe the different impacts of independent variables before and after 2008 global financial crisis.

We first analyze the relationship between the average density of the industries without a comparative advantage in a city in 2001 (or 2007) and the probability of this city developing a comparative advantage in a new industry in 2006 (or 2013). As notified above, city *c* is considered as having a comparative advantage in industry *i* if $RCA_{i,c}$ is above 1. Industry *i* is considered as a developed industry in city *c*. Otherwise, it is an undeveloped industry. If industry *i* is an undeveloped industry in 2001 (or 2007) and becomes a developed industry in 2006 (2013), it is named as a transition industry (Hidalgo et al. 2007). We also differentiate cities in East, Central, Southwest, Northeast and Northwest China (Figure 1).

[3] As there are more 4-digit HS industries than 4-digit SIC industries, a SIC industry is often bigger than a HS industry and the latter is often a subset of the former. In some cases, we have to use the R&D value in a big SIC industry as a proxy of that in a small HS industry. We admit this concordance is not precise. However, due to unavailability of R&D data compiled in HS system we have to use data compiled in SIC system and the results do provide some interesting insights.

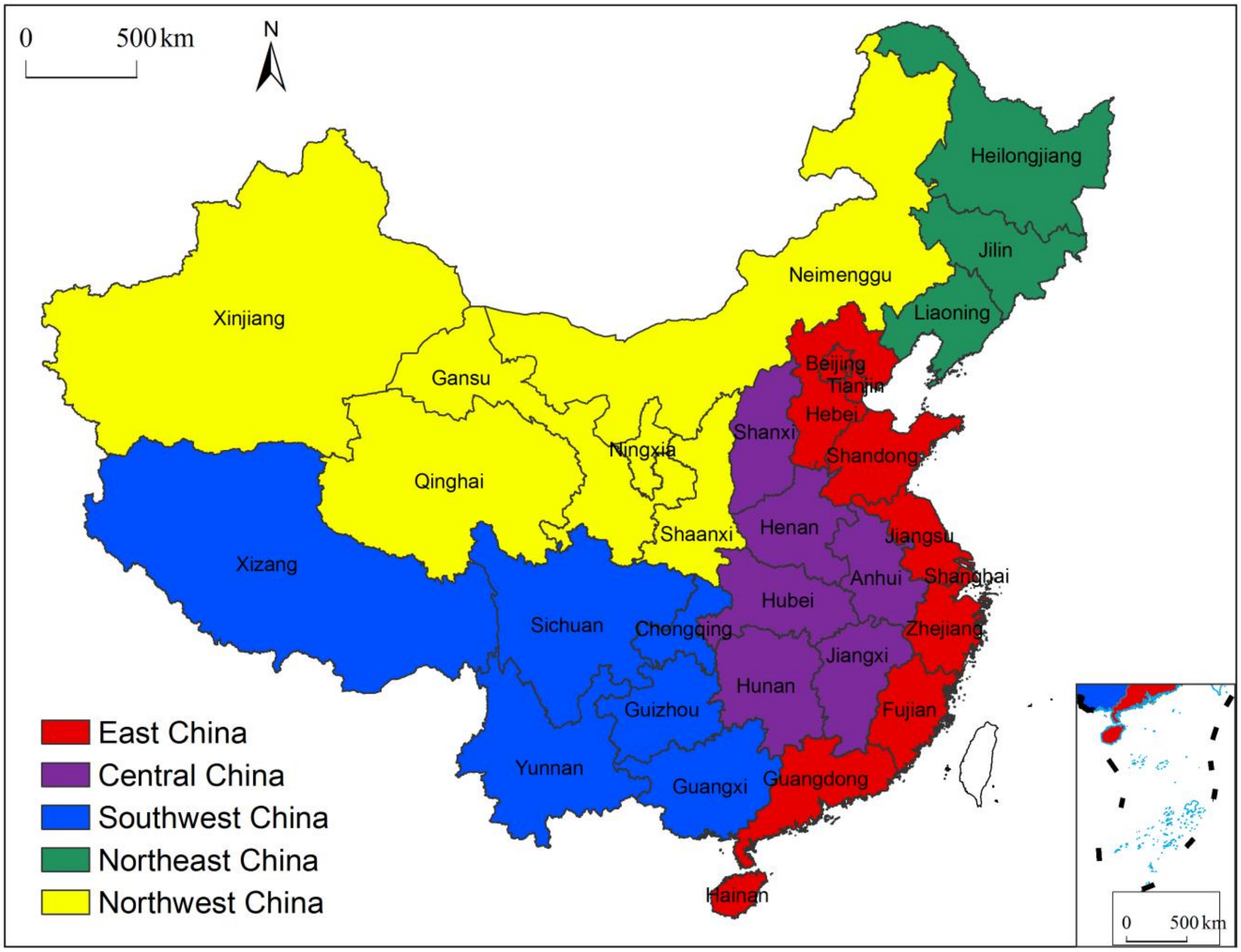
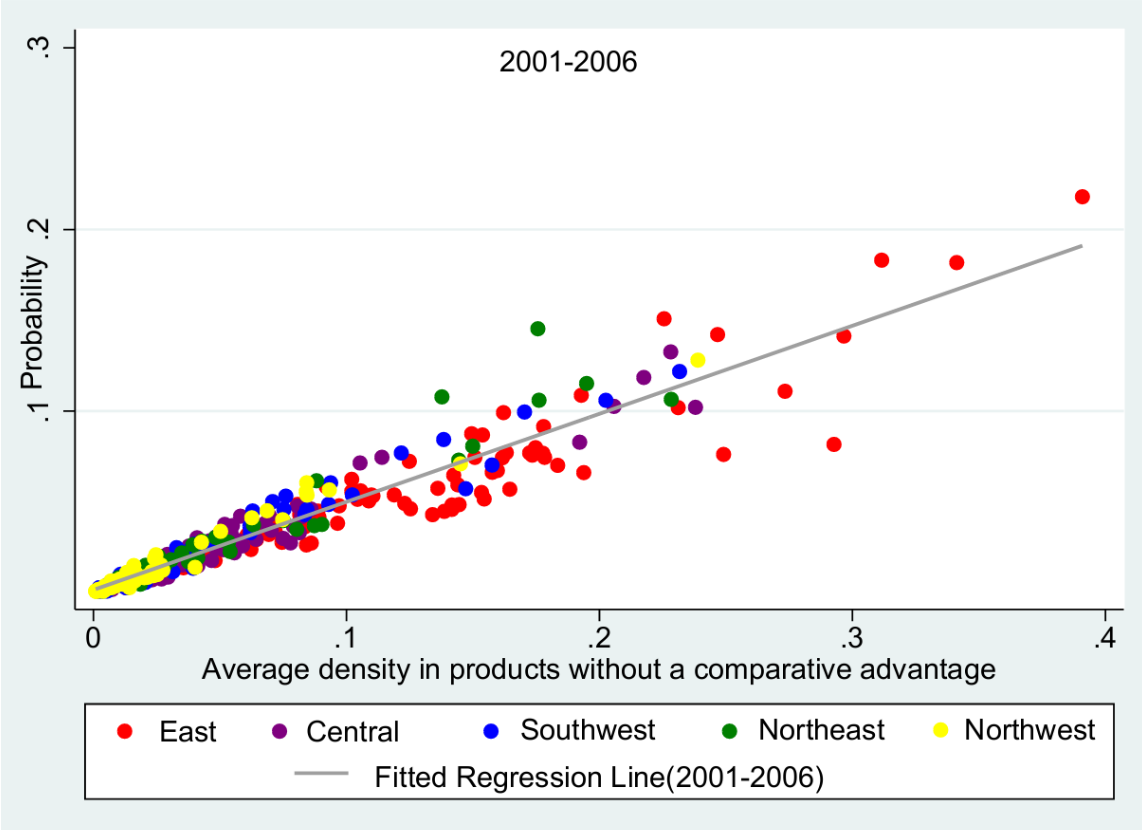


