Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-025-05771-2 OPEN

Creating path breakthroughs with skill endowment and complementarity: evidence from the regional industrial evolution of China, 2000–2015

Xiaoqi Zhou[1,2], Dieter Franz Kogler[3], Junsong Wang[4], Jing Chen[1,2]✉ & Rongjun Ao[1,2]✉

This study explores how skill endowments and complementarity influence industrial development trajectories, particularly those leading to path-breaking economic activities. By drawing on the multidimensional cross-relatedness approach of evolutionary economic geography (EEG), industry-occupation cross-relatedness networks were constructed for 264 Chinese prefecture-level cities on the basis of Chinese industry and occupation data, and subsequently these were analysed to examine their interactions. Empirical evidence suggests that regions are inclined to develop new industries closely associated with skill structure. More importantly, regions with high-skilled labour agglomeration and high-low skilled labour co-agglomeration have a stronger ability to develop new industries less related to their existing industrial structures and achieve path breakthroughs. The effects of skill endowment and complementarity vary across different regions and industries. Specifically, the impact of high-low skilled labour agglomeration on industrial path breakthroughs is more pronounced in eastern China. Moreover, in regions with greater economic complexity, industrial path breakthroughs are more closely associated with skill complementarity. Labour-intensive industrial diversification exhibits the strongest relationship with skill structure, whereas the regional path dependence effect is more pronounced in areas with higher industrial complexity.

[1] Key Laboratory for Geographical Process Analysis and Simulation of Hubei Province, Central China Normal University, Wuhan, China. [2] College of Urban and Environmental Sciences, Central China Normal University, Wuhan, China. [3] Spatial Dynamics Lab, School of Architecture Planning and Environmental Policy, University College Dublin, Dublin, Ireland. [4] School of Geographic Sciences, East China Normal University, Shanghai, China. ✉email: jing.chen@ccnu.edu.cn; aorongjun8600@163.com

Introduction

The concept of relatedness has been widely used by evolutionary economic geographers to explain a variety of regional economic dynamics (Kogler et al. 2023a; Whittle and Kogler, 2020). Regional economic development is thus conceptualized as a metabolic process and characterized by the continuous emergence of new economic activities and the decline of older ones (Martin and Sunley, 2022; Xu et al. 2025). As far as regional evolution is concerned, current studies on economic structure mainly focus on the industry itself and propose different forms of evolutionary paths like path dependence and breakthroughs in the formation of new industries by industrial relatedness (Balland et al. 2019; Frenken and Boschma, 2007).

Recent research has indicated that human capital has turned into a major driver of urban growth (Duranton and Puga, 2023). Based on the Marshall labour pool hypothesis, the entry of new industries is strongly linked to labours' skill structure. However, some scholars have pointed out that existing studies lack sufficient attention to occupations or skill structures (Sunley et al. 2020). In addition, the spatial layout of the original global value chain and production network has been affected in the context of the current technological progress gradually replacing manual labour and the deepening international division of labour. Therefore, the regional industrial structure cannot fully describe its development foundation and capabilities (Feser, 2003; Thompson and Thompson, 1987). As such, focusing only on industries can lead to a biased understanding of regional development (Barbour and Markusen, 2007; Nolan et al. 2011).

In general, skills are closely related to human capital (Lise and Postel-Vinay, 2020). In empirical studies, skills are usually defined as the ability to perform a certain job (Florida et al. 2012) or have a high educational attainment. However, recent research shows that educational attainment does not fully reflect skill levels (Deegan et al. 2024). Occupations, as the types of tasks and responsibilities undertaken by workers in the production process, are increasingly being used to measure skills (Santoalha et al. 2021). Evolutionary economic geographers take occupations as a window into skill endowments and begin to explore occupational relatedness based on industrial relatedness (Elekes et al. 2023; Hane-Weijman et al. 2022), which better captures the relatedness of economic activities in terms of labour skills. Occupational relatedness involves the relatedness of different economic activities in terms of labour skill demand, which can accurately describe the relatedness degree of industrial activities based on labour skill. Existing studies on occupational relatedness have neither effectively separated the impact of labour skill from industrial relatedness nor considered the heterogeneity of labour skill.

Furthermore, emerging research on multidimensional cross-relatedness provides new ideas for understanding the impact of labour skill structure on industrial evolution. The study of relatedness in early evolutionary economic geography (EEG) focused on a single factor and lacked attention to the cross-relatedness between different elements (Catalán et al. 2022). Given that measures such as technological relatedness are aggregate and rely on *ex post* co-occurrence data, they cannot fully assess the impact of labour skill elements on regional industrial evolution. Chen et al. (2025), for instance, attempt to construct an industry–occupation cross-space to highlight the critical role of skill structures in shaping regional industrial dynamics. They find that Beijing not only has an advantageous industrial portfolio for biopharmaceutical development but also a matching occupational structure, while Dalian, despite lacking relevant industries, possesses sufficient occupational capabilities to support potential breakthroughs in that field. These cases suggest that regional industrial upgrading depends not only on existing industrial bases but also heavily on the configuration of skill endowments. Building on this foundation, this study advances the discussion by focusing on skill endowment and complementarity mechanisms, thereby providing a more nuanced understanding of how skill configurations contribute to regional industrial path breakthroughs.

By taking this approach, this article contributes to the literature in the following aspects. First, the environment in which evolutionary economic geographers explore various changes in the economic landscape has entered a state of "new normal" or "polyresins" (Kogler et al. 2023a). In the context of "new realities", a tension remains between transformation and "lock-in". Existing studies have examined the relationship between skill structures and industrial activities (Buyukyazici et al. 2024; Galetti et al. 2021), and have verified the skill-path dependency of industrial development (Chen et al. 2024; Zhou et al. 2025). However, the measurement of human capital in these studies has largely remained at the level of educational attainment (Zhu et al. 2017), and has not revealed the internal role of skill structures and their relationships in driving path breakthroughs. In this article, it was hypothesized that regions with high-skilled labour agglomeration and high-low skilled labour co-agglomeration are more capable of achieving industrial path breakthroughs.

Second, this article contributes to the literature on the impact of heterogeneous labour mobility on industrial structure upgrading by analysing the potential impact of skill endowment and complementary on regional industrial path breakthroughs. This is crucial because the positive effect of the increase of high-skilled labour on the upgrading of industrial structure has received more attention. In recent years, many cities in China have introduced various policies for attracting high-skilled labour. These high-skill-biased policies have led to increasing concentrations of high-skilled workers in large cities, while restricting the inflow of medium- and low-skilled labour, especially low-skilled workers (Gu et al. 2021). Whether large cities can absorb more low-skilled labour while attracting high-skilled workers remains an open question. Based on the data of Chinese industries and occupations from 2000 to 2015, therefore, an industry-occupation cross-relatedness network was constructed and incorporated into the industrial evolution mechanism as an influencing factor along with the spatial agglomeration of the heterogeneous labour force. Notably, since medium-skilled labour is often at risk of being replaced or faces "skill polarization" (Autor et al. 2006; Henning and Eriksson, 2021), this paper focuses on the two ends of the skill distribution—high-skilled and low-skilled workers. Focusing on these groups allows us to isolate how skill complementarity drives path-breaking industrial trajectories.

The results of this study indicate that an increase in the level of high-skilled labour agglomeration or high-low skilled labour co-agglomeration in a region will raise the likelihood of developing new industries less related to their existing industrial structures. It is applicable to different industries and regions, although the degree of impact is different. More specifically, the skill structure in eastern China has the most prominent impact on industrial path breakthroughs. More skill complementary effects will be needed by industrial path breakthroughs when the economic complexity of a city is higher. Meanwhile, industrial diversification in labour-intensive cities has the greatest correlation with skill structure. The path breakthroughs of capital-intensive industries are more related to high-skilled labour, and those of technology-intensive industries are more dependent on the skill complementary effect. In addition, the lower the industrial complexity is, the more it will be necessary to introduce highly skilled labour to achieve path breakthroughs.

The remainder of this study was organized as follows. In the second section, relevant literature was reviewed, and research hypotheses were developed. In the third section, data and methods were introduced. In the fourth section, descriptive analyses and empirical results were presented. In the final section, the results were concluded and discussed in detail, and possible research directions were provided.

Literature review and research hypotheses

Relatedness, cross-relatedness and regional diversification. Industrial relatedness is extensively used by evolutionary economic geographers to explain regional industrial dynamics (Boschma, 2017; Pylak and Kogler, 2021). Firms as specific regional actors are employed to answer how industrial relatedness drives regional diversification (Hervas-Oliver et al. 2024; Kogler et al. 2023b). Regional branching literature has focused on the micro mechanism of industrial relatedness from the angle of firm heterogeneity (Neffke et al. 2018; Neffke and Henning, 2013; Tanner, 2014). For one thing, firms generally choose to integrate multiple resources such as technology, institutions and knowledge to reduce risks and save costs and engage in industrial activities related to existing economic activities (Neffke et al. 2011; Teece, 1980). For another, a cognitive basis exists between enterprises with moderate cognitive distance, which can effectively promote knowledge spillover. If the cognitive distance is too close or far, it is hard to recombine existing technologies and knowledge, which thus hinders the generation of new knowledge (Nooteboom et al. 2007). Moreover, firm spin-offs have been observed to be a direct source of regionally unrelated diversification (Klepper and Sleeper, 2005), especially in the development of emerging industries, where they may be the main mechanism for regional branching under certain circumstances.

Unlike the micro mechanism of industrial relatedness, occupational relatedness holds that both matching and learning theories emphasize individuals rather than firms. Knowledge and information are thought not to spill over between firms, but between employees, who then transfer knowledge to firms (Wixe and Andersson, 2017). On the one hand, human capital mainly refers to knowledge, skills and labour ability concentrated in individual workers, and the main ways to improve human capital are health care, on-the-job training, formal education, etc. (Schultz, 1961). To avoid the waste of human capital and save the opportunity cost of finding new jobs, the individual labour force tends to participate in economic activities requiring similar skills as existing ones, which thus matches individual labour skills with the production activities of firms (Penrose, 2009). On the other hand, cognitive proximity provides theoretical support for individual-scale organizational learning as well. Labour individuals with moderate cognitive distance have similar, complementary and shared human capital capabilities (Chen et al. 2024; Frenken et al. 2007). Through channels like collaboration and labour mobility, organizational learning is promoted, and knowledge spillover is generated, which in turn affects the evolution of skill structure and regional economic development (Mawdsley and Somaya, 2016). If the cognitive distance between labour individuals is too far or close, it is not easy to produce knowledge spillover. It should be added that knowledge spillover at the individual level does not necessarily contribute to firm development since the micro-subject of skill relatedness is the individual labour force. Only when new knowledge is applied to the production activities of firms can it be reflected in the economic performance of firms (Cappelli et al. 2019). In addition, spin-offs are another direct source of skill relatedness that drives regional branching. Spin-offs occur when employees leave the workplace to establish new firms based on the skills they acquired at the parent firm (Tanner, 2014).

