Is the occupational evolution of Chinese cities driven by industrial structures? Insights from industry-occupation cross-relatedness

Rongjun Ao, Ling Zhong, Jing Chen, Xiaojing Li, Xiaoqi Zhou

Papers in Evolutionary Economic Geography

25.33

Utrecht University Human Geography and Planning

Is the occupational evolution of Chinese cities driven by industrial structures? Insights from industry-occupation cross-relatedness

Rongjun Ao[1], Ling Zhong[1], Jing Chen[2]*, Xiaojing Li[2], Xiaoqi Zhou[2]

Affinitions:

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

Funding:

The authors would like to acknowledge funding from the Natural Science Foundation of China (grant numbers 42001136 and 42271188).

Correspondence author:

Jing Chen, College of Urban and Environmental Sciences, Central China Normal University; China; Email: jing.chen@ccnu.edu.cn

Note:

This is the working paper version of a published paper. Please cite the published version:

Rongjun Ao, Ling Zhong, Jing Chen*, Xiaojing Li, Xiaoqi Zhou. (2025). Is the occupational evolution of Chinese cities driven by industrial structures? Insights from industry-occupation cross-relatedness. *Journal of Regional Science*. https://doi.org/10.1111/jors.70026

Is the occupational evolution of Chinese cities driven by industrial structures? Insights from industry-occupation cross-relatedness

Abstract: While prior research has emphasized the path-dependent nature of occupational systems, it has paid limited attention to how local industrial structures contribute to occupational change. To address this gap, this study examines how regional occupational evolution is shaped by two distinct mechanisms: (1) path-dependent skill and knowledge transfer, whereby new occupations emerge through the recombination of existing occupational structures; and (2) industry-driven task reconfiguration, through which industrial upgrading reshapes the demand for occupations. To operationalize these dynamics, the concept of industry–occupation cross-relatedness is introduced, capturing the proximity between new occupations and a region's existing industrial portfolio. Drawing on panel data from 241 Chinese cities between 2000 and 2015, the study estimates the effects of both occupational relatedness and cross-relatedness on new occupation specialization. The results reveal that both mechanisms significantly promote occupational evolution, yet they tend to function as substitutes rather than complements. Furthermore, their effects differ across skill levels: high-skilled occupations are more responsive to industrial transformation, low-skilled occupations to occupational pathways, while medium-skilled occupations exhibit relatively weak responsiveness to both. These findings underscore the importance of structural conditions and skill heterogeneity in shaping regional patterns of occupational change.

Keywords: Occupational Evolution; Path Dependence; Chinese Cities; Industry-Occupation Cross-Relatedness; Skill Heterogeneity

JEL classification: R11, O14, N95

1 Introduction

As no one ever "sees" what a regional economy is, economic geographers, regional scientists and urban researchers often describe it through the composition of either industries or occupations (Barbour and Markusen, 2007). Despite the long-time domination of literature by industry-based approaches (Boschma, 2017), occupation-based perspectives are gaining prominence. This shift is driven by two major trends: the recognition of human capital as a key driver of endogenous growth (Martin and Sunley, 1998; Romer, 1990), and the growing functional and spatial separation of economic activities under the background of globalization and technological change (Consoli et al., 2021). Since Thompson and Thompson (1987) argued that regional development should consider not only what regions *make* but also what they *do*, occupational composition has become an essential lens in regional studies. This perspective was reinforced by the development of the Occupational Information Network (O*NET) and Dictionary of Occupational Titles (DOT) by the United States Department of Labor in the 1990s, which enabled a wave of empirical research on occupational structures (Fusillo et al., 2022; Sunley et al., 2020). More recently, global production fragmentation and skill-biased technological change have revealed that regions with similar industrial structures may exhibit substantially different occupational profiles. Thus, scholars increasingly call for a more integrated approach that takes into account both industrial and occupational structures in regional analysis and policy design (Chrisinger et al., 2012; Zhou et al., 2025).