Figure 1 Provinces and five big regions in China

As can be seen in Figure 2a, on average, cities in relatively developed, wealthy East China have higher level of average density of the industries without a comparative advantage in 2001 and higher level of probability of developing a comparative advantage in a new industry in 2006 than do cities in the rest of China, while cities in Northwest China have the lowest scores on both aspects. There is a clear positive relationship between the average density of undeveloped industries in a city in 2001 and the probability of this city developing a transition industry in 2006. For example, the three dots in the upper, rightmost area are Shanghai, Nanjing and Beijing from East China that have the highest level of average density of undeveloped industries in 2001 (0.391, 0.341 and 0.312, respectively) and the highest probability of developing a transition industry in 2006 (0.218, 0.182 and 0.183, respectively). In contrast, Guyuan and Dingxi in Northwest China have the lowest level of average density of undeveloped industries in 2001 (0.00089 and 0.00091 respectively) and the lowest level of probability of developing a transition industry in 2006 (0.00092 and 0.00093 respectively). In Figure 2a, all dots are close to the fitted regression line, R^2 of which is as high as 0.9109, suggesting that Chinese regions were heavily reliant on industrial relatedness while jumping into new industries during 2001-2006.


2001-2006
Probability
Average density in products without a comparative advantage
East
Central
Southwest
Northeast
Northwest
Fitted Regression Line(2001-2006)

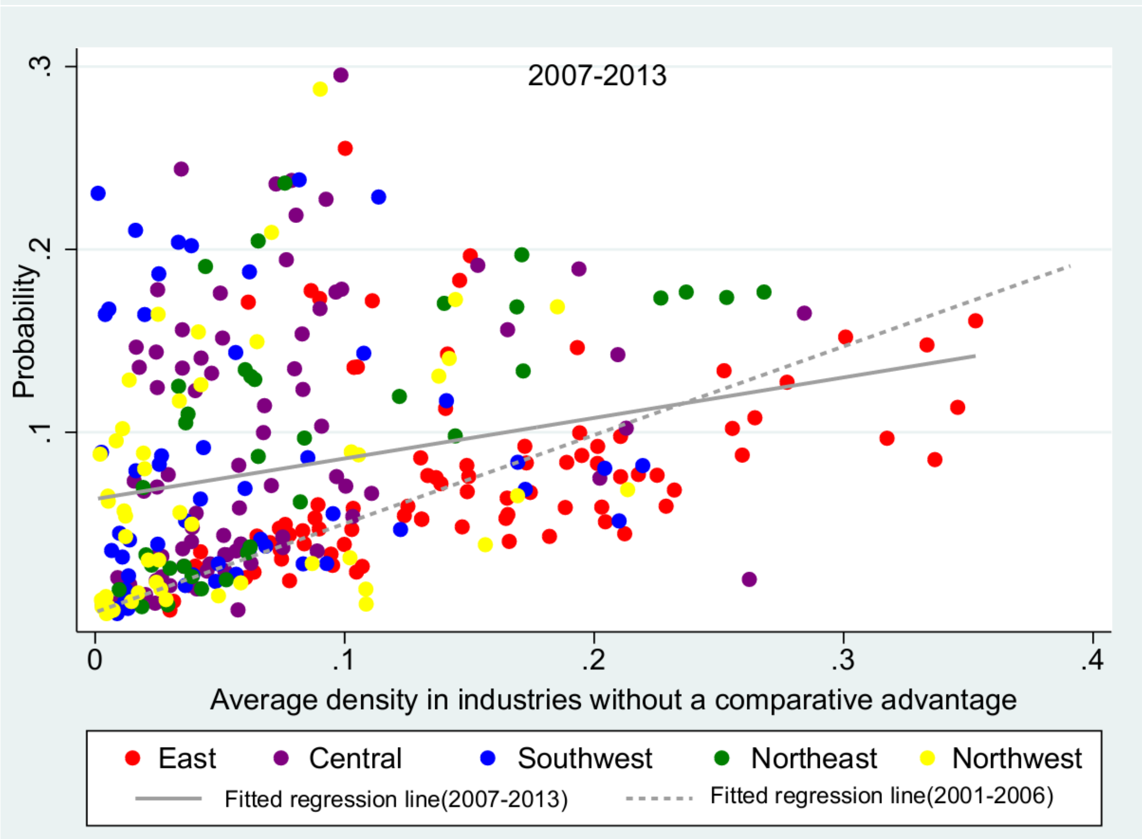

2007-2013
Probability
Average density in industries without a comparative advantage
East
Central
Southwest
Northeast
Northwest
Fitted regression line(2007-2013)
Fitted regression line(2001-2006)