Essentially, emerging multidimensional cross-relatedness represents a specific application of multi-layer complex networks within the field of network science. With the advancement of network science, traditional single-layer complex networks still exhibit inevitable limitations, as they encompass only one type of object and connection and overlook the heterogeneity inherent in complex systems. This limitation has prevented single-layer complex networks from being widely applied in real-world complex systems. In recent years, building on the foundational theory of single-layer networks, researchers have extensively explored multi-layer networks that incorporate both inter-layer and intra-layer interactions (Artime et al. 2024). Some scholars have examined the impact of scientific research and industrial activities on the evolution of regional patent technologies (Catalán et al. 2022; Balland and Boschma, 2022), while others, based on the concept of cross-relatedness, have investigated the internal interactions among various components of patent technologies (Qin et al. 2024; Chun and Hwang, 2025).

At the same time, cross-relatedness also offers new perspectives for understanding the co-evolution of regional industrial and occupational structures. Industrial relatedness generally refers to the connections between different products or industries, without tracing the source of such relatedness. Occupational relatedness, by contrast, emphasizes the relatedness at the level of individual labour skills. Chen et al. (2025) attempted to construct an industry-occupation cross-relatedness space to decompose the source of industrial relatedness into occupational skill-level origins, and to determine whether regional industrial evolution paths are based on skill-based path dependence or leapfrogging path breakthroughs through industry-occupation connections. Building on this, the bidirectional interaction between industries and occupations has attracted increasing attention. For instance, Deegan et al. (2024) used micro-level individual data to directly measure the interaction between industries and occupations, finding that industrial composition may influence occupational evolution and vice versa. Zhou et al. (2025) and Shutters (2024) further emphasized that industries and occupations represent different manifestations of economic activity and distinguished between self-relatedness (i.e., industrial or occupational relatedness) and cross-relatedness between industries and occupations.

Skill structure and industrial path dynamics. Unlike the micro-foundations underlying industry or occupation relatedness, industry-occupation cross-relatedness is essentially the local interaction between individual workers and firms' industrial activities. Firms and labour individuals serve as the micro-agents of this interaction, which is reflected in the skill matching effect between workers and firm production activities. Based on this, we propose a conceptual framework to depict the relationships among occupation-industry cross-relatedness, heterogeneous labour force spatial agglomeration and industrial evolution (Fig. 1). According to the theory of spatial self-selection in labour economics (Guasch and Weiss, 1981), labour individuals normally choose firms matching their skills for work, and firms are likely to choose labour meeting the skills needs of their jobs (Behrens et al. 2014; Venables, 2011). Firms, a micro-agent of industrial evolution, tend to choose labour matching the skills required for their production activities, and labour per se is inclined to choose work matching its skills to avoid wasting human capital. From the angle of collective learning, the recombination of labour and industrial knowledge may result in

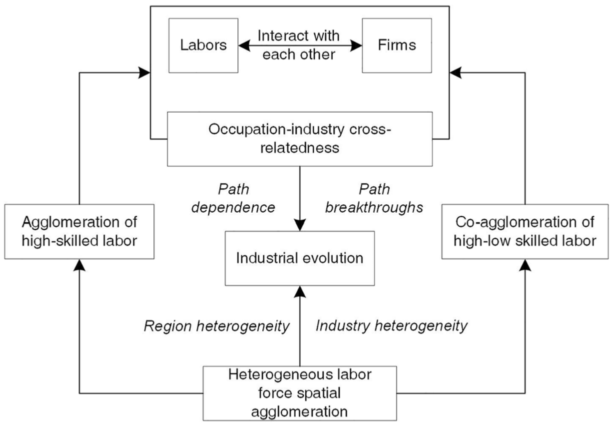

Fig. 1 Conceptual framework. This figure illustrates the conceptual framework used in the study, highlighting the relation between occupation-industry cross-relatedness, heterogeneous labour force spatial agglomeration and industrial evolution. It emphasizes the impact of high-skilled labour agglomeration and the co-agglomeration of high-low skilled labour on occupation-industry cross-relatedness, which in turn influence industrial evolution.

knowledge spillover, which thereby increases productivity (Laursen et al. 2020).

In evolutionary economic geography, the diversity of regional capabilities has long been seen as a key factor influencing path development and transformation. The notion of related variety promotes incremental innovation through shared cognitive structures, while unrelated variety facilitates radical innovation by enabling novel combinations across distant knowledge bases (Frenken et al. 2007; Boschma, 2017). Building on this, emerging literature emphasizes cross-dimensional relatedness—such as industry–occupation or science–technology linkages—as crucial mechanisms of regional renewal (Balland and Boschma, 2022; Catalán et al. 2022). By adopting this cross-relatedness perspective, our study connects local skill structures to industrial path creation, offering a nuanced understanding of regional diversification. Thus, the following hypothesis was proposed:

Hypothesis 1: Regions are likely to develop new industries closely related to local skill structures.

Proponents of development economics argue that labour as a key input factor exerts a huge impact on industrial structure transformation and upgrading (Teixeira and Queirós, 2016). The impact of high-skilled labour on industrial structure upgrading primarily promotes production efficiency by improving skill levels and thus becomes the basis for supporting industrial innovation-driven development (Goos et al. 2014). Especially when labour skills and regional industries are reasonably allocated, the effect of promoting industrial structure upgrading will be significantly enhanced (Acemoglu and Autor, 2011). Studies have shown that the agglomeration of high-skilled labour not only facilitates the transformation and upgrading of existing industries but also equips regions with greater adaptability and absorptive capacity for new knowledge (Reiman et al. 2023), thereby increasing their potential to develop emerging industries. In fact, industrial evolution involves both path-dependent upgrading and breakthroughs, as well as diversification beyond existing trajectories (Neffke et al. 2011; Wu et al. 2023). High-skilled labour, with its capacity for cross-domain collaboration and technological transfer (Kapur, 2001; Kofler et al. 2020), plays a crucial role in driving path breakthroughs (Zhu et al. 2017).

In parallel, the creative class literature in urban geography has highlighted the crucial role of high-skilled, knowledge-intensive labour in shaping urban innovation and economic upgrading (Florida, 2003; Shutters et al. 2016). Creative workers contribute to regional growth not only through their technical expertise but also by fostering institutional openness, cultural vibrancy, and collaborative learning environments (Sunley et al. 2020). These urban conditions enhance a region's ability to attract diverse talent and recombine knowledge, thereby increasing its potential for industrial transformation. This perspective complements the evolutionary economic geography framework by emphasizing the agency of high-skilled individuals in enabling regional diversification and innovation. Therefore, the following hypothesis was proposed:

Hypothesis 2: Regions with high-skilled labour agglomeration are more capable of developing new industries less related to their existing industrial structures.

Based on the spatial self-selection theory of heterogeneous labour location selection, the influx of labour in large cities is the result of "self-selection". The purpose of labour migration between cities is to maximize utility (Glaeser and Resseger, 2010; Roback, 1982). That is, assuming that labour can move freely between cities, high-skilled labour will choose labour markets with high-skilled returns and high inequality, while low-skilled labour will select labour markets with low-skilled returns and low inequality. Meanwhile, large cities also have selection effects on labour, and a certain degree of classification effect is formed by screening out low-skilled labour.

However, will a city be better when the proportion of high-skilled labour in the city is higher and the proportion of low-skilled labour is lower? Many scholars have given a negative answer. They point out the complementarity between high- and low-skilled labour, with their co-agglomeration contributing to the overall increase in urban productivity (Moretti, 2010; Eeckhout et al. 2014). On one hand, low-skilled labour can improve their skills through the knowledge spillovers from high-skilled labour (Carlsen et al. 2016; Glaeser, 1999); on the other hand, the productivity of high-skilled labour is also dependent on the collaborative division of low-skilled support roles, such as in basic services (Goos and Manning, 2007). More importantly, the complementarity of skills not only enhances the internal production efficiency of a region but also lays the foundation for industrial path breakthroughs. In the development of complex industries and emerging technologies, overcoming path dependency often requires the reorganization of cross-disciplinary knowledge and the collaboration of diverse skills (Frenken et al. 2007; Chen and Wu, 2025). This process depends not only on the creative capabilities of high-skilled labour but also on the coordination of basic roles in institutional support and production services (Moretti, 2010; Eeckhout et al. 2014). Therefore, skill complementarity may not only function as an "efficiency tool" but could also act as an enabling mechanism that potentially increases the likelihood of path breakthroughs. Ideas developed in this part lead to the following hypothesis:

Hypothesis 3: Regions with high-low skilled labour co-agglomeration are more capable of developing new industries less related to their existing industrial structures.

The demand for labour varies from industry to industry. Labour-intensive industries are more concerned about labour costs and tend to search for cheap labour to reduce production costs, while technology- and capital-intensive industries attach more importance to human capital. In the meantime, industry types are further differentiated in space. With the transfer and upgrading of industries across regions, the industrial structure in eastern China has been upgraded from low- and middle-end manufacturing to high-end manufacturing and service industries, which further strengthens the demand for high-skilled labour. Industries in central and western China gradually undertake the transferred industries in eastern China to complete the upgrading, and the demand for the labour force with different skill levels is more urgent. In addition, the emerging economic and industrial complexity of EEG is more advantageous compared with that of the traditional static regional division based on location factors and industrial categories considering heterogeneity. The indicators of industrial and economic complexity change dynamically over time, which can precisely reflect the dynamic differences in the capacity of different regions and industries (Hidalgo and Hausmann, 2009; Li et al. 2024) to analyse the root causes of differences in economic performance. Differences in production capacity in different regions cause differences in regional economic output. To better capture the spatial and sectoral variation in industrial evolution, we conduct heterogeneity analyses based on two dimensions: (1) the eastern-central-western division, which reflects distinct regional development trajectories and labour market conditions (Cai et al. 2002; Dai et al. 2022); and (2) industry types categorized by factor intensity, which capture differing skill demands (Han et al. 2022). These dimensions are closely related to the mechanisms of skill endowment and complementarity at the core of our study. Therefore, the following hypothesis was proposed:

Hypothesis 4: The impact of labour skill structure on regional industrial evolution has the heterogeneity of industry and region.