Meanwhile, the evolution of skills and occupations has garnered growing attention from scholars in the field of evolutionary economic geography (EEG) in recent years (Cappelli et al., 2019; Neffke and Henning, 2013). Regional economic diversification is widely understood as a path-dependent branching process, whereby new economic activities emerge by building on and recombining existing related activities rather than at random (Boschma, 2017; Zhu et al., 2017). Within this framework, it is generally assumed that the existing occupational structure of a region plays a critical role in shaping its future developmental trajectory. This assumption is supported by a range of empirical studies notwithstanding variations in dependent variables (e.g., new occupation entry and firm survival), measures of relatedness (e.g., skill or occupation relatedness), spatial analysis units (e.g., regions, cities and labor markets), and time periods (Farinha et al., 2019; Muneepeerakul et al., 2013).

A major limitation in current literature is that a majority of studies focus on the role of occupation relatedness in shaping occupational dynamics, while largely overlooking the influence of local industrial activities (Castaldi and Drivas, 2023; Catalán et al., 2022; Hidalgo, 2023). More specifically, it is widely acknowledged that regions tend to develop new industrial specializations in the presence of related industries (Chen et al., 2024; Hidalgo et al., 2007; Zhu et al., 2017). Likewise, regions are more inclined to diversify into new occupations if already specializing in related ones (Farinha et al., 2019; Muneepeerakul et al., 2013). However, whether industrial structures influence the evolution of occupational structures remains largely underexplored.

Two recent studies have attempted to address this gap. Deegan et al. (2024) examined the short-term effects of industrial activities and skill capabilities on the emergence of complex jobs in Norway and highlighted their complementary roles. Nevertheless, their analysis was confined to a short period (2009–2014) and failed to address long-term dynamics. In contrast, Zhou et al. (2025) investigated the long-term co-evolution of industries and occupations in

China (2000–2015) and centered on self- and cross-relatedness. However, their study hardly analyzed the interaction between industrial and occupational structures in shaping occupational evolution. Moreover, these two studies both overlooked labor market polarization essential for understanding occupational change. More specifically, research on labor economics suggests that technological change is skill-biased: it displaces routine tasks concentrated in medium-skilled jobs (e.g., clerks and machine operators) while increasing productivity and demand for high- and low-skilled non-routine occupations, which is a process known as de-routinization or job polarization (Consoli et al., 2021; Fusillo et al., 2022; Henning and Eriksson, 2021). Based on these insights, this study aimed to explore how industrial activities and occupational portfolios—individually and interactively—co-shape regional occupational diversification across different skill levels among Chinese regions.

This study has the following two contributions. First, regional occupational evolution was linked with industrial dynamics through the lens of cross-relatedness, a concept widely applied in recent research on regional economic diversification (Castaldi and Drivas, 2023; Catalán et al., 2022; Chen et al., 2025). While prior studies have largely treated industrial and occupational trajectories separately, this research is one of the early studies to systematically examine how industrial and skill capabilities co-shape regional occupational evolution. In the present study, the analysis was extended by exploring the interaction between industrial and occupational structures over a longer time span, which is in contrast to existing studies (Deegan et al., 2024; Zhou et al., 2025). Unlike most studies on skill relatedness that concentrate on developed economies (Farinha et al., 2019; Hane-Weijman et al., 2022), this study draws on the case of China, a transitional economy that has undergone rapid industrial and technological transformation in recent decades (He et al., 2018; Zhu et al., 2017). This context provides novel empirical insights into the relationship between skill composition and industrial development.

Second, two closely related strands of literature were bridged as well, namely regional economic diversification and labor market polarization. Most studies on regional occupational diversification have focused on skill complexity(Buyukyazici, 2023; Deegan et al., 2024; Hane-Weijman et al., 2022), but they often overlook skill heterogeneity in terms of abstract, routine and manual tasks. By comparison, recent work in EEG has attached increasing importance to the polarization of the labor market and the unequal impact of occupational evolution across different skill levels (Consoli et al., 2021; Weinstein and Patrick, 2020). According to Zhou et al. (2023) and Fusillo et al. (2022), occupations were classified into three categories: low-, medium- and high-skilled, based on their task content (i.e., abstract, routine or manual tasks). In this way, the dynamic evolution of occupational space was traced, and how industrial and occupational structures co-shape occupational transitions across varying skill levels was compared.

The remaining part of this study is structured as follows. Section 2 reviews the current literature, proposes the theoretical framework and develops the research hypotheses. Section 3 introduces corresponding data sources and methods. Section 4 presents the results, while the final section discusses the results and concludes the findings of this study.