Figure 2a (top) and 2b (bottom) Relationship between the average density of the industries without a comparative advantage in a city in 2001 (or 2007) and the probability of this city developing a comparative advantage in a new industry in 2006 (or 2013)

As is shown in Figure 2b, during 2007-2013, even though the average density of undeveloped industries in a city in 2007 is still positively related with the probability of this city developing a transition industry in 2013, this relationship has been weakened. First, the gradient of 2001-2006 fitted regression line is 0.485, while that of 2007-2013 fitted regression line is 0.222, indicating a decreasing impact of average density over probability of developing transition industries. Second, a considerable number of dots are located far away from the fitted regression line, and R^2 of the 2007-2013 fitted regression line is 0.0686, much lower than that of the 2001-2006 fitted regression line. This means some other factors have reduced regions' reliance on density in the process of industrial diversification, and enhanced regions' jumping capabilities. Third, the increase of jumping capability is particularly phenomenal in Central, Northwest, Northeast and Southwest China, as some cities with low level of density (for example below 0.05) in 2007 still have high probability of developing transition industries in 2013. In contrast, most cities in East China are still located close to the 2001-2006 fitted regression line (the dotted line in Figure 2b). In a word, the evolution of regional industrial structure should not be understood solely based on industrial relatedness as well as the distance between new industries and regions' existing productive structure, since the interaction between industrial relatedness and regional diversification is increasingly inflected by other factors that enable regions to jump further, particularly in Central, Northwest, Northeast and Southwest China.

5. Empirical Results

Since the dependent variable is binary and the econometric equation includes a large number of dummy variables, we follow Boschma et al. (2013) and estimate equation (4) with a linear probability-OLS model. Correlation analysis indicates that correlations of most independent variables are moderate or low, suggesting no serious problem of multi-collinearity. Some highly correlated terms are separated into different models[4]. Table 2 reports the econometric results. First, density has a positive effect, which is consistent with theoretical prediction. The parameter of comparative advantage in 2001 (2006) is also positive and significant, suggesting that having a comparative advantage at the beginning of the period raises the probability of having a comparative advantage at the end of the period. This echoes with the findings of Boschma et al. (2013). Second, as notified above, we distinguish the role that independent variables and interaction terms play in developing a comparative advantage in *new* industries ("on new" variables in Table 2) from their contribution in keeping a comparative advantage in *current* industries ("on current" variables). Although the coefficients of some are insignificant, most independent variables in Model (1), (2), (5) and (6) have positive and significant effects, indicating the important role of extra-regional linkages and internal innovation in developing transition industries and maintaining developed industries.

[4] For example, FDI and Import; PCAP and HCAP. In addition, $x_{i,c,t1}$ is highly correlated with some of its interaction terms, therefore it is excluded in model (3), (4), (7) and (8) in Table 2. Density is highly correlated with PCAP, it is thus not included in model (1), (2), (5) and (6) in Table 2.

Moving on to the results connected more closely with the central argument (Model (3), (4), (7) and (8)), we first focus on the role of the interaction terms between density and *EXT*/*INT* variables in developing a comparative advantage in *new* industries ("on new" variables in Model (3), (4), (7) and (8)). The interaction term between FDI and density presents a positive and significant sign in the first stage and an insignificant sign in the second, indicating that FDI has not brought into path-breaking knowledge that can further leads to radical innovation. Instead, most foreign-owned firms in China either engage in industries that are closely related to China's existing industrial structure or are not greatly embedded in China's economy which inhibits knowledge spillover from foreign-owned firms to domestic firms (also see (Wei 2010) for an example of weak local embeddedness of foreign-owned firms in China). In contrast, the interaction term between density and the other *EXT* variable, Import, has a negative and statistically significant sign throughout these two stages, suggesting that Chinese firms are able to obtain path-breaking knowledge from imports. The learning opportunity provided by new knowledge embedded in imported products reduces regions' reliance on industrial relatedness and enables the latter to jump further into less related industries. This echoes with other empirical studies (Zhu and Fu 2013).

The interaction term between R&D and density has a negative and significant sign in 2001-2006 models, but a positive sign in 2007-2013 models. This means Chinese cities have benefited from radical innovation in 2001-2006 and managed to jump further, whereas in 2007-2013 most enterprises' investment in R&D has only generated path-dependent new knowledge, and technological changes have been largely driven by incremental, gradual innovation into related industries. Government supports in the form of export tax rebate only play a supportive and secondary role in the creation of radical innovation and of course could not generate path-breaking knowledge on its own. High level of physical and human capital are likely to enable regions to achieve more radical innovation and jump further into less related industries, which is consistent with our expectation.

We move on to the "on current" variables in Model (3), (4), (7) and (8). Another problem plaguing developing countries/regions is even if they manage to jump further and develop a comparative advantage in new industries that are not close to their existing productive structure, it is still difficult for them to *maintain* a comparative advantage in these technologically distant and less related industries (Neffke, Henning, and Boschma 2011). Table 2 shows that regions can overcome this problem by improving local physical and human capital. Even though government supports in the form of export tax rebate could not generate path-breaking new knowledge, they help regions to stand firm in technologically distant industries, once regions jump into these industries. Finally, regions should keep importing path-breaking knowledge in order to maintain a comparative advantage in less related industries. In a word, extra-regional linkages and internal innovation have the potential to reduce regions' reliance on industrial relatedness in the process of developing a comparative advantage in new industries and keeping a comparative advantage in current industries.