Data and methods

Data sources. This study relied on three types of data for analysis. To begin with, export products were used to calculate industry comparative advantages and cross-relatedness, and corresponding data were derived from the *Chinese Customs Trade Statistics* (CCTS) in 2000, 2005, 2010 and 2015. This dataset is widely used in the research on industrial relatedness in China (Guo and He, 2017; Li et al. 2024). The data preprocessing process is as follows. Firstly, the six-digit product codes of different versions were unified into the 2007 version of the code. Secondly, the unified six-digit product codes were manually matched with the four-digit industry codes of the National Economic Industrial Classification (GB/T 4754-2017). Thirdly, trade firms not engaged in actual production activities were excluded since not all industries export products (Boschma et al. 2013). Finally, the total export value of prefecture-level cities was added, and the processed data contained 601 industrial sub-categories.

Then, occupations were used to indicate the superiority of skill structure over education level (Florida et al. 2012; Sunley et al. 2020). Occupational data were adopted to calculate cross-relatedness and the agglomeration level of skill labour in this study. Corresponding data were obtained from the 2000 Chinese Census, the 2005 Population Sampling Survey, the 2010 Chinese Census and the 2015 Population Sampling Survey. These two datasets have seen wide applications in the study of the evolution of skill structure in China (Shen and Liu, 2016). The data preprocessing process is as follows. The classification standards adopted in the Chinese Census and Population Sampling Survey are basically in line with the International Standard Classification of Occupations (ISCO-08). First, the national standards for occupational classification used in two datasets, including GB/T6565-1999 (for the 2000 Chinese Census) and GB/T6565-2015 (for the 2015 Population Sampling Survey), were used following the research of Zhou et al. (2023). Additionally, the data of sub-major and secondary groups were manually matched according to occupational codes to form 13 occupational categories and 63 occupational subcategories. Second, the occupational data were aggregated to prefecture-level cities; occupational data of prefecture-level cities were matched with industrial data; the occupational and industrial data of 264 prefecture-level cities and 601 industrial sub-categories were obtained. Third, the occupations were grouped into three skill levels—high, medium, and low—based on the widely recognized and established task-based classification approach in labour economics (Autor et al. 2006; Acemoglu and Autor, 2011). This widely adopted framework categorizes occupations according to the type of tasks they primarily involve: abstract tasks (typically requiring analytical or cognitive problem-solving) are associated with high-skilled occupations; routine tasks (which follow set procedures and are susceptible to automation) correspond to medium-skilled occupations; and manual tasks (involving physical activity and relatively limited cognitive demand) are linked to low-skilled occupations. Building on this well-established framework and following Zhou et al. (2023), occupational subgroups were reassigned into skill levels according to their dominant task content and skill requirements. The assignment of major occupational categories to the three skill levels is summarized in Supplementary Table A1.

In addition to occupational and industrial data, some other socioeconomic variable data were derived from statistical yearbooks, including *the China Urban Statistical Yearbook*, *the China Regional Economic Statistical Yearbook*, and provincial and municipal statistical yearbooks.

Measuring relatedness, cross-relatedness and complexity. The measure of industrial relatedness has been widely used in the literature on regional economic structure. First, the location quotient of industry i in region c at time t is calculated by the share of the number of industrial products in region c at time t divided by the proportion of the number of national products in the industry in the total number of national products:

$$RCA_{i,c,t} = \frac{Exp_{i,c,t}/\sum_i Exp_{i,c,t}}{\sum_c Exp_{i,c,t}/\sum_{c,i} Exp_{i,c,t}} \tag{1}$$

$$x_{i,c,t} = \begin{cases} 1, & RCA_{i,c,t} \geq 1 \\ 0, & RCA_{i,c,t} < 1 \end{cases} \tag{2}$$

As suggested by Hidalgo et al. (2007), a region has a comparative advantage in an industry if its location quotient is above 1. Then, $x_{i,c,t}$ is assigned 1 and otherwise 0. On this basis, a city-industry binary matrix M_{ci} is established according to the comparative advantage. The region used for analysis here refers to one of the 264 prefecture-level cities.

Second, industrial proximity is measured by the minimum co-occurrence probability method (Hidalgo et al. 2007). This index computes the conditional probability that a region has a comparative advantage in industries i and j, and takes the minimum value. The specific calculation formula is:

$$\varphi_{i,j,t} = \min\left\{P\left(x_{i,t}|x_{j,t}\right), P\left(x_{j,t}|x_{i,t}\right)\right\} \tag{3}$$

The larger the value is, the closer the two industries will be. The value ranges between 0 and 1, and trending towards 1 indicates that the higher the proximity of the two industries is, the higher the similarity will be, and vice versa.

Third, relatedness density is an established measure of the relationship between economic activities and regional portfolios (Hidalgo, 2021). If an industry has high relatedness density, more industries with comparative advantages are around the industry. The specific calculation formula is:

$$Density_{i,c,t} = \frac{\sum_j \phi_{i,j,t} x_{i,c,t}}{\sum_j \phi_{i,j,t}} \tag{4}$$

In addition to relatedness, two steps are taken to measure occupation-industry cross-relatedness density. This study followed Pugliese et al. (2019), Catalán et al. (2022), and Castaldi and Drivas (2023), and considered industry-occupation cross-relatedness between industry i and occupation m if they both have comparative advantages in the same region[1].

First, occupation-industry proximity is calculated. Below is the specific formula:

$$\varphi^X_{i,m,t} = \min\left\{P\left(x_{i,t}|x_{m,t}\right), P\left(x_{m,t}|x_{i,t}\right)\right\} \tag{5}$$

The value range is between 0 and 1, and trending towards 1 indicates that the higher the proximity of industry i and occupation m is, the higher the similarity will be, and vice versa.

Second, occupation-industry cross-relatedness density is calculated. Similar to industrial relatedness and occupational relatedness density, the cross-relatedness density here is the degree of relatedness between industry and occupation with a comparative advantage. The specific formula is as follows:

$$CrossDensity_{i,c,t} = \frac{\sum_j \phi^X_{i,m,t} x_{m,c,t}}{\sum_j \phi^X_{i,m,t}} \tag{6}$$

$CrossDensity_{i,c,t}$ corresponds to the occupation-industry cross-relatedness of a potential new industry i in relation to the occupational structure of region c at time t. The value ranges from 0 to 1, and being closer to 1 indicates a high degree of relatedness between industry i and occupational skill structure in region c.

In this study, economic and industrial complexity were used to analyse heterogeneity. This paper drew on the method of reflection proposed by Hidalgo and Hausmann (2009). According to the city-industry comparative advantage matrix M_{ci} constructed by Eqs. (1) and (2), the complexity is calculated by iteratively calculating the average value of the attributes of node neighbours at the upper level. This approach is well-recognized in the literature and has been widely used in recent studies (Balland et al. 2019; Hane-Weijman et al. 2022; Li and Rigby, 2023; Qiao et al. 2024). Meanwhile, comparative work also shows that the principal complexity indices—ECI, ECI+, and Fitness—are highly correlated (Cristelli et al., 2013; Fritz & Manduca, 2021). As such, economic and industrial complexity are calculated as follows:

$$\begin{aligned} ECI_c &= \frac{1}{k_c}\sum_i M_{ci} ICI_i \\ ICI_i &= \frac{1}{k_i}\sum_i M_{ci} ECI_c \end{aligned} \tag{7}$$

$$\begin{aligned} k_c &= \sum_i M_{ci} \\ k_i &= \sum_c M_{ci} \end{aligned} \tag{8}$$

where k_c represents the export diversity of the city; k_i stands for industry ubiquity; ECI_c denotes the economic complexity of the city c; ICI_i refers to the complexity of industry i.

Measuring labour agglomeration levels. The main variables of interest in this study also included labour agglomeration and co-agglomeration levels. In this study, measurements were conducted from the perspective of skilled labour. This study built on the work of industrial agglomeration by Ellison et al. (2010), and attempted to use the relative difference of the agglomeration index to construct a measure of agglomeration of high- and low-skilled labour.

First, the agglomeration levels of the two are considered by the location entropy index. The specific calculation formulas are as follows:

$$H = \frac{\sum_{m=11}^{13} T_{m,c}/\sum_m T_{m,c}}{\sum_{m=11}^{13}\sum_c T_{m,c}/\sum_m\sum_c T_{m,c}} \tag{9}$$

$$L = \frac{\sum_{m=1}^{4} T_{m,c}/\sum_m T_{m,c}}{\sum_{m=1}^{4}\sum_c T_{m,c}/\sum_m\sum_c T_{m,c}} \tag{10}$$

where H and L represent the agglomeration level of high- and low-skilled labour, respectively; m stands for 13 categories of skilled labour ($m = 1, 2, 3\ldots 13$, according to the previous division; 1–4 and 11–13 categories are low- and high-skilled labour, respectively); $T_{m,c}$ denotes the corresponding total number of occupations m in region c. The larger the index is, the higher the agglomeration level of high- or low-skilled labour relative to the whole country will be, and the more concentrated the geospatial distribution will be.

Second, the agglomeration index difference is used to construct the index of high-low skilled labour co-agglomeration:

$$Co\_HL = 1 - |H - L|/|H + L| + |H - L| \tag{11}$$

The higher the value of this index is, the higher the level of high-low skilled labour co-agglomeration will be, and the higher the degree of skill complementarity of the labour force will be.

Econometric model. According to Balland et al. (2019), Castaldi and Drivas (2023), and Catalán et al. (2022), the econometric model in this study is considered as follows:

$$\begin{aligned} x_{i,c,t} = {} & \alpha_0 + \alpha_1 x_{i,c,t-1} + \alpha_2 CrossDensity_{i,c,t-1} \\ & + \delta X + \varphi_c + \varphi_i + \varphi_t + \varepsilon_{c,i,t-1} \end{aligned} \tag{12}$$

where i represents industry category; t stands for time; t-1 is the previous period time; $x_{i,c,t}$ denotes whether the city industry category has a comparative advantage at time t. According to the way of model setting by Boschma (2017) and Zhu et al. (2017), $x_{i,c,t-1}$ is added. $CrossDensity_{i,c,t-1}$ denotes occupation-industry cross-relatedness density and X is a set of control variables. In addition, φ_i refers to industry fixed effects; φ_c means city fixed effects; φ_t means time fixed effects; $\varepsilon_{i,c,t-1}$ is regression residuals.