2 Literature review, theoretical framework and research hypotheses

2.1 From industry relatedness to skill relatedness

For the past few years, the concept of relatedness has seen extensive application across different dimensions of economic activities, such as industries (Deegan et al., 2024; He et al., 2018), products (Jin et al., 2025; Zhu et al., 2017), occupations (Hane-Weijman et al., 2022; Muneepeerakul et al., 2013) and technologies (Castaldi and Drivas, 2023; Kogler et al., 2023). Among them, industrial relatedness refers to the similarity or complementarity between industries in terms of infrastructure, production technologies, managerial practices and input factors (Boschma, 2017). Based on theories of cognitive proximity and knowledge spillovers (Boschma, 2005; Frenken et al., 2007), it is further argued that knowledge spillovers are most effective when industries are neither too similar nor distant in cognitive terms. If the cognitive distance is too large—such as between automobile manufacturing and dairy production—there is little room for learning. Conversely, the shared knowledge base of industries limits the scope for knowledge spillovers when they are too close.

EEG scholars have further extended the notion of relatedness to labor market dynamics by developing the concept of skill relatedness, which is analyzed through three types of relatedness networks, namely industry, occupation and skill spaces (Chen et al., 2025). First, industry space captures skill relatedness indirectly via the mapping of inter-industry labor mobility: frequent worker flows between industries indicate shared knowledge requirements, such as those between hotels and restaurants (Neffke and Henning, 2013). Scholars have visualized this through networks where nodes represent industries and links reflect mobility intensity (Cappelli et al., 2019; Neffke and Henning, 2013). Second, occupation space is built using co-occurrence methods, where occupations frequently appearing together are considered skill-related (Farinha et al., 2019; Zhou et al., 2025). This approach shifts the focus from educational attainment to occupation-based skill indicators (Sunley et al., 2020), and represents occupations as nodes connected by shared tasks or transitions (Hane-Weijman et al., 2022). Third, skill space maps relationships between individual skills directly based on whether they are commonly required across different occupations. For instance, spatial orientation and peripheral vision are normally co-required, while spatial orientation and stamina are less co-required, which indicates weak relatedness (Alabdulkareem et al., 2018; Xu et al., 2021). Different from industry and occupation spaces, which capture skill relatedness indirectly, skill space reveals it explicitly.

Empirical studies have mainly zoomed in on two strands once skill relatedness is defined and operationalized (see Table A1). One line of research investigates the impact of skill relatedness on economic outcomes, including employment, productivity, output, etc. (Boschma et al., 2014; Hane-Weijman et al., 2022; Wixe and Andersson, 2017). These studies generally find that optimal cognitive distances promote knowledge spillovers, although some also note potential downsides like wage compression and job poaching (Rørheim and Boschma, 2022). Another line of work uses skill relatedness to analyze regional diversification processes and lays emphasis on the path-dependent nature of skill accumulation and transformation over time (Alabdulkareem et al., 2018; Farinha et al., 2019).

Based on these insights, recent literature has highlighted at least three distinct advantages of shifting the analytical focus from industrial to occupational or skill relatedness. First of all, occupational relatedness emphasizes what regions do, whereas industrial relatedness focuses on what regions make, which thus highlights the role of human capital in regional development (Chrisinger et al., 2012). Second, occupation- and task-based data provide a finer-grained, micro-level lens for analyzing regional diversification, which is typically neglected in firm- or industry-based approaches (Boschma, 2017). Third, occupational relatedness encompasses both manufacturing and service sectors, whereas industrial relatedness generally privileges manufacturing (Farinha et al., 2019; Hane-Weijman et al., 2022). Nevertheless, the wider availability and comparability of industry-level data across time and space means that industrial approaches are still more commonly used in empirical research (Barbour and Markusen, 2007). These considerations point to the value of integrating both perspectives to more comprehensively understand regional economic evolution.

2.2 From relatedness to cross-relatedness

The concept of relatedness, which is typically regarded as the similarity or complementarity of economic activities (Hidalgo et al., 2018), has become central to explaining the path dependence of regional economic diversification across products, industries and occupations (Boschma, 2017; Hidalgo, 2023). In this framework, the emergence of new economic activities is not random but builds on existing regional capabilities, which leads to what is commonly referred to as related diversification. Although related diversification tends to dominate in empirical patterns, unrelated diversification also plays a vital role in avoiding regional lock-in and reducing inequality (Jin et al., 2025; Zhu et al., 2017). In the Chinese context, for example, studies, such as He et al. (2018) and Qiao et al. (2024), demonstrated how extra-regional linkages and institutional factors help break path-dependent trajectories and foster more diverse regional development paths.