Table 2 Determinants of having developed industries in China

	2001-2006				2007-2013			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Comparative advantage	0.477***	0.492***			0.493***	0.539***		
density			1.071***	1.407***			0.984***	1.197***

FDI			0.171***				0.132***	
Import				3.25e-10***				1.28e-10***
Rebate			-0.003***	-0.003***			-0.002***	-0.002***
R&D			-0.285	-0.049			-0.075	-0.158
PCAP			-0.025***				0.0569***	
HCAP				-0.00003*				0.0001***
FDI on current	0.132***				0.155***			
FDI on new	0.103***				0.065***			
Import on current		1.98e-10***				1.89e-11***		
Import on new		1.54e-10***				4.56e-11***		
Rebate on current	0.014***	0.012***			0.0145***	0.016***		
Rebate on new	0.003***	0.003***			0.002***	0.002***		
R&D on current	-0.000005	-0.000009			0.000002*	0.000001		
R&D on new	0.000001	-0.0000005			0.000001**	0.0000003		
PCAP on current	0.037***				0.015***			
PCAP on new	0.040***				-0.003			
HCAP on current		0.0002***				0.00004***		
HCAP on new		0.0002***				0.00006***		
FDI*density on current			7.96e-10***				8.66e-11***	
FDI*density on new			9.52e-10*				-6.84e-11	
Import*density on current				-5.44e-10**				-2.99e-10***
Import*density on new				-4.43e-10**				-2.84e-10***
Rebate*density on current			-0.228***	-0.233***			-0.069***	-0.064***
Rebate*density on new			0.102***	0.100***			0.050***	0.054***
R&D*density on current			0.00004*	0.00003*			0.000007**	0.000006*
R&D*density on new			-0.00003***	-0.00002***			0.0000009	0.000005*
PCAP*density on current			0.014				-0.041*	
PCAP*density on new			-0.059*				-0.514***	
HCAP*density on current				-0.0004***				-0.0002***
HCAP*density on new				-0.001***				-0.001***
HS	included	included	included	included	included	included	included	included
Province	included	included	included	included	included	included	included	included
_cons	0.153***	0.102***	-0.0679***	-0.0585***	0.111***	0.0890***	0.0241	0.0111
N	281960	277720	281960	277720	281960	281960	281960	281960
R^2	0.228	0.227	0.198	0.190	0.202	0.199	0.164	0.159
F	699.5	685.8	564.1	530.5	598.6	590.1	448.3	434.1

Notes: *significant at the 10% level; **significant at the 5% level; and ***significant at the 1% level.

It is also worthwhile to point out that the absolute value of the coefficient of "PCAP*density on new" has increased dramatically from the first stage to the second (Model (3) and (7) in Table 2), indicating that China's investment on infrastructure particularly during 2007-2013 has enabled its regions to achieve more radical innovations and become less reliant on industrial relatedness while entering into new industries. This may also contribute to the temporal change of the relationship between density and the probability of developing transition industries, as presented in Figure 2. To better understand the temporal change showed in Figure 2, we compare estimation results in East and Northeast China since such a temporal change varies across China's big regions[5].

[5] As notified above, the temporal change of the relationship between density and the probability of developing transition industries is relatively small in East China, but is quite remarkable in Central,

Since we seek to examine the temporal change of the relationship between density and the probability of developing a comparative advantage in new industries here, we only focus on the coefficients of "on new" variables. As is shown in Table 3, in East China, even though the role of physical capital in reducing regions' reliance on industrial relatedness has been enhanced from the first stage (coefficient: -0.082) to the second (coefficient: -0.397), other factors remain relatively stable. The role of R&D in reducing regions' reliance on industrial relatedness has even faded away during this time period. In contrast, in Northeast China, extra-regional linkage variables have changed from strengthening the impact of industrial relatedness on industrial diversification towards reducing regions' reliance on relatedness, even though FDI's impact is still not significant. The impact of physical and human capital remains the same. However, the former's impact has drastically increased (the absolute value of the coefficient has increased from 0.674 to 3.578). In addition, not only has human capital's impact risen greatly, but it has also turned significant in the second stage.

During 2007-2013, import, physical and human capital have reduced regions' reliance on industrial relatedness in Northeast China to a greater extent than in East China (see the coefficients of Import*density on new, PCAP*density on new and HCAP*density on new in Model (3), (4), (7) and (8) in Table 3). This provides an explanation to the regional disparity shown in Figure 2 to some extent. In China's relatively less developed inland regions, establishing extra-regional linkages and investing in education and infrastructure both contribute to reducing regions' reliance on relatedness, generating path-breaking new knowledge, and enabling regions to jump into technologically distant industries that are less related or unrelated to regions' existing industrial structure. Given this, industrial evolution in inland China exhibits a much more path-breaking pattern during 2007-2013, probably due not only to the industrial relocation from coastal regions to inland China but also to China's central government's efforts to alleviate regional disparity and boost economic development in relatively less developed inland China through for example Western Development Strategy, Revitalize Northeast China Initiative, and Rise of Central China strategy.

Table 3 Determinants of having developed industries in East and Northeast China

	East China				Northeast China			
	2001-2006		2007-2013		2001-2006		2007-2013	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
density	1.275***	1.618***	1.246***	1.396***	0.914***	0.905***	1.214***	1.147***
FDI	0.147***		0.114***		0.160***		0.139***	
Import		2.07e-10*		1.06e-10***		1.44e-09*		1.03e-09***
Rebate	-0.00955***	-0.00971***	-0.009***	-0.009***	0.000443	0.00115	-0.001	-0.0009
R&D	-0.674	0.0843	-3.209	-3.255	-0.275	-0.375	1.131	0.806
PCAP	-0.0174*		0.045***		0.0311		0.482***	
HCAP		-0.000113***		0.0002***		-0.00003		0.0001***
FDI*density on current	8.45e-10***		7.55e-11***		3.88e-10		1.18e-09**	
FDI*density on new	9.02e-10		-1.18e-10		4.43e-09		-2.39e-09	

Southwest, Southeast and Northeast China. Here, due to space limitation, we only compare East China and Northeast China. Estimation results on other regions are available on request.