This paper introduced the interaction term variables of heterogeneous labour skill structure variables (skilled labour agglomeration and high-low skilled labour co-agglomeration) and industry-occupation cross-relatedness density. This approach follows Boschma and Capone (2015) and Zhu et al. (2017). It also enables us to test how skill endowment and complementarity drive path-breaking industrial evolution; accordingly, we specify the following model:

$$\begin{aligned} x_{i,c,t} = {} & \alpha_0 + \alpha_1 x_{i,c,t-1} + \alpha_2 CrossDensity_{i,c,t-1} \\ & + \beta_1 H_{c,t-1} + \beta_2 Co\_HL_{c,t-1} \\ & + \gamma_1 (H_{c,t-1} \times CrossDensity_{i,c,t-1}) \\ & + \gamma_2 (Co\_HL_{c,t-1} \times CrossDensity_{i,c,t-1}) \\ & + \delta X + \varphi_c + \varphi_i + \varphi_t + \varepsilon_{c,i,t-1} \end{aligned} \tag{13}$$

where $CrossDensity_{i,c,t-1}$ represents industry-occupation cross-relatedness density; $H_{c,t-1}$ stand for the agglomeration level of high-skilled labour; $Co\_HL_{c,t-1}$ denotes the co-agglomeration level of high-low skilled labour; X, φ_c, φ_t, φ_i and $\varepsilon_{c,i,t-1}$ have the same meaning as that of the previous model.

Six variables were selected as control variables. (1) Five regional-level variables. The population controls the city size. The total gross domestic product (GDP) controls the level of urban economic development. The degree of foreign-trade dependence, measured as the ratio of total imports and exports to GDP, characterizes the regional globalization index. The marketization index, also known as the National Economic Research Institute (NERI) index, reflects the degree of market-oriented economic transformation across China and is used to

Table 1 Descriptive statistics.

Variables	Observation	Mean	Std. Dev.	Min	Max
$x_{i,c,t}$	475992	0.277	0.447	0.000	1.000
$x_{i,c,t-1}$	475992	0.242	0.428	0.000	1.000
$CrossDensity_{i,c,t-1}$	475992	0.337	0.179	0.000	0.907
$H_{c,t-1}$	475992	0.982	0.484	0.209	4.172
$Co\_HL_{c,t-1}$	475992	1.257	0.261	0.948	3.863
Population	475992	425.038	300.581	27.630	3303.450
GDP	475992	8.867e + 06	1.433e + 07	270594	1.717e + 08
FTD	475992	0.909	5.540	0.001	143.027
Marketization	475992	6.662	2.603	1.217	13.420
Industry stock	475992	145.408	75.629	3.000	394.000
Industry size	475992	3.185e + 10	7.235e + 10	41430	7.275e + 11

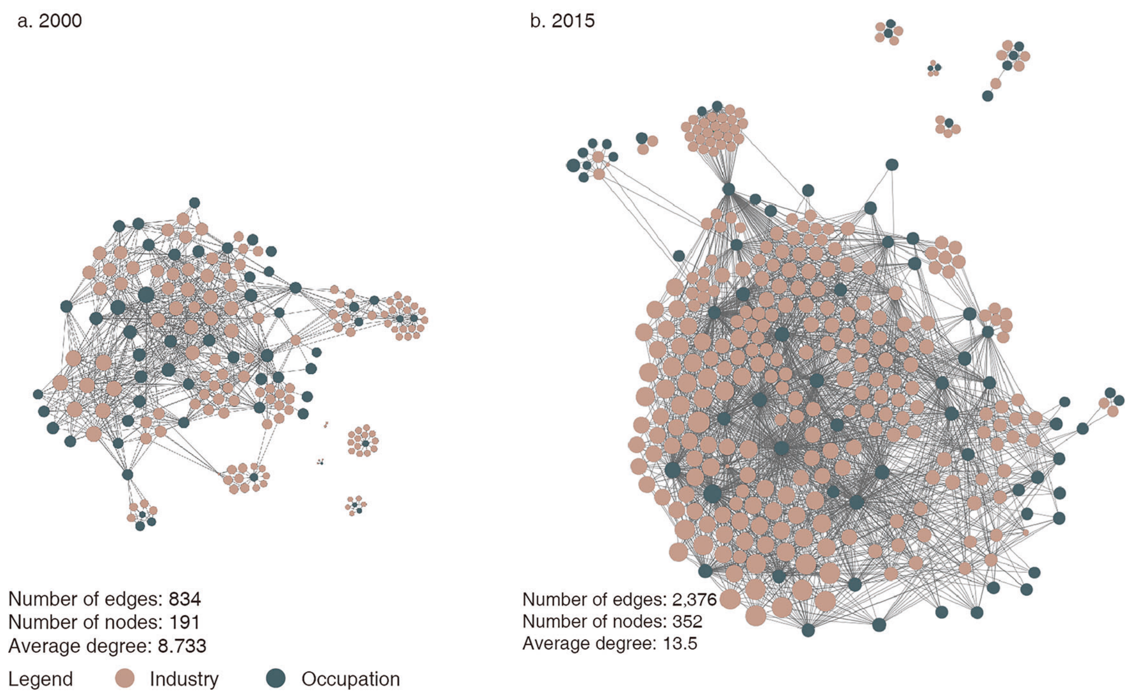

Fig. 2 Industry-occupation cross-space. This figure presents the changes in the industry-occupation cross-space from 2000 to 2015. The red nodes represent industries, and the green nodes represent occupations. This figure highlights the increased complexity and interconnections between industries and occupations over the 15-year period.

examine institutional environment differences (Fan et al. 2019). In this study, the marketization index was calculated at the municipal level. Additionally, industry stock (the total number of specialized industries in a region) is included as a control variable. (2) At the industry level, industry size (the total number of regions specializing in industry *i*) is included. All control variables were taken as their natural logarithms.

Concerning the large number of fixed effects in the model, a linear probability model (LPM) was selected for a regression analysis (Boschma et al. 2013; Castaldi and Drivas, 2023). Descriptive statistics are illustrated in Table 1. The correlation matrices for the variables are summarized in Supplementary Table A2 online. No obvious sign of multicollinearity can be found, as all of the independent variables have a variance inflation factor (VIF) value of below 4.5.

Results

Descriptive results. The evolution of the industry-occupation cross-space in China from 2000 to 2015 is shown in Fig. 2. To make the visualization of the cross-space clearer, only edges with weights greater than 0.45 were included. Node size represents the specialization degree of the industry or occupation. By comparing the two subgraphs of Fig. 2, it can be seen that the number of nodes and edges all exhibited a significant increase. This indicates that the relatedness of industries and occupations in China became closer during the study period.

The spatial distribution of the agglomeration level of high-skilled labour is presented in Fig. 3. According to the algorithm, the larger index indicates the higher agglomeration level of high-skilled labour in the city and the more concentrated geospatial distribution. As can be seen, the index presents an overall downward trend, which indicates that more cities are attractive to high-skilled labour nationwide. High-skilled labour force is mainly distributed in the eastern cities and provincial capitals, probably because these cities have high industrial complexity and more developed economies. The spatial distribution of high-low skilled labour co-agglomeration is demonstrated in Fig. 4. During the study period, the co-agglomeration levels in various cities in China changed greatly. According to the algorithm, the skill complementarity degree of the labour force will be higher when the index is larger. Except for less economically developed cities

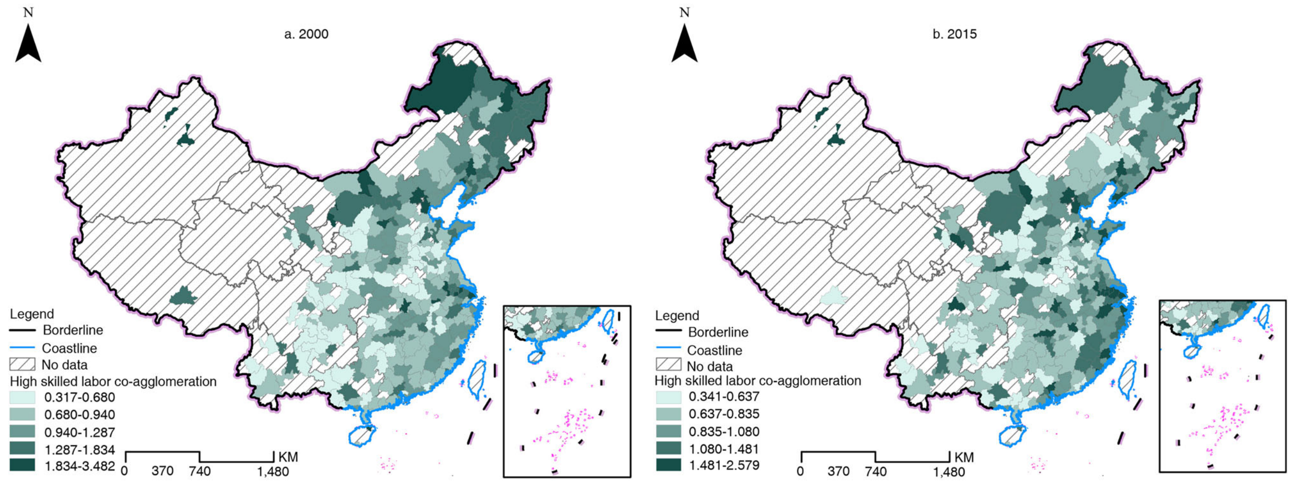

Fig. 3 The spatial distributions of high-skilled labour agglomeration. This figure illustrates the spatial distributions of high-skilled labour agglomeration across various regions. The figure emphasizes more cities have become attractive to high-skilled labour, with the eastern coastal areas showing a distinct agglomeration advantage for high-skilled labour.

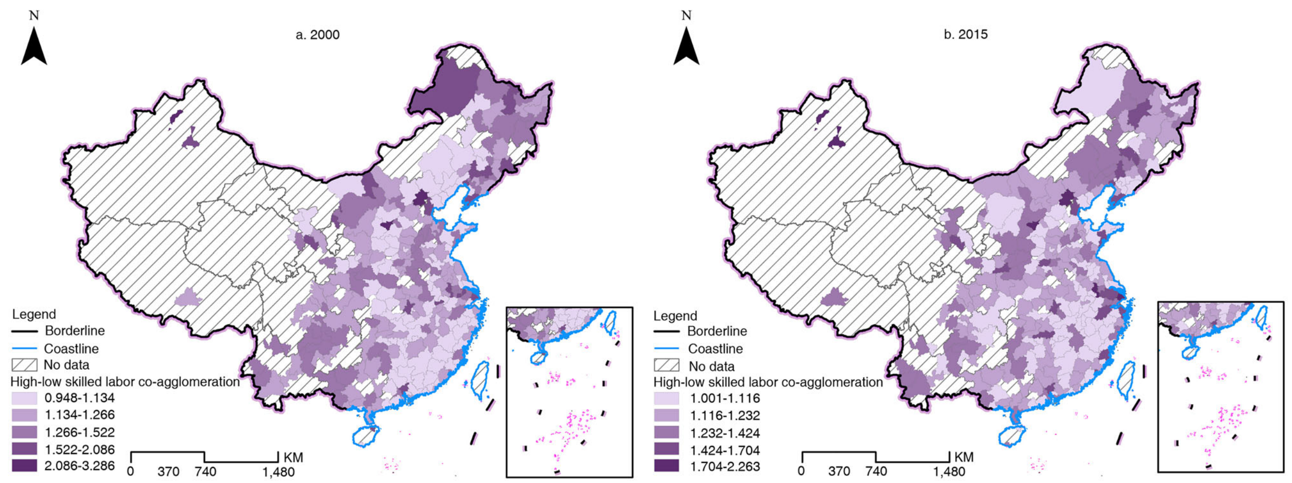

Fig. 4 The spatial distributions of high-low skilled labour co-agglomeration. This figure illustrates the spatial distributions of high-low skilled labour co-agglomeration across various regions. This figure illustrates the spatial dispersion of skill complementarity levels, indicating that there is no significant agglomeration advantage in the eastern coastal regions.

such as Beijing and Shanghai, other cities with high-skilled labour force do not have high skill complementarity effect. Namely, the concentration of high-skilled labour force and the level of skill complementarity effect do not match in space.