Despite a growing body of literature on relatedness, most studies have focused on interactions within a single domain: products, industries or occupations, while largely ignoring the effects across these domains. This lack of integration limits the understanding of the full spectrum of regional diversification processes and challenges the conventional binary of related versus unrelated diversification. In the case of this study, occupational evolution driven by local industrial activities, like the expansion of manufacturing-related jobs, should not be seen as unrelated diversification. Instead, it reflects a form of cross-related diversification, as both industrial and occupational structures are interdependent and representative of disparate dimensions of embedded capabilities of a region (Barbour and Markusen, 2007; Sunley et al., 2020). This suggests the necessity of extending the principle of relatedness to encompass multiple knowledge dimensions.

Recent studies have responded to this call by exploring cross-relatedness—interactions that span across domains, such as technology, industry and occupation. For instance, Pugliese et al. (2019) introduced a tri-layered network capturing the interplay among scientific publications, patents and industrial activities at the country level, thereby opening up a new direction in research on technological diversification. Similarly, Catalán et al. (2022), Castaldi and Drivas (2023), and Qin et al. (2024) contributed to advancing the understanding of cross-domain relationships in regional technology diversification.

Beyond studies on technological diversification, several contributions have turned attention to the occupation–industry nexus. Deegan et al. (2024) advanced this line of inquiry by defining a job as a unique combination of an industry and an occupation. They demonstrated that regions can diversify into new jobs via relevant industries and/or occupations. Their example comparing biomedical and aerospace engineers within the same occupational group but different industrial domains illustrates the need for a more granular and multidimensional perspective. Based on these, Zhou et al. (2025) introduced the concepts of self- and cross-relatedness to distinguish within- and between-domain linkages, while Chen et al. (2025) developed a cross-relatedness and complexity framework to map regional industrial diversification pathways.

2.3 Theoretical framework and research hypotheses

Based on the above literature review, a theoretical framework was proposed to understand regional occupational evolution through two coexisting channels: occupational relatedness and industry–occupation cross-relatedness (Figure 1). The former captures the proximity of knowledge and skills between occupations and lays the foundation for path-dependent occupational evolution. At the micro level, this path dependence operates via three potential mechanisms: (1) skill similarity, where overlapping skillsets enable transitions between occupations; (2) skill complementarity, where interdependent roles co-evolve within organizational structures; (3) local synergy, where co-located occupations mutually support functionally (Farinha et al., 2019; Galetti et al., 2021).

[INSERT FIGURE 1]

In parallel, industry–occupation cross-relatedness reflects how changes in the industrial structures of firms influence occupational dynamics. Despite operating at a more structural level, this channel is similarly grounded in micro-level behavioral mechanisms—specifically, the mobility decisions of workers and the diversification strategies of firms (Xu et al., 2025). From the perspective of labor, workers are disposed to shift toward occupations and industries related to their prior experience to preserve skill value (Chen et al., 2024). From the perspective of firms, diversification decisions are usually rooted in current industrial and occupational capabilities (Klepper, 2007; Teece et al., 1994), or fueled by new demands that reshape regional occupational structures.

Together, occupational relatedness and industry–occupation cross-relatedness stand for two distinct but simultaneously operating mechanisms of occupational change, each of which takes root in micro-level behavioral dynamics. Their independent effects were first examined in this study through the following hypothesis:

- **Hypothesis 1**: *Regions tend to exhibit new occupation specializations in fields related to their existing occupational and industrial structures.*

The relationship between both channels may vary by context though they are expected to independently promote occupational evolution. In this study, two macro-level pathways were further conceptualized: (1) path-dependent skill and knowledge transfer, which are grounded in occupational relatedness and supported by micro-level mechanisms such as skill similarity, complementarity and local synergy (Farinha et al., 2019); and (2) industry-driven task reconfiguration, where industrial upgrading alters occupational demand. In relatively stable environments, firms normally leverage existing occupational capabilities to support industrial transformation, which suggests complementarity (Deegan et al., 2024). On the contrary, firms may bypass traditional occupational structures by hiring or redesigning task systems against the backdrop of rapid technological and industrial change, which results in substitution. To take one example, the rise of electric vehicle manufacturing has spurred demand for new roles like battery engineers and software developers that emerge within established industries but require novel skills (Gong and Hansen, 2023; Zhang et al., 2025). Accordingly, the following competing hypotheses were proposed:

- **Hypothesis 2a**. *Industrial activities and occupational skills are complementary in the process of occupational evolution.*
- **Hypothesis 2b.** *Industrial activities and occupational skills are mutually substitutable in the process of occupational evolution.*

In addition to the interaction between occupational and industrial dynamics, the effects of (cross)-relatedness mechanisms are considered to vary across occupations with different skill levels. Literature on labor market polarization and skill-based technological change was used for reference (Autor et al., 2013; Sunley et al., 2020). Medium-skilled occupations, which typically comprise routine automatable tasks and lack strong professional anchoring, are expected to exhibit weaker responsiveness to both occupational relatedness and industry–occupation cross-relatedness. Their standardized and inflexible task structures restrict the potential for skill recombination across occupations, while their low adaptability constrains their ability to absorb changes in industrial demand (Chen et al., 2024; Consoli et al., 2021). As such, medium-skilled jobs are more susceptible to displacement under exogenous technological and industrial shifts instead of evolving through endogenous relatedness-based mechanisms. Contrarily, high- and low-skilled occupations are prone to showing stronger path dependence, albeit in different ways. High-skilled occupations, which are often embedded in innovation-driven and knowledge-intensive sectors, have a stronger ability to absorb industrial transformation and thus respond more to industry-driven task reconfiguration (Fusillo et al., 2022; Henning and Eriksson, 2021). Low-skilled occupations, which are more deeply rooted in local service provision and labor-intensive segments (Consoli and Sánchez-Barrioluengo, 2019), show a tendency to evolve along established occupational paths. Accordingly, the following hypotheses were proposed:

- **Hypothesis 3a**. *Occupational relatedness has a stronger impact on occupational evolution in high- and low-skilled occupations than in medium-skilled occupations.*
- **Hypothesis 3b.** *Industry–occupation cross-relatedness has a weaker impact on occupational evolution in medium-skilled occupations than in high- and low-skilled occupations.*
- **Hypothesis 3c***: Industry–occupation cross-relatedness has a stronger impact on occupational evolution in high-skilled occupations than in low-skilled occupations.*
- **Hypothesis 3d***: Occupational relatedness has a stronger impact on occupational evolution in low-skilled occupations than in high-skilled occupations.*

Taken collectively, these hypotheses reflect the layered nature of occupational change, which is shaped by the interaction between structural conditions and occupational task characteristics. They also underline the importance of considering skill heterogeneity when evaluating the mechanisms of regional occupational evolution. The following section describes the data sources, variable construction and econometric approach used to test these hypotheses.

3 Data sources and methods

3.1 Data sources

The empirical scope of this study covers mainland China. Due to data availability, the analysis focuses on 241 prefecture-level cities that are consistently observed across four benchmark years: 2000, 2005, 2010, and 2015. Although the original datasets can cover up to 288 cities in the initial and final years (i.e., 2000 and 2015), only 241 cities can be reliably matched across all four waves. To ensure a balanced panel structure, the regression sample is restricted to these 241 cities, while relatedness network analyses based on the full 288-city sample are provided for comparative purposes.

This study relies on three main data sources. First, occupational data are drawn from the 2000 and 2010 Population Censuses and the 2005 and 2015 1% Population Sampling Surveys. These micro-datasets have been widely used in studies of regional labor structures in China (Gu and Jie, 2024; Zhou et al., 2023). Following Muneepeerakul et al. (2013) and Sunley et al. (2020), occupations are used as proxies for workforce skills. In line with Fusillo et al. (2022) and Zhou et al. (2023), occupations are grouped into three skill levels based on task content: abstract (e.g., managers and professionals), routine (e.g., clerks and machine operators), and manual (e.g., cleaners and janitors) are classified as high-, medium-, and low-skilled, respectively.