Import*density on current		-1.49e-10		-2.44e-10***		-7.48e-09***		-2.76e-09***
Import*density on new		-2.01e-10		-2.20e-10**		3.63e-11		-5.29e-09***
Rebate*density on current	-0.186***	-0.206***	-0.039***	-0.037***	-0.192***	-0.171***	-0.0602***	-0.052***
Rebate*density on new	0.137***	0.144***	0.088***	0.0899***	0.0548***	0.046***	0.0252***	0.024***
R&D*density on current	0.00003	0.00004	0.000002	0.0000004	0.0002*	0.0002	0.000154***	0.0001***
R&D*density on new	-0.00004***	-0.00004**	0.000002	0.000005	0.0000008	0.000006	0.0000113	0.00001
PCAP*density on current	-0.025		0.014		3.181***		-1.246***	
PCAP*density on new	-0.082*		-0.397***		-0.674*		-3.578***	
HCAP*density on current		-0.0007***		-0.00006		0.005***		0.0006***
HCAP*density on new		-0.0009***		-0.001***		-0.0003		-0.002***
HS	included	included	included	included	included	included	included	included
Province	included	included	included	included	included	included	included	included
_cons	-0.0838***	-0.0639***	0.00568	-0.0189	-0.0467*	-0.0397	-0.0420	-0.0802**
N	91160	91160	91160	91160	33920	33920	33920	33920
R^2	0.227	0.221	0.221	0.215	0.148	0.144	0.189	0.184
F	259.9	250.5	251.3	242.6	61.28	59.37	81.85	79.19

Notes: *significant at the 10% level; **significant at the 5% level; and ***significant at the 1% level.

Table 4a and 4b report the empirical results on whether the impact of the articulation between density indicator and *EXT*/*INT* variables vary across industries. We here focus on four industries: HS11 (textile and apparel), HS16 (mechanical and electrical equipment), HS17 (transport) and HS18 (high-tech instruments and equipment). Again, since the central question is how to jump further into new industries, this part also mainly focuses on the “on new” variables, unless otherwise stated. FDI in technology- and capital-intensive HS16 and HS18 plays a negligible impact over regional path-breaking, indicating that China’s technology-intensive industries have not been able to achieve radical innovation and upgrading through learning from foreign-owned firms, due either to the weak local embeddedness of FDI or to the fact that transnational corporations, particularly in high-tech industries, have only relocated their relatively low-end and low-tech production to China while leaving the high-end functions (R&D and headquarters) in the North during this time period. In HS11, FDI reinforces the effect of density over regional diversification, indicating that most foreign-owned firms in China’s textile and apparel industry have not brought in path-breaking knowledge and are manufacturing products that are closely related to China’s existing productive structure. However, the impact of the interaction term between density and FDI on developing a comparative advantage in new industries in HS 17 is negative and significant (Model (3), Table 4b), indicating the joint-venture firms in China’s automobile industry may bring in more distant and less related knowledge and result in more radical technological changes (see also (Liu and Dicken 2006) for an example of the importance of joint-venture firms in the upgrading of China’s automobile industry). Import is particularly crucial in path-breaking regional diversification in the textile and apparel industry, as global buyers are increasingly outsourcing higher-value-added and high-end functions (e.g., original design manufacturing) to their Chinese suppliers (see also (Zhu and Pickles 2015)). In the meantime, import in the other three technology-intensive industries only has minor effects.

Government support in the form of export tax rebates has a quite consistent effect in all models, strengthening relatedness’s effect in generating new transition industries but enabling regions maintain a comparative advantage in technologically distant and less related industries. R&D plays a key role in breaking old path in the mechanical and electrical industry, while in the textile and apparel industry R&D investment has centered on incremental innovation. Infrastructure allows regions to jump further, in all models except in the textile and apparel industry model. Similarly, human capital also reduces regions' reliance on density and relatedness in all models except in the textile and apparel industry models. One reason for this inconsistency is that the human capital indicator we used is measured as the share of college students, who are likely to become white-collar workers engaging in knowledge- and technology-intensive industries. In contrast, industrial diversification in the textile and apparel industry may be more related to the share of blue-collar labor force that is largely trained by vocational and technical schools in China.

Table 4a (top) and b (bottom) Determinants of having developed industries in different industries

	HS11. Textile and textile articles				HS16. Machinery and mechanical appliances and electrical equipment			
	2001-2006		2007-2013		2001-2006		2007-2013	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
density	0.874***	1.227***	0.583***	0.774***	1.445***	1.692***	1.157***	1.296***
FDI	0.132***		-0.0121		0.144***		0.118***	
Import		2.02e-09***		1.44e-09**		7.30e-11		5.91e-11*
Rebate	-0.003	-0.00422*	-0.005***	-0.00615***	-0.00844***	-0.00859***	-0.0192***	-0.0191***
R&D	-0.138	-1.014	-2.701	-0.941	-1.206	-1.773	-0.249	-0.344
PCAP	-0.0404*		0.0892***		-0.0437*		0.0334*	
HCAP		-0.0002***		-0.00004*		0.0000119		0.000126***
FDI*density on current	1.70e-09*		4.58e-09***		1.64e-10		2.90e-12	
FDI*density on new	-8.24e-10		3.33e-09**		1.75e-09**		1.44e-10	
Import*density on current		-9.82e-09***		-1.95e-09		-5.54e-11		-1.36e-10*
Import*density on new		-4.87e-09**		-3.76e-09*		2.77e-10		-8.99e-11
Rebate*density on current	-0.342***	-0.442***	-0.250***	-0.275***	-0.250***	-0.281***	-0.239***	-0.262***
Rebate*density on new	0.0933***	0.114***	0.0914***	0.107***	0.216***	0.213***	0.250***	0.247***
R&D*density on current	-0.000364	0.000431	0.0001***	0.0001***	0.002***	0.00267***	-0.00002	0.00002
R&D*density on new	0.000181	0.000230	0.00003**	0.00003*	-0.0004*	-0.000320*	-0.00006**	-0.00004*
PCAP*density on current	-0.385***		-0.0990		-0.0862		0.0953	
PCAP*density on new	0.0543		-0.0299		0.160		-0.322***	
HCAP*density on current		-0.00297***		-0.0005***		-0.001***		-0.0002*
HCAP*density on new		0.00006		-0.00007		-0.0007***		-0.0008***
HS	included	included	included	included	included	included	included	included
Province	included	included	included	included	included	included	included	included
_cons	-0.154***	-0.162***	0.0313	0.0238	-0.134***	-0.129***	0.160***	0.150***
N	40166	39562	40166	40166	34314	33798	34314	34314
R^2	0.242	0.247	0.229	0.225	0.242	0.235	0.197	0.191
F	232.8	235.1	216.7	211.7	255.0	241.2	195.3	188.5