Regression analysis. The baseline regression results are shown in Table 2. In Models 1–5, the industrial comparative advantage index and cross-relatedness density in the previous stage were both significant and positive, which revealed a significant path-dependent effect in the development of city industries in China. This is consistent with existing research results (Guo and He, 2017; He et al. 2018; Chen et al. 2025). In addition, regional industrial diversification is related to the skill structure of the local labour force, which verifies Hypothesis 1. In Models 4 and 5, the coefficients on both high-skilled labour agglomeration and high–low skill co-agglomeration are significantly positive, confirming that richer skill endowments and stronger skill complementarity each foster industrial development. Model 6 introduces their interaction with cross-relatedness density. The interaction coefficients are significantly negative, indicating a substitutive relationship: the positive effect of skill agglomeration is strongest where the new industry shares fewer capabilities with the region's incumbent industries. In other words, abundant and complementary skills compensate for limited relatedness, enabling the emergence of industries that break away from the existing industrial path. This pattern substantiates Hypotheses 2 and 3 by demonstrating that skill endowment and skill complementarity facilitate path-breaking industrial trajectories precisely when the relatedness foundation is weak.

Heterogeneity analysis. A heterogeneity analysis was conducted from regional and industrial levels[2], and the results are illustrated in Tables 3 and 4. These two dimensions of heterogeneity are closely aligned with the theoretical framework of the study. The eastern-central-western division captures spatial variation in development stages, institutional contexts, and labour market configurations (Cai et al. 2002; Dai et al. 2022), all of which condition how skill endowment and complementarity influence industrial evolution in Chinese regional context. Meanwhile, classifying industries by factor intensity reflects underlying technological and organizational characteristics that determine skill demands and knowledge absorption capacities (Han et al. 2022). Examining heterogeneity along these lines thus enables us

Table 2 Baseline regressions.

	(1)	(2)	(3)	(4)	(5)
$x_{i,c,t-1}$	0.479^{***} (0.002)		0.473^{***} (0.002)	0.472^{***} (0.002)	0.472^{***} (0.002)
$CrossDensity_{i,c,t-1}$		0.737^{***} (0.011)	0.336^{***} (0.009)	0.389^{***} (0.010)	0.517^{***} (0.028)
$H_{c,t-1}$				0.053^{***} (0.003)	0.078^{***} (0.005)
$Co\_HL_{c,t-1}$				0.069^{***} (0.005)	0.117^{***} (0.007)
$H_{c,t-1}CrossDensity_{i,c,t-1}$					-0.123^{***} (0.019)
$Co\_HL_{c,t-1}CrossDensity_{i,c,t-1}$					-0.205^{***} (0.029)
Population	−0.006 (0.009)	0.004 (0.010)	−0.002 (0.009)	0.002 (0.009)	0.003 (0.009)
GDP	-0.041^{***} (0.005)	-0.064^{***} (0.005)	-0.050^{***} (0.005)	-0.044^{***} (0.005)	-0.047^{***} (0.005)
FTD	-0.033^{***} (0.005)	-0.022^{***} (0.005)	-0.032^{***} (0.005)	-0.025^{***} (0.005)	-0.025^{***} (0.005)
Marketization	-0.013^{***} (0.005)	0.002 (0.005)	-0.010^{**} (0.005)	-0.010^{**} (0.005)	-0.009^{*} (0.005)
Industry stock	-0.060^{***} (0.002)	0.011^{***} (0.002)	-0.061^{***} (0.002)	-0.059^{***} (0.002)	-0.059^{***} (0.002)
Industry size	-0.012^{***} (0.001)	-0.026^{***} (0.002)	-0.012^{***} (0.001)	-0.012^{***} (0.001)	-0.012^{***} (0.001)
Industry FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Constant	1.559^{***} (0.112)	1.485^{***} (0.121)	1.451^{***} (0.112)	1.240^{***} (0.113)	1.265^{***} (0.113)
N	475992	475992	475992	475992	475992
R^2	0.340	0.172	0.341	0.342	0.342
Adjusted R^2	0.338	0.171	0.340	0.341	0.341

Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

to test the robustness and contextual dependence of our core findings.

The specific analysis steps are as follows: First, at the regional level, prefecture-level cities were grouped for regression to explore the spatial heterogeneity of labour force spatial agglomeration on regional industrial evolution, given the impact of economic development level and location, and according to the GDP level and locations (eastern, central and western China). The results are shown in Models 1–3 in Table 3. Second, at the industrial level, this study referred to the report on global industrial development for 2013 by the United Nations Industrial Development Organization. According to the method of "technology group as the standard manufacturing classification" and combined with the actual situation of China, 29 two-digit manufacturing industries were classified into labour-, capital- and technology-intensive categories for grouping regression (see Supplementary Table A3 online). Among them, C13-24 and C41-43, C25-26 and C28-33, as well as C27 and C34-40 are labour-, capital- and technology-intensive industries, respectively. The results are presented in Models 1–3 in Table 4. Third, according to the three-digit economic and industrial complexity, regions and industries were grouped into three categories: high, medium and low. The results are shown in Models 4–6 in Tables 3 and 4.

The coefficients of $x_{i,c,t-1}$, cross-relatedness density and interaction variables are shown in Tables 3 and 4. The city industrial development of China has significant path-dependent effects at different regional and industrial levels. Regional industrial diversification is correlated with the skill structure of the local labour force. Labour skill endowment and skill complementation have a positive impact on industrial path breakthroughs to varying degrees.

The regional heterogeneity of the results is shown in Table 3. It can be seen from Models 1–3 that interaction terms are significantly negative. The impact of high-skilled labour on industrial path breakthroughs is more prominent in central China and western China. Skill complementary effect has a greater impact on eastern China. In other words, in the eastern China, where there is a large number of high-skilled labourers, the skill complementary effect is more likely to promote regional transformation and upgrading, while in regions with relatively few high-skilled labourers, the direct impact of the introduction of high-skilled labour is more conducive to the creation of new industries. It can be seen from Models 4–6 that industrial path breakthroughs in regions with higher economic complexity are more related to the skill complementarity effect, and regions with medium and lower economic complexity need to introduce high-skilled labour to achieve path breakthroughs. To be more specific, high-complexity cities have stronger production and learning capabilities, and urban vibrancy is beneficial to enhancing knowledge spillover (Balland et al. 2020). Due to labour preferences in cities where labour demand matches their skill portfolios (Ewers and Dicce, 2018), high-skilled labour is more concentrated in high-complexity cities. As a result, it is more affected by high-low skilled complementary effect when industrial path breakthroughs are promoted. To put it another way, it is more important to improve the optimal allocation of skilled labour. In low-complexity cities with poor production and learning capabilities, high-skilled labour accounts for a relatively low proportion, and introducing high-skilled labour will bring new knowledge and technology to these regions. If cities have the ability to use these possibilities, it is possible to promote industrial path breakthroughs and improve regional innovation performance (Li and Rigby, 2023). Cities of different economic complexity are in eastern, central and western China. The results of Models 4–6 allow a better explanation of regional heterogeneity. This extends current understandings of how labour structures interact with existing industrial bases to shape regional innovation trajectories.

In the aspect of industrial heterogeneity (Table 4), it can be seen from Models 1–3 that the positive correlation between labour-intensive industrial diversification and skill structure is stronger, and labour skill endowment and skill complementation play the strongest role in promoting their path breakthroughs. Capital-intensive industry path breakthroughs are more affected by high-skilled labour, and technology-intensive industry path breakthroughs are more dependent on the skill complementation effect. As shown in Models 4–6, the regional path dependence effect is stronger with higher industrial complexity. The lower the industrial complexity of a region is, the more its industrial diversification will be related to skill structure, and the more necessary it will be to introduce high-

Table 3 Heterogeneity analysis at the regional level.

	(1) Eastern China	(2) Central China	(3) Western China	(4) High economic complexity	(5) Medium economic complexity	(6) Low economic complexity
$x_{i,c,t-1}$	0.536^{***} (0.002)	0.434^{***} (0.003)	0.379^{***} (0.004)	0.501^{***} (0.004)	0.477^{***} (0.003)	0.434^{***} (0.003)
$CrossDensity_{i,c,t-1}$	1.043^{***} (0.048)	0.323^{***} (0.052)	0.219^{***} (0.057)	1.391^{***} (0.105)	0.296^{***} (0.068)	0.582^{***} (0.058)
$H_{c,t-1}$	0.134^{***} (0.013)	0.117^{**} (0.057)	0.119^{***} (0.007)	0.096^{***} (0.026)	0.047^{***} (0.010)	0.068^{***} (0.008)
$Co\_HL_{c,t-1}$	0.236^{***} (0.021)	0.155^{***} (0.015)	0.148^{***} (0.015)	0.108^{**} (0.049)	0.042^{*} (0.025)	0.094^{***} (0.018)
$H_{c,t-1}CrossDensity_{i,c,t-1}$	-0.270^{***} (0.042)	-0.311^{**} (0.132)	-0.391^{***} (0.034)	−0.011 (0.093)	-0.087^{**} (0.042)	-0.059^{*} (0.033)
$Co\_HL_{c,t-1}CrossDensity_{i,c,t-1}$	-0.657^{***} (0.057)	-0.544^{***} (0.051)	-0.494^{***} (0.056)	-0.196^{***} (0.038)	-0.194^{***} (0.071)	-0.160^{***} (0.057)
Population	-0.238^{***} (0.017)	0.027 (0.035)	0.104^{***} (0.012)	-0.168^{***} (0.039)	−0.007 (0.042)	-0.365^{***} (0.038)
GDP	-0.033^{***} (0.012)	−0.006 (0.009)	-0.096^{***} (0.007)	−0.017 (0.046)	-0.142^{***} (0.014)	−0.007 (0.014)
FTD	0.015 (0.012)	0.035^{***} (0.008)	-0.065^{***} (0.008)	−0.073 (0.060)	-0.041^{***} (0.014)	−0.004 (0.010)
Marketization	0.054^{***} (0.009)	-0.098^{***} (0.008)	0.011 (0.009)	0.105^{***} (0.019)	-0.080^{***} (0.010)	-0.109^{***} (0.011)
Industry stock	-0.115^{***} (0.004)	-0.049^{***} (0.003)	-0.032^{***} (0.003)	-0.105^{***} (0.008)	-0.041^{***} (0.003)	-0.098^{***} (0.004)
Industry size	0.011^{***} (0.002)	-0.022^{***} (0.002)	-0.032^{***} (0.003)	0.007^{*} (0.004)	-0.018^{***} (0.002)	-0.010^{***} (0.002)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.258^{***} (0.260)	1.505^{**} (0.234)	1.773^{***} (0.131)	0.778 (0.755)	2.708^{***} (0.391)	3.518^{***} (0.376)
N	174891	174891	126210	75125	187512	213355
R^2	0.384	0.344	0.280	0.364	0.351	0.364
Adjusted R^2	0.382	0.342	0.276	0.358	0.348	0.361

Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4 Heterogeneity analysis at the industry level.