To address inconsistencies in occupational coding across years, a dual strategy is employed. While the 2000, 2010, and 2015 datasets report occupations at the 3-digit level, the 2005 census includes only 2-digit codes (Chen et al., 2025; Fan et al., 2025). For relatedness network analysis in the initial and final years (2000 and 2015), the more detailed 3-digit data (250 occupations) are used to construct (cross-)relatedness networks. For econometric analysis spanning the entire study period, a harmonized classification of 63 occupations at the 2-digit level is adopted to ensure temporal consistency.

Second, industrial data at the four-digit level are derived from the Annual Survey of Industrial Firms for the years 2000, 2005, 2010, and 2015. This dataset provides detailed firm-level information on employment and industry classification. By matching firm records with four-digit industry codes and city-level administrative codes, we construct an industry–city matrix covering 407 industries across the 241 cities in the balanced panel.

Finally, supplementary city-level variables were gathered from the *China Urban Statistical Yearbook* and the *China Regional Economic Statistical Yearbook*.

3.2 Measurement of occupational relatedness

A key variable in this study is the measure of skill relatedness between occupations. In existing literature, relatedness is commonly assessed using methods like hierarchical classification structures and co-occurrence probabilities of economic activities (Neffke and Henning, 2013). In this study, the co-occurrence approach estimating the conditional probability of two occupations co-located within the same region was adopted. This method is particularly suitable because it enables researchers to capture relatedness both within the occupational domain and across occupations and industries, which is a requirement that alternative approaches cannot meet. The remainder of this subsection outlines the three main steps used to calculate occupational relatedness.

In the first step, the revealed comparative advantage (RCA) is calculated. RCA measures the comparative advantage of a certain occupation in a region when it is compared to the national average as follows:

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

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

where $RCA_{i,c,t}$ represents the comparative advantage of region *c* in occupation *i* at time *t*, and $E_{i,c,t}$ stands for the number of employees engaged in occupation *i* in city *c* at time *t*.

In the second step, according to Hidalgo et al. (2007), the conditional probability that a region has a comparative advantage in occupation *i* given that it has a comparative advantage in occupation *j*, and *vice versa*, was computed. Specifically, two conditional probabilities were calculated: (1) the probability of observing a comparative advantage in occupation *i* given the presence of a comparative advantage in *j*, and (2) the probability of observing a comparative advantage in *j* given the presence of a comparative advantage in *i*. The final relatedness intensity between occupations *i* and *j* was defined as the minimum of these two conditional probabilities, as shown below:

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

where $\phi_{i,j,t}$ stands for the degree of proximity between occupations *i* and *j* at time *t*. A large value of $\phi_{i,j,t}$ indicates the similar requirements of these two occupations for skills and knowledge.

Thirdly, relatedness density measures the degree of relatedness between occupation *i* and the overall occupation structure of city *c* with the form:

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

where $Density_{c,i,t}$ represents the density of occupation *i* in city *c* at time *t*; $\phi_{i,j,t}$ stands for the degree of proximity between occupations *i* and *j* at time *t*; $x_{c,i,t}$ takes the value of 1 if city *c* has a comparative advantage in occupation *i* at time *t* and 0 otherwise.

3.3 Measurement of industrial and occupational cross-relatedness

Cross-proximity between industry *k* and occupation *j* corresponds to the minimum conditional probability that a city develops occupation *j* as it has some degree of comparative advantage in industry *k*, and *vice versa*. This form is based on the original proximity concept proposed by Hidalgo et al. (2007) and further extended by recent scholars like Catalán et al. (2022) and Castaldi and Drivas (2023) with the following form:

$$\phi^X_{k,j,t} = min\{P(x_{k,t}|x_{j,t}), P(x_{j,t}|x_{k,t})\} \quad (5)$$

where $P(x_{k,t}|x_{j,t})$ represents the conditional probability of having a comparative advantage in industry *k*, given that the city has a comparative advantage in occupation *j*; $P(x_{j,t}|x_{k,t})$ stands for the conditional probability of having a comparative advantage in occupation *j*, given that the city has a comparative advantage in industry *k*. The range of $\phi^X_{k,j,t}$ is between 0 and 1. A value close to 0 suggests no co-occurrence between occupation *j* and industry *k*, whereas a value close to 1 indicates the systematic co-occurrence of the industry-occupation pair[1].