	HS17. Vehicles, aircraft, vessels and associated transport equipment				HS18. Optical, photographic, cinematographic, measuring, checking, precision, medical or surgical instruments and apparatus			
	2001-2006		2007-2013		2001-2006		2007-2013	
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
density	1.215***	1.393***	0.939***	0.875***	1.464***	1.734***	0.887***	1.131***

FDI	0.021		0.230***		0.126***		0.121***	
Import		1.49e-09***		8.99e-11		1.77e-09*		2.62e-10*
Rebate	0.002	0.00338	-0.006***	-0.005***	-0.0143***	-0.0155***	-0.0208***	-0.0215***
R&D	-0.019	-0.0614	1.127	0.0964	-1.733	-1.273	0.536	0.601
PCAP	0.238***		0.0883***		-0.004		0.0910***	
HCAP		0.00004		0.0002***		-0.000131**		0.000186***
FDI*density on current	1.42e-09*		5.75e-10**		1.87e-09		3.96e-10***	
FDI*density on new	-3.35e-09		-2.22e-09*		5.66e-08***		2.82e-10	
Import*density on current		-3.24e-09***		2.97e-10		-3.87e-09		-6.44e-10
Import*density on new		-3.39e-09**		-3.27e-10		7.17e-09**		-3.06e-10
Rebate*density on current	-0.240***	-0.239***	-0.130***	-0.144***	-0.208***	-0.266***	-0.153***	-0.150***
Rebate*density on new	0.106***	0.0696***	0.0757***	0.055***	0.283***	0.304***	0.147***	0.155***
R&D*density on current	-0.00002	0.000004	0.00009***	0.000008	0.00008	0.0001	-0.0000008	0.0000181
R&D*density on new	-0.00001	-0.000001	-0.000006	0.000001	-0.00009**	-0.00006	0.0000006	0.0000157
PCAP*density on current	-0.0616		0.295**		0.135		-0.0879	
PCAP*density on new	-0.515***		-0.705***		0.242*		-0.397***	
HCAP*density on current		0.0009***		0.001***		-0.0008**		-0.0003
HCAP*density on new		-0.002***		-0.001***		0.0002		-0.001***
HS	included	included	included	included	included	included	included	included
Province	included	included	included	included	included	included	included	included
_cons	0.130**	-0.00161	0.265***	0.191***	-0.154***	-0.158***	0.0414	0.0180
N	9842	9694	9842	9842	14098	13886	14098	14098
R^2	0.218	0.215	0.210	0.210	0.274	0.267	0.171	0.167
F	60.80	58.69	58.00	57.79	120.2	114.8	65.79	63.83

Notes: *significant at the 10% level; **significant at the 5% level; and ***significant at the 1% level.

As a robustness check, all models are also estimated by probit model, by using data in different years (for example, 2002-2006 and 2007-2012), and by using different threshold values (0.8 and 1.2) to determine a comparative advantage. Compared with the results presented above, these changes produce only minor effects. Finally, we examine the transition industries in 2001-2006 and 2007-2013 (Figure 3). First, average density of transition industries in all five regions has increased from the first stage to the second, as a result of rapid regional diversification and branching in the entire country throughout this time period. Second, transition industries in East China have higher density level than those in the rest of China, suggesting that China's coastal regions are more path-dependent while inland China is relatively more path-breaking. This resonates with our findings based on Figure 2 and econometric analyses, pointing out the need to pay attention to a broader range of factors besides industrial relatedness. Third, in the upper-left areas of Figure 3a and 3b, transition industries with low density and high unit value include HS16 (Machinery and mechanical appliances and electrical equipment) and HS17 (Vehicles, aircraft, vessels and associated transport equipment) mainly located in inland China, most of which are technology-intensive, high-value, and high-end industries that are distant from the existing regional industrial structure of the less developed inland China. Inland China's diversification into high-value industries, as notified above, could be attributed not only to the industrial relocation from coastal regions to inland China but also to China's central government's efforts to alleviate regional disparity and boost economic development in relatively less developed inland China. Whether or not inland China is able to maintain a comparative advantage in these technologically distant and less related industries may depend on the extent to which regions in inland China can continuously improve infrastructure and education, provide government supports, and promote the technology content of imports. In doing so, less developed regions may be able to jump from the peripheral, deserted areas to the core, dense areas in the uneven industry space/network.