	(1) labor-intensive industries	(2) capital-intensive industries	(3) technology-intensive industries	(4) High industrial complexity	(5) Medium industrial complexity	(6) Low industrial complexity
$x_{i,c,t-1}$	0.455*** (0.003)	0.466*** (0.003)	0.450*** (0.003)	0.497*** (0.003)	0.474*** (0.002)	0.455*** (0.004)
$CrossDensity_{i,c,t-1}$	0.750*** (0.051)	0.608*** (0.056)	0.574*** (0.042)	0.410*** (0.050)	0.566*** (0.041)	0.869*** (0.073)
$H_{c,t-1}$	0.066*** (0.009)	0.051*** (0.010)	0.095*** (0.006)	0.105*** (0.008)	0.068*** (0.007)	0.061*** (0.011)
$Co\_HL_{c,t-1}$	0.127*** (0.017)	0.147*** (0.018)	0.095*** (0.013)	0.116*** (0.016)	0.112*** (0.013)	0.197*** (0.023)
$H_{c,t-1}CrossDensity_{i,c,t-1}$	−0.101*** (0.034)	−0.244* (0.139)	−0.164*** (0.028)	−0.084** (0.035)	−0.111*** (0.027)	−0.305*** (0.043)
$Co\_HL_{c,t-1}CrossDensity_{i,c,t-1}$	−0.152*** (0.053)	−0.211*** (0.058)	−0.290*** (0.043)	−0.243** (0.125)	−0.240*** (0.042)	−0.179*** (0.074)
Population	−0.007 (0.017)	−0.022 (0.019)	0.024* (0.014)	0.005 (0.014)	0.007 (0.014)	0.032 (0.034)
GDP	−0.064*** (0.009)	−0.105*** (0.010)	0.003 (0.007)	−0.047*** (0.008)	−0.046*** (0.007)	−0.018 (0.016)
FTD	−0.036*** (0.009)	−0.055*** (0.010)	0.001 (0.007)	−0.029*** (0.008)	−0.034*** (0.007)	0.029** (0.013)
Marketization	0.012 (0.008)	0.007 (0.009)	−0.030*** (0.007)	−0.018** (0.007)	0.004 (0.007)	0.018 (0.015)
Industry stock	−0.072*** (0.003)	−0.061*** (0.003)	−0.042*** (0.002)	−0.076*** (0.003)	−0.059*** (0.002)	−0.021*** (0.005)
Industry size	0.018*** (0.003)	−0.028*** (0.003)	−0.028*** (0.003)	−0.025*** (0.004)	0.002 (0.002)	−0.019 (0.019)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.958*** (0.198)	2.937*** (0.231)	0.446*** (0.166)	1.730*** (0.197)	0.869*** (0.168)	0.035 (0.489)
N	152064	128304	195624	173712	229680	72600
R^2	0.370	0.351	0.311	0.344	0.335	0.439
Adjusted R^2	0.368	0.349	0.309	0.341	0.333	0.435

Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

skilled labour to achieve path breakthroughs. This is mainly because long-term industrial development in these regions is relatively simple, and the technology spillover benefits generated by the agglomeration of the high-skilled workforce will be greater. The skill complementarity effect promotes the regional path breakthroughs of high-sized industrial complexity most obviously. The reason may be that more low-skilled labour flows into areas with lower industrial complexity, while the demand for low-skilled labour is overlooked as cities with higher industrial complexity continue to introduce high-skilled labour. By incorporating both regional and industrial heterogeneity, the study deepens our understanding of how differentiated skill policies may unlock path-creating opportunities across diverse local contexts.

Robustness check. Several robustness tests of the baseline regression model were carried out. First, the calculation method of industrial data was changed. In this study, three-digit industrial data were used for recalculation according to Castaldi and Drivas (2023). Second, the definition method of regional comparative advantages was changed. The calculation of the original regional comparative advantage was defined according to whether the location quotient was greater than 1. This study was based on the robustness test method of Zhu et al. (2017) and He et al. (2018), and utilized different thresholds (0.8 and 1.2) to define the regional comparative advantage. Finally, we replaced the minimum co-occurrence probability metric with Hidalgo et al.'s (2007) cosine-similarity measure to calculate relatedness proximity. All robustness checks confirm the findings in Tables 2 and 3; full results are provided in Supplementary Tables A4–A7 (online Appendix).

Discussion and conclusions

The study of regional industrial evolution and relatedness in EEG has been extensively discussed (Frenken and Boschma, 2007; Kogler et al. 2023a). However, two gaps still exist. First, while empirical research on EEG mainly examines the role of relatedness and regional development trajectories in the context of industry sector evolution, limited attention has been paid to the cross-relatedness between industrial and occupational structures. Although recent studies have called for integrating industry-occupation similarity networks (Chen et al. 2025), this direction remains underexplored. Building on the concept of multidimensional relatedness (Balland and Boschma, 2022; Castaldi and Drivas, 2023; Catalán et al. 2022), this study extends the literature by constructing an industry-occupation cross-relatedness network, which enriches the understanding of labour skill influences beyond traditional industrial relatedness. Second, although the role of skill structures in shaping industrial development paths is well established (Neffke and Henning, 2013), existing research has not sufficiently examined how different forms of skill endowment and complementarity jointly influence regional path breakthroughs. This paper contributes to the empirical literature by offering a refined analytical perspective on how heterogeneous labour forces, especially the interaction between high- and low-skilled workers, shape path-breaking trajectories in regional industrial transformation.

The main research results of industrial and occupational data analytics in Chinese prefecture-level cities between 2000 and 2015 show that the industrial development of China has a significant path-dependent feature, which is aligned with existing research results (Muneepeerakul et al. 2013; Neffke et al. 2011). The relatedness between industries and occupations in China has become closer, and the skill structure is closely bound up with regional industrial diversification. Both high-skilled labour agglomeration and high-low skilled labour co-agglomeration can help regions break industry-occupation cross-relatedness and achieve industrial path breakthroughs. The heterogeneity analysis and robustness test results strengthen the above conclusions. First, it was found in this study that the impact of high-low skilled labour agglomeration on industrial path breakthroughs is more pronounced in eastern China. Second, industrial path breakthroughs in regions with higher economic complexity are more related to the skill complementarity effect. Third, labour-intensive industrial diversification has the greatest relatedness to skill structure. Finally, the regional path dependence effect is stronger in regions with higher industrial complexity. Regions with lower industrial complexity need to introduce high-skilled labour to achieve path breakthroughs. These findings further validate the hypotheses of this study.

The above results are significant as they not only reveal a positive correlation between industry-occupation cross-relatedness and regional industrial diversification but also highlight an indirect effect of labour, indicating opportunities for regional economic transformation. More precisely, the findings of this study suggest that those regions seeking to create new paths to industrial specialization can achieve their goals by optimizing workforce structure. In specific terms, pre-existing regional industrial structures and industry-occupation cross-relatedness may limit the speed and direction of regional industrial diversification, but the optimal allocation of regional high-skilled labour agglomeration and high-skilled labour co-agglomeration can ease this dependence and provide new development paths. These findings help study how regions can further "jump" beyond the "limits" of industrial relatedness. Moreover, they contribute to enriching the literature on the impact of heterogeneous labour mobility on industrial structure upgrading. This considers the role of the spatial matching effect of skilled labour in the process of industrial structure upgrading of high-skilled labour, and discusses the impact of high-low skilled labour co-agglomeration on industrial structure upgrading.

These empirical results have clear policy implications. First, the research results show that the agglomeration of high-skilled labour and the co-agglomeration of low-skilled labour in a region exert a positive impact on industrial path breakthroughs to varying degrees, which deserves the attention of policymakers. Regional talent policies focus more on introducing high-skilled labour (Yang, 2023). It is important to fully exploit the positive externalities of high-skilled labour, but how to coordinate the co-agglomeration of high-low skilled labour may be another path breakthrough for government policymaking. To achieve this, governments should implement inclusive regional population policies. Expanding employment opportunities for diverse labour groups can help cities attract and retain workers across skill levels in a more balanced and sustainable manner. Second, the heterogeneity in labour demand across cities calls for differentiated and targeted policy responses. In regions where high-skilled labour is relatively abundant, preferential housing policies—such as providing affordable rental housing—may enhance the attraction and retention of low-skilled workers. Conversely, cities with a relative shortage of high-skilled labour should strengthen their talent attraction mechanisms, for example, by offering financial subsidies or settlement incentives to newly recruited professionals. Third, these efforts could be accompanied by improved labour mobility. In China, the restrictions of the household registration system and the existence of urban settlement thresholds are the non-negligible causes of related institutional "distortions" (Gu et al. 2021; Qi et al. 2024). Given the complexity of reforming these systems, particularly in mega-cities facing fiscal pressure, gradual and flexible approaches are recommended.