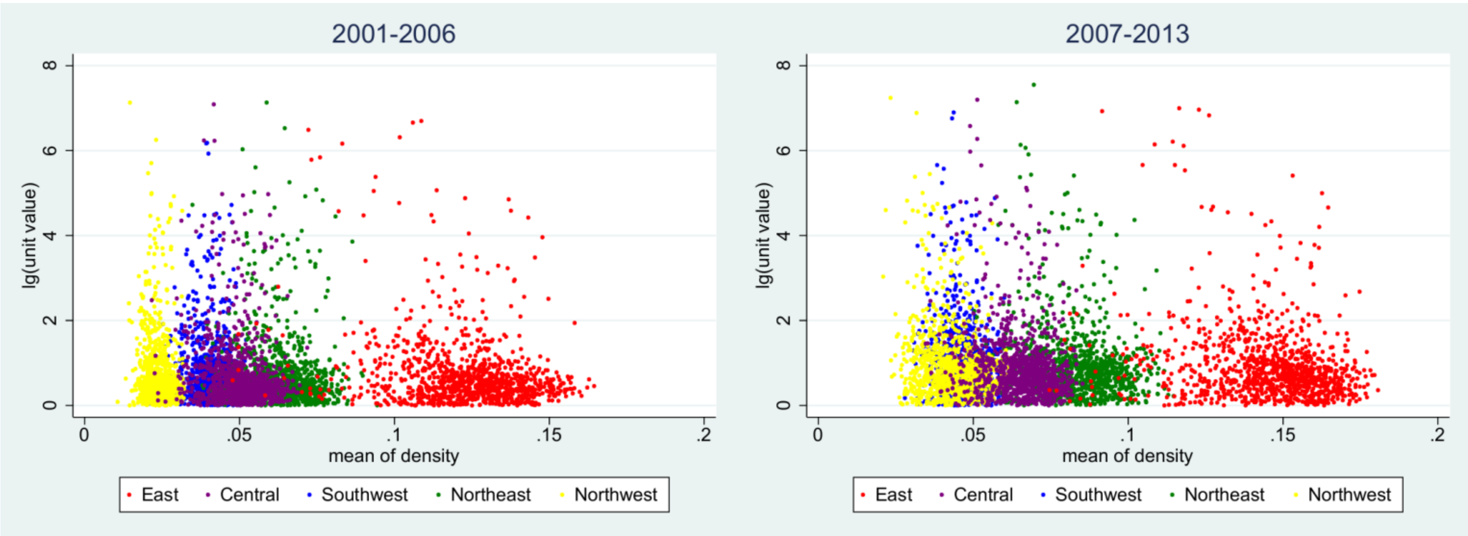


Figure 3a (left) and b (right) Unit value and average density of transition industries in China’s five regions during 2001-2006 and 2007-2013

6. Conclusion

By examining the regional diversification in China’s different industries and different regions through the proximity product index developed by Hidalgo et al. (2007), we seek to shed new light on the debates about the articulation between industrial relatedness and regional evolution. Technological relatedness, as a key driving force of regional economic development, not only affects the growth of existing industries through externalities derived from related variety, but is also responsible for regions’ entry into new industries and formation of new growth pathways. Recent studies tend to emphasize the ways in which regional diversification emerged as a path-dependent process, since regions often branch into industries that are technologically close and related to their current industrial structure. Such a path-dependent regional diversification can be thought as regions/countries jumping in a heterogeneous and uneven industry space where they are only allowed to jump certain distance. This distance is determined by technological relatedness among industries. It is thereafter argued that developed countries that start from the core, dense areas in the uneven industry space have more opportunities to jump to new related industries and therefore have more opportunities to sustain economic growth than do developing countries that jump from peripheral, deserted areas.

Even though researches based on industrial relatedness do not exclude the possibility that developing countries can reach the core area in the industry space, empirical studies are mostly centered on regional diversification in developed countries and highlight the key role of path-dependence in the process of regional economic development. In some extreme cases, it is impossible for developing countries to enter the core area, and the divergence between developed and developing countries persists due to the path-dependent development trajectories. Such a conclusion is quite pessimistic particularly for developing countries/regions, since it predicates too much on an assumption that regional diversification is affected or even confined by relatedness among industries, but pays less attention to whether or not countries/regions' jumping capabilities in the uneven industry space can be changed.

In this paper, we differentiate two types of new growth path creation—path-dependent and path breaking, and focus on the second type in order to examine how developing countries/regions can jump further in a more path-breaking way, so that the "confinement" of technological relatedness can be transcended. Based on China's export data, this paper testifies that regions can jump further by investing in extra-regional linkages and internal innovation. Not only do these two sets of factors promote regions' jumping capability and reduce regions' reliance on industrial relatedness, but they also contribute to regions' capability of maintaining a comparative advantage in these technologically distant and less related industries. In addition, different extra-regional linkages and internal innovation factors have affected path-breaking regional diversification to different extents in China, and these effects also vary across regions and industries.

Empirically, this research seeks to find a brighter future for developing countries/regions in an increasingly competitive global economy by pointing out that the seemingly dominant path-dependent development trajectories can be broken through continuously improving infrastructure and education, providing government supports, and promoting the technology content of imports. In doing so, less developed regions may be able to catch up and jump from the peripheral, deserted areas further into the core, dense areas in the uneven industry space/network.

Theoretical studies in EEG have long pointed out extra-regional linkages and internal innovation driven by purposeful actions of regional actors have the potential to bring in and generate fresh know-how and technologies that are less related to region's existing productive structure, therefore may enable regions to dislodge path-dependence and develop in a more path-breaking way. Empirical works based on case studies also testify this argument. This paper thus provides a quantitative research that is complementary to existing theoretical and qualitative studies. In addition, our research is based on recent studies (Boschma and Capone 2014a; Boschma, Minondo, and Navarro 2013; Neffke, Henning, and Boschma 2011), but asks questions from a different angle and includes some economic and institutional factors that have been left out in previous studies. Not only have we investigated the role of economic and institutional factors at the regional level, but we also explore the ways in which their impacts over path-breaking regional diversification vary across industries and regions. Finally, it is also pointed out that changes of economic and institutional factors also have long-lasting consequences on the process of industrial diversification, through comparing China's regional economic development in two stages.

References

Amighini, A. and M. Sanfilippo. 2014. Impact of South–South FDI and Trade on the Export Upgrading of African Economies. *World Development* **64:1-17**.

Arthur, W.B. 1989. Competing Technologies, Increasing Returns, and Lock-In by Historical Events. *The Economic Journal* **99:116-131**.

———. 1994. *Increasing returns and path dependence in the economy*. Ann Arbor: University of Michigan Press.

Bathelt, H. and P.-F. Li. 2013. Global cluster networks—foreign direct investment flows from Canada to China. *Journal of Economic Geography*.

Bathelt, H.; A. Malmberg; and P. Maskell. 2004. Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress In Human Geography* **28:31-56**.

Baum, J.A.C. 1996. Organizational ecology. In *Handbook of Organization Studies*, ed. S.R. Clegg; C. Hardy; T. Lawrence; and W.R. Nord, 77-114. London: Sage.