This research also has the following limitations and may constitute future research directions. First, it revolves around prefecture-level cities. As argued by Fotheringham and Wong (1991) and Castaldi and Drivas (2023), the same analysis for smaller geographical units can reveal important differences. The movement of labour in geographical units at different levels and the comparative analysis of different geographical levels may reveal more information. Second, this research explored the impact of high-low skilled labour complementation on the mechanism of industrial evolution, and future research can further discuss the matching of specific labour types. In addition, selecting typical regions to explore deeper mechanisms using in-depth qualitative case study methods will be interesting. Third, due to changes in the statistical standards of key databases, this study relies on data only up to 2015. Future research should aim to update the dataset and explore alternative data sources to better reflect recent developments in industrial evolution and skill dynamics. In particular, the rapid advancement of artificial intelligence and digital technologies may significantly reshape skill demand, which in turn could further influence the mechanisms of industrial evolution examined in this study (Buyukyazici et al. 2024; Santoalha et al. 2021).

Data availability

The datasets generated and/or analysed during the current study are available on request from the authors.

Received: 15 October 2024; Accepted: 18 August 2025;
Published online: 28 August 2025

Notes

1 Industries and occupations are interconnected. We take a top-down approach to measuring the relatedness between industries and occupations based on co-location patterns among regions. By comparison, several papers like Deegan et al. (2024) and Boschma (2024) took a bottom-up approach to establishing the industry-occupation link based on micro data of individuals. The results of these two approaches might have differences.

2 For brevity, only complete models with fixed effects were included.

References

Acemoglu D, Autor D (2011) Skills, tasks and technologies: implications for employment and earnings. Handb Labor Econ 4:1043–1171. https://doi.org/10.1016/S0169-7218(11)02410-5

Artime O, Grassia M, De Domenico M et al. (2024) Robustness and resilience of complex networks. Nat Rev Phys 6:114–131. https://doi.org/10.1038/s42254-023-00676-y

Autor D, Katz L, Kearney M (2006) The polarization of the US labor market. Am Econ Rev 96:189–194. https://doi.org/10.1257/000282806777212620

Balland P-A, Boschma R (2022) Do scientific capabilities in specific domains matter for technological diversification in European regions? Res Policy 51:104594. https://doi.org/10.1016/j.respol.2022.104594

Balland P-A, Boschma R, Crespo J et al. (2019) Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification. Reg Stud 53:1252–1268. https://doi.org/10.1080/00343404.2018.1437900

Balland P-A, Jara-Figueroa C, Petralia S et al. (2020) Complex economic activities concentrate in large cities. Nat Hum Behav 4:248–254. https://doi.org/10.1038/s41562-019-0803-3

Barbour E, Markusen A (2007) Regional occupational and industrial structure: does one imply the other? Int Reg Sci Rev 30:72–90. https://doi.org/10.1177/0160017606296727

Behrens K, Duranton G, Robert-Nicoud F (2014) Productive cities: sorting, selection, and agglomeration. J Polit Econ 122:507–553. https://doi.org/10.1086/675534

Boschma R (2017) Relatedness as driver of regional diversification: a research agenda. Reg Stud 51:351–364. https://doi.org/10.1080/00343404.2016.1254767

Boschma R, Capone G (2015) Institutions and diversification: related versus unrelated diversification in a varieties of capitalism framework. Res Policy 44:1902–1914. https://doi.org/10.1016/j.respol.2015.06.013

Boschma R, Minondo A, Navarro M (2013) The emergence of new industries at the regional level in Spain: a proximity approach based on product relatedness. Econ Geogr 89:29–51. https://doi.org/10.1111/j.1944-8287.2012.01170.x
Boschma R (2024) An evolutionary approach to regional studies on global value chains. Reg Stud 58:1492–1500. https://doi.org/10.1080/00343404.2024.2323035
Buyukyazici D, Mazzoni L, Riccaboni M et al. (2024) Workplace skills as regional capabilities: relatedness, complexity and industrial diversification of regions. Reg Stud 58:469–489. https://doi.org/10.1080/00343404.2023.2206868
Cai F, Wang D, Du Y (2002) Regional disparity and economic growth in China: the impact of labor market distortions. China Econ Rev 13(2–3):197–212. https://doi.org/10.1016/S1043-951X(02)00072-X
Cappelli R, Boschma R, Weterings A (2019) Labour mobility, skill-relatedness and new plant survival across different development stages of an industry. Environ Plan Econ Space 51:869–890. https://doi.org/10.1177/0308518X18812466
Carlsen F, Rattsø J, Stokke HE (2016) Education, experience, and urban wage premium. Reg Sci Urban Econ 60:39–49. https://doi.org/10.1016/j.regsciurbeco.2016.06.006
Castaldi C, Drivas K (2023) Relatedness, cross-relatedness and regional innovation specializations: an analysis of technology, design, and market activities in Europe and the US. Econ Geogr 99:253–284. https://doi.org/10.1080/00130095.2023.2187374
Catalán P, Navarrete C, Figueroa F (2022) The scientific and technological cross-space: Is technological diversification driven by scientific endogenous capacity? Res Policy 51:104016. https://doi.org/10.1016/j.respol.2020.104016
Chen J, Li X, Zhu Y (2024) Shock absorber and shock diffuser: the multiple roles of industrial diversity in shaping regional economic resilience after the great recession. Ann Reg Sci 72:1015–1045. https://doi.org/10.1007/s00168-023-01233-2
Chen Y, Wu K (2025) Integrating artificial intelligence into regional technological domains: the role of intra- and extra-regional AI relatedness. Camb J Reg Econ Soc 18:111–130. https://doi.org/10.1093/cjres/rsae041
Chen J, Li X, Zhou X et al. (2025) Skill capabilities behind the scenes: the role of occupational portfolio in regional industrial evolution. Reg Stud 59:2416535. https://doi.org/10.1080/00343404.2024.2416535
Chun Y, Hwang J (2025) The nexus of artificial intelligence and green innovation: a cross-density analysis at the country level. J Knowl Econ 16:1688–1716. https://doi.org/10.1007/s13132-024-02076-8
Cristelli M, Gabrielli A, Tacchella A et al. (2013) Measuring the intangibles: a metrics for the economic complexity of countries and products. PloS One 8:e70726. https://doi.org/10.1371/journal.pone.0070726
Dai X, Yan L, Li J et al. (2022) A research on the threshold effect of human capital structure upgrading and industrial structure upgrading—based on the perspective of path dependence. Qual Quant 56:2645–2674. https://doi.org/10.1007/s11135-021-01230-7
Deegan J, Broekel T, Haus-Reve S et al. (2024) How regions diversify into new jobs: from related industries or related occupations? Reg Stud 58:1965–1980. https://doi.org/10.1080/00343404.2024.2335265
Duranton G, Puga D (2023) Urban growth and its aggregate implications. Econometrica 91:2219–2259. https://doi.org/10.3982/ECTA17936
Eeckhout J, Pinheiro R, Schmidheiny K (2014) Spatial Sorting. J Polit Econ 122:554–620. https://doi.org/10.1086/676141
Elekes Z, Baranowska-Rataj A, Eriksson R (2023) Regional diversification and labour market upgrading: local access to skill-related high-income jobs helps workers escaping low-wage employment. Camb J Reg Econ Soc 16:417–430. https://doi.org/10.1093/cjres/rsad016
Ellison G, Glaeser EL, Kerr WR (2010) What causes industry agglomeration? Evidence from coagglomeration patterns. Am Econ Rev 100:1195–1213. https://doi.org/10.1257/aer.100.3.1195
Ewers MC, Dicce R (2018) High-skilled migration and the attractiveness of cities. High-skilled migration: Drivers and policies. Oxford University Press, London
Fan G, Ma G, Wang X (2019) Institutional reform and economic growth of China: 40-year progress toward marketization. Acta Oeconomica 69:7–20. https://doi.org/10.1556/032.2019.69.s1.2
Feser EJ (2003) What regions do rather than make: a proposed set of knowledge-based occupation clusters. Urban Stud 40:1937–1958. https://doi.org/10.1080/0042098032000116059
Florida R (2003) Cities and the creative class. City Community 2(1):3–19. https://doi.org/10.1111/1540-6040.00034
Florida R, Mellander C, Qian H (2012) China's development disconnect. Environ Plan Econ Space 44:628–648. https://doi.org/10.1068/a44284
Fotheringham AS, Wong DWS (1991) The modifiable areal unit problem in multivariate statistical analysis. Environ Plan Econ Space 23:1025–1044. https://doi.org/10.1068/a231025
Frenken K, Boschma RA (2007) A theoretical framework for evolutionary economic geography: industrial dynamics and urban growth as a branching process. J Econ Geogr 7:635–649. https://doi.org/10.1093/jeg/lbm018
Frenken K, Van Oort F, Verburg T (2007) Related variety, unrelated variety and regional economic growth. Reg Stud 41:685–697. https://doi.org/10.1080/00343400601120296
Fritz B, Manduca R (2021) The economic complexity of US metropolitan areas. Reg Stud 55:1299–1310. https://doi.org/10.1080/00343404.2021.1884215
Galetti JR-B, Tessarin MS, Morceiro PC (2021) Skill relatedness, structural change and heterogeneous regions: evidence from a developing country. Pap Reg Sci 100:1355–1376. https://doi.org/10.1111/pirs.12629
Glaeser EL (1999) Learning in cities. J Urban Econ 46:254–277. https://doi.org/10.1006/juec.1998.2121
Glaeser EL, Resseger MG (2010) The complementarity between cities and skills. J Reg Sci 50:221–244. https://doi.org/10.1111/j.1467-9787.2009.00635.x
Goos M, Manning A (2007) Lousy and lovely jobs: the rising polarization of work in Britain. Rev Econ Stat 89:118–133. https://doi.org/10.1162/rest.89.1.118
Goos M, Manning A, Salomons A (2014) Explaining job polarization: routine-biased technological change and offshoring. Am Econ Rev 104:2509–2526. https://doi.org/10.1257/aer.104.8.2509
Gu H, Jie Y, Li Z, Shen T (2021) What drives migrants to settle in Chinese cities: a panel data analysis. Appl Spat Anal Policy 14:297–314. https://doi.org/10.1007/s12061-020-09358-z
Guasch JL, Weiss A (1981) Self-selection in the labor market. Am Econ Rev 71:275–284. https://www.jstor.org/stable/1802778
Guo Q, He C (2017) Production space and regional industrial evolution in China. GeoJournal 82:379–396. https://doi.org/10.1007/s10708-015-9689-4
Han J, Li G, Shen Z et al. (2022) Manufacturing transfer and environmental efficiency: evidence from the spatial agglomeration of manufacturing in China. J Environ Manag 314:115039. https://doi.org/10.1016/j.jenvman.2022.115039
Hane-Weijman E, Eriksson RH, Rigby D (2022) How do occupational relatedness and complexity condition employment dynamics in periods of growth and recession? Reg Stud 56:1176–1189. https://doi.org/10.1080/00343404.2021.1984420
He C, Yan Y, Rigby D (2018) Regional industrial evolution in China. Pap Reg Sci 97:173–198. https://doi.org/10.1111/pirs.12246
Henning M, Eriksson R (2021) Labour market polarisation as a localised process: evidence from Sweden. Camb J Reg Econ Soc 14:69–91. https://doi.org/10.1093/cjres/rsaa030
Hervas-Oliver J, Boronat-Moll C, Sempere-Ripoll F et al. (2024) Industry diversification in industrial districts: Is it about embedded regional or firm-level capabilities? Reg Stud 58:756–767. https://doi.org/10.1080/00343404.2023.2244527
Hidalgo CA (2021) Economic complexity theory and applications. Nat Rev Phys 3:92–113. https://doi.org/10.1038/s42254-020-00275-1
Hidalgo CA, Hausmann R (2009) The building blocks of economic complexity. Proc Natl Acad Sci 106:10570–10575. https://doi.org/10.1073/pnas.0900943106
Hidalgo CA, Klinger B, Barabási AL et al. (2007) The product space conditions the development of nations. Science 317:482–487. https://doi.org/10.1126/science.1144581
Kapur D (2001) Diasporas and technology transfer. J Hum Dev 2:265–286. https://doi.org/10.1080/14649880120067284
Klepper S, Sleeper S (2005) Entry by spinoffs. Manag Sci 51:1291–1306. https://doi.org/10.1287/mnsc.1050.0411
Kofler I, Innerhofer E, Marcher A et al. (2020) The future of high-skilled workers: regional problems and global challenges. Springer Nature
Kogler DF, Evenhuis E, Giuliani E et al. (2023a) Re-imagining evolutionary economic geography. Camb J Reg Econ Soc 16:373–390. https://doi.org/10.1093/cjres/rsad029
Kogler DF, Whittle A, Kim K et al. (2023b) Understanding regional branching: Knowledge diversification via inventor and firm collaboration networks. Econ Geogr 99:471–498. https://doi.org/10.1080/00130095.2023.2242551
Laursen K, Leten B, Nguyen NH et al. (2020) Mounting corporate innovation performance: the effects of high-skilled migrant hires and integration capacity. Res Policy 49:104034. https://doi.org/10.1016/j.respol.2020.104034
Li S, Li X, Lang W et al. (2024) What are the drivers of enterprise diversification in Chinese cities: coming from path-dependence or path-creation forces? J Clean Prod 447:141143. https://doi.org/10.1016/j.jclepro.2024.141143
Li Y, Rigby D (2023) Relatedness, complexity, and economic growth in Chinese cities. Int Reg Sci Rev 46:3–37. https://doi.org/10.1177/01600176221082308
Lise J, Postel-Vinay F (2020) Multidimensional skills, sorting, and human capital accumulation. Am Econ Rev 110:2328–2376. https://doi.org/10.1257/aer.20162002
Martin R, Sunley P (2022) Making history matter more in evolutionary economic geography. ZFW – Adv Econ Geogr 66:65–80. https://doi.org/10.1515/zfw-2022-0014
Mawdsley JK, Somaya D (2016) Employee mobility and organizational outcomes: an integrative conceptual framework and research agenda. J Manag 42:85–113. https://doi.org/10.1177/0149206315616459