Boschma, R. and G. Capone. 2014a. Institutions and diversification: related versus unrelated diversification in a varieties of capitalism framework. Urban & Regional Research Center, Utrecht University, Utrecht.

———. 2014b. Relatedness and diversification in the EU-27 and ENP countries. Urban & Regional Research Center, Utrecht University, Utrecht.

Boschma, R. and S. Iammarino. 2009. Related Variety, Trade Linkages, and Regional Growth in Italy. *Economic Geography* **85:289-311**.

Boschma, R.; A. Minondo; and M. Navarro. 2012. Related variety and regional growth in Spain. *Papers in Regional Science* **91:241-256**.

———. 2013. The Emergence of New Industries at the Regional Level in Spain: A Proximity Approach Based on Product Relatedness. *Economic Geography* **89:29-51**.

Bryce, D.J. and S.G. Winter. 2009. A General Interindustry Relatedness Index. *Management Science* **55:1570-1585**.

David, P.A. 1985. Clio and the Economics of QWERTY. *The American Economic Review* **75:332-337**.

Dollar, D. 1992. Outward-Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976-1985. *Economic Development and Cultural Change* **40:523-544**.

Frenken, K.; F. Van Oort; and T. Verburg. 2007. Related Variety, Unrelated Variety and Regional Economic Growth. *Regional Studies* **41:685-697**.

Fu, X. and Y. Gong. 2011. Indigenous and Foreign Innovation Efforts and Drivers of Technological Upgrading: Evidence from China. *World Development* **39:1213-1225**.

Görg, H. and D. Greenaway. 2004. Much Ado about Nothing? Do Domestic Firms Really Benefit from Foreign Direct Investment? *The World Bank Research Observer* **19:171-197**.

Goldberg, P.K.; A.K. Khandelwal; N. Pavcnik; and P. Topalova. 2010. Imported Intermediate Inputs and Domestic Product Growth: Evidence from India. *The Quarterly Journal of Economics* **125:1727-1767**.

Hassink, R. 2005. How to unlock regional economies from path dependency? From learning region to learning cluster. *European Planning Studies* **13:521-535**.

Hausmann, R. and B. Klinger. 2007. The Structure of the Product Space and the Evolution of Comparative Advantage. Center for International Development, Harvard University, CID Working Paper No. 146. Cambridge, Massachusetts.

Henning, M.; E. Stam; and R. Wenting. 2013. Path Dependence Research in Regional Economic Development: Cacophony or Knowledge Accumulation? *Regional Studies***:1-15**.

Hidalgo, C.A. and R. Hausmann. 2008. A Network View of Economic Development. *Developing Alternatives* **12:5-10**.

Hidalgo, C.A.; B. Klinger; A.-L. Barabási; and R. Hausmann. 2007. The Product Space Conditions the Development of Nations. *Science* **317:482-487**.

Lall, S. and R. Narula. 2004. Foreign Direct Investment and its Role in Economic Development: Do We Need a New Agenda? *The European Journal of Development Research* **16:447-464**.

Liu, W. and P. Dicken. 2006. Transnational Corporations and 'Obligated Embeddedness': Foreign Direct Investment in China's Automobile Industry. *Environment and Planning A* **38:1229-1247**.

Martin, R. 2010. Roepke Lecture in Economic Geography—Rethinking Regional Path Dependence: Beyond Lock-in to Evolution. *Economic Geography* **86:1-27**.

Martin, R. and P. Sunley. 2006. Path dependence and regional economic evolution. *Journal of Economic Geography* **6:395-437**.

———. 2011. Conceptualizing Cluster Evolution: Beyond the Life Cycle Model? *Regional Studies* **45:1299-1318**.

Maskell, P. and A. Malmberg. 2007. Myopia, knowledge development and cluster evolution. *Journal of Economic Geography* **7:603-618**.

Neffke, F.; M. Henning; and R. Boschma. 2011. How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. *Economic Geography* **87:237-265**.

Poncet, S. and F. Starosta de Waldemar. 2013. Export Upgrading and Growth: The Prerequisite of Domestic Embeddedness. *World Development* **51:104-118**.

Porter, M. 2003. The Economic Performance of Regions. *Regional Studies* **37:545-546**.

Rigby, D.L. and J. Essletzbichler. 1997. Evolution, Process Variety, and Regional Trajectories of Technological Change in U.S. Manufacturing*. *Economic Geography* **73:269-284**.

Romer, P.M. 1990. Endogenous Technological Change. *Journal of Political Economy* **98:S71-S102**.

Schiff, M. and Y. Wang. 2006. North-South and South-South trade-related technology diffusion: an industry-level analysis of direct and indirect effects

La diffusion de la technologie reliée au commerce Nord-Sud ou Sud-Sud: une analyse industrielle des effets directs et indirects. . *Canadian Journal of Economics/Revue canadienne d'économique* **39:831-844**.

Schumpeter, J.A. 1947. The Creative Response in Economic History. *The Journal of Economic History* **7:149-159**.

Sturgeon, T.; J. Van Biesebroeck; and G. Gereffi. 2008. Value chains, networks and clusters: reframing the global automotive industry. *Journal of Economic Geography* **8:297-321**.

Sydow, J.; F. Lerch; and U. Staber. 2010. Planning for Path Dependence? The Case of a Network in the Berlin-Brandenburg Optics Cluster. *Economic Geography* **86:173-195**.

Teece, D.J.; R. Rumelt; G. Dosi; and S. Winter. 1994. Understanding corporate coherence: Theory and evidence. *Journal of Economic Behavior & Organization* **23:1-30**.

Wei, Y.H.D. 2010. Beyond new regionalism, beyond global production networks: remaking the Sunan model, China. *Environment and Planning C: Government and Policy* **28:72-96**.

Zhu, S. and X. Fu. 2013. Drivers of Export Upgrading. *World Development* **51:221-233**.

Zhu, S. and J. Pickles. 2015. Turkishization of a Chinese apparel firm: fast fashion, regionalisation and the shift from global supplier to new end markets. *Cambridge Journal of Regions, Economy and Society*.