Moretti E (2010) Local multipliers. Am Econ Rev 100:373–377. https://doi.org/10.1257/aer.100.2.373

Muneepeerakul R, Lobo J, Shutters ST et al. (2013) Urban economies and occupation space: Can they get "there" from "here"? PloS One 8:e0073676. https://doi.org/10.1371/journal.pone.0073676

Neffke F, Henning M (2013) Skill relatedness and firm diversification. Strateg Manag J 34:297–316. https://doi.org/10.1002/smj.2014

Neffke F, Henning M, Boschma R (2011) How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. Econ Geogr 87:237–265. https://doi.org/10.1111/j.1944-8287.2011.01121.x

Neffke F, Hartog M, Boschma R et al. (2018) Agents of structural change: the role of firms and entrepreneurs in regional diversification. Econ Geogr 94:23–48. https://doi.org/10.1080/00130095.2017.1391691

Nolan C, Morrison E, Kumar I et al. (2011) Linking industry and occupation clusters in regional economic development. Econ Dev Q 25:26–35. https://doi.org/10.1177/0891242410386781

Nooteboom B, Van Haverbeke W, Duysters G et al. (2007) Optimal cognitive distance and absorptive capacity. Res Policy 36:1016–1034. https://doi.org/10.1016/j.respol.2007.04.003

Penrose ET (2009) The Theory of the Growth of the Firm. Oxford University Press

Pugliese E, Cimini G, Patelli A et al. (2019) Unfolding the innovation system for the development of countries: coevolution of Science, Technology and Production. Sci Rep. 9:16440. https://doi.org/10.1038/s41598-019-52767-5

Pylak K, Kogler DF (2021) Successful economic diversification in less developed regions: long-term trends in turbulent times. Reg Stud 55:465–478. https://doi.org/10.1080/00343404.2020.1862782

Qi H, Zhao M, Liu Z (2024) Geographic transformation and its determinants of China's interprovincial skilled migration from 2005 to 2015. Appl Spat Anal Policy 17:175–205. https://doi.org/10.1007/s12061-023-09539-6

Qiao Y, Ascani A, Morrison A (2024) External linkages and regional diversification in China: the role of foreign multinational enterprises. Environ Plan A Econ Space 56:1077–1101. https://doi.org/10.1177/0308518X231223621

Qin X, Hui E, Shen J (2024) The emergence of Industry 4.0 technologies across Chinese cities: the roles of technological relatedness/cross-relatedness and industrial policy. Appl Geogr 173:103455. https://doi.org/10.1016/j.apgeog.2024.103455

Reiman A, Kaivo-oja J, Parviainen E et al. (2023) Human work in the shift to Industry 4.0: a road map to the management of technological changes in manufacturing. Int J Prod Res 62:5613–5630. https://doi.org/10.1080/00207543.2023.2291814

Roback J (1982) Wages, rents, and the quality of life. J Polit Econ 90:1257–1278. https://doi.org/10.1086/261120

Santoalha A, Consoli D, Castellacci F (2021) Digital skills, relatedness and green diversification: a study of European regions. Res Policy 50:104340. https://doi.org/10.1016/j.respol.2021.104340

Schultz TW (1961) Investment in human capital. Am Econ Rev 51:1–17

Shen J, Liu Y (2016) Skilled and less-skilled interregional migration in China: a comparative analysis of spatial patterns and the decision to migrate in 2000–2005. Habitat Int 57:1–10. https://doi.org/10.1016/j.habitatint.2016.06.007

Shutters ST (2024) Urban economic structures as multidimensional networks: a complex systems framework for analyzing economic development trajectories. Complexity 2024:5521625. https://doi.org/10.1155/2024/5521625

Shutters ST, Muneepeerakul R, Lobo J (2016) Constrained pathways to a creative urban economy. Urban Stud 53(16):3439–3454. https://doi.org/10.1177/0042098015616892

Sunley P, Martin R, Gardiner B et al. (2020) In search of the skilled city: Skills and the occupational evolution of British cities. Urban Stud 57:109–133. https://doi.org/10.1177/0042098019834249

Tanner AN (2014) Regional branching reconsidered: emergence of the fuel cell industry in European regions. Econ Geogr 90:403–427. https://doi.org/10.1111/ecge.12055

Teece DJ (1980) Economies of scope and the scope of the enterprise. J Econ Behav Organ 1:223–247. https://doi.org/10.1016/0167-2681(80)90002-5

Teixeira AAC, Queirós ASS (2016) Economic growth, human capital and structural change: a dynamic panel data analysis. Res Policy 45:1636–1648. https://doi.org/10.1016/j.respol.2016.04.006

Thompson WR, Thompson PR (1987) National industries and local occupational strengths: the cross-hairs of targeting. Urban Stud 24:547–560. https://doi.org/10.1080/00420988720080781

Venables AJ (2011) Productivity in cities: Self-selection and sorting. J Econ Geogr 11:241–251. https://doi.org/10.1093/jeg/lbq040

Whittle A, Kogler DF (2020) Related to what? Reviewing the literature on technological relatedness: Where we are now and where can we go? Pap Reg Sci 99:97–113. https://doi.org/10.1111/pirs.12481

Wixe S, Andersson M (2017) Which types of relatedness matter in regional growth? Industry, occupation and education. Reg Stud 51:523–536. https://doi.org/10.1080/00343404.2015.1112369

Wu K, Chen Y, Zhang H et al. (2023) ICTs capability and strategic emerging technologies: evidence from Pearl River Delta. China Appl Geogr 157:103019. https://doi.org/10.1016/j.apgeog.2023.103019

Xu J, Li D, Yeh AG, Zhu S (2025) Degree of technological relatedness, firm heterogeneity, and survival condition: evidence from emerging high-tech firms in the Pearl River Delta, China. Habit Int 159:103378. https://doi.org/10.1016/j.habitatint.2024.103378

Yang Z (2023) Human capital space: a spatial perspective of the dynamics of people and economic relationships. Humanit Soc Sci Commun 10:145. https://doi.org/10.1057/s41599-023-01639-5

Zhou X, Ao R, Yierfanjiang A et al. (2023) Influence of skill relatedness on the location choice of heterogeneous labor force in Chinese prefecture-level cities. PloS One 18:e0289803. https://doi.org/10.1371/journal.pone.0289803

Zhou X, Ao R, Li X et al. (2025) Self-relatedness or cross-relatedness: the co-evolution of industries and occupations among Chinese cities. Tijdschr Econ Soc Geogr 116:109–128. https://doi.org/10.1111/tesg.12662

Zhu S, He C, Zhou Y (2017) How to jump further and catch up? Path-breaking in an uneven industry space. J Econ Geogr 17:521–545. https://doi.org/10.1093/jeg/lbw047

Acknowledgements

This work is supported by the Natural Science Foundation of China (Grant numbers 42001136 and 42271188).

Author contributions

XZ: Conception and design of the study, data analysis and interpretation, literature review, figure drawing and visualization, interpretation of results, revision of the manuscript, optimization of manuscript structure, editing and proofreading of the manuscript. DFK: Drafting of the manuscript, revision of the manuscript, optimization of manuscript structure, editing and proofreading of the manuscript. JC: Conception and design of the study, revision of the manuscript, optimization of manuscript structure, editing and proofreading of the manuscript, funding acquisition. RA: Supervision and guidance, project administration, funding acquisition, provision of resources. JW: Optimization of manuscript structure, revision of the manuscript, technical support. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-025-05771-2.

Correspondence and requests for materials should be addressed to Jing Chen or Rongjun Ao.