Likewise, the industry-occupation cross-density measure corresponds to the average

[1] This differs from the use of micro-level data on occupation–industry pairs in recent studies (Deegan et al., 2024; Hernández-Rodríguez et al., 2025).

industry-occupation cross-proximity of a potential occupation *j* relative to the industrial structure of region *r*, which takes the form:

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

where $x_{k,c,t}$ equals 1 if $RCA_{i,c,t}$ is greater than one, and 0 otherwise. The range of $CrossDensity_{c,j,t}$ is between 0 and 1. A value close to 0 indicates that region *c* has high-level industrial development in areas close to occupation *j*. That is, the current state of the industrial structure features a high degree of proximity to the occupation to be developed.

3.4 Model specifications and variables

To evaluate the roles of industrial and occupational structures on regional occupation evolution, the following equation was constructed:

$$\begin{aligned} Entry_{c,j,t} = \beta_0 + \beta_1 \times Density_{c,j,t-1} + \beta_2 \times CrossDensity_{c,j,t-1} \\ + \beta_3 \times Density_{c,j,t-1} \times CrossDensity_{c,j,t-1} + \varphi_c + \gamma_j \\ + \alpha_t + \varepsilon_{c,j,t} \end{aligned} \tag{7}$$

where the dependent variable $Entry_{c,j,t}$. takes the value of 1 if city *c* has a comparative advantage in occupation *j* at time *t* but no comparative advantage at time *t* -1. As for the independent variables, $Density_{c,j,t-1}$ stands for the relatedness density of occupation *j* in city *c*; $CrossDensity_{c,j,t-1}$ denotes the degree of the association between the occupation class *j* and the current industrial portfolio of city *c* at time *t*-1. The interaction term $Density_{c,j,t-1} \times CrossDensity_{c,j,t-1}$ was included to examine whether the relationship between these two density variables is one of substitution or complementarity (Chen and Wu, 2025; Deegan et al., 2024). Considering that the residual relatedness of different regions in different occupations across different years will lead to the deviation of the estimated results, the fixed effects φ_c, γ_j and α_t were estimated directly using dummy variables for city, occupation and year fixed effects. Additionally, $\varepsilon_{c,i,t}$ is the residual of the applied regression.

Based on the research of Balland et al. (2019), a linear probability model (LPM) was employed as the main estimation strategy. To ensure the robustness of the results, non-linear specifications were also estimated using logit and probit models. Descriptive statistics for all variables are shown in Table 1. The variance inflation factor (VIF) values for all the covariates are below the commonly accepted threshold of 5, which indicates that multicollinearity is not a concern in the models.

[INSERT TABLE 1]

4 Results

4.1 Relatedness network analysis

Several relatedness networks were used for descriptive analysis, including occupation space and industry-occupation cross-space for the years 2000 and 2015[2]. The skill networks of occupations are presented in Figure 2. These networks revealing the proximities between occupations were established using Gephi, an open-source network analysis and visualization software (Bastian et al., 2009). For visualization purposes, only proximities higher than 0.45 were included. Notably, for a given node, the color reflects its skill level and the size reflects

[2] Although Hidalgo (2023) emphasizes the analytical advantages of disaggregated data in studying relatedness, data availability imposes constraints on our temporal coverage. For relatedness network analysis, we use the more detailed 3-digit occupational data to illustrate the concept of cross-relatedness in the starting and ending years. This has also been noted in recent studies dealing with similar data challenges (Fan et al., 2025; Guo et al., 2025). Additionally, Figures B2 and B3 present relatedness networks based on the 2-digit occupation classification, and the main findings remain robust even when using this more aggregated structure.

the number of cities with an RCA greater than one. Similarly, the edge between the two nodes indicates the strength of skill relatedness between the two occupations. Based on these understandings, the networks generally display a core-peripheral pattern in both years. Interconnected nodes, including high-, medium-, and low-skilled occupations, cluster in the core area, while nodes in the peripheral area are predominantly medium-skilled occupations. In addition, medium-skilled occupations comprise the largest proportion of all the nodes. They are found in both core and peripheral areas. Moreover, from 2000 to 2015, the number of edges decreased sharply from 1,833 to 1,239, which made the overall network sparser.

[INSERT FIGURE 2]

Connections between occupations and industries in both years are illustrated in Figure 3. As in previous network visualizations, only proximities above 0.45 are displayed to enhance readability. In 2000, most industry nodes were closely connected with medium-skilled occupations, while relatively few high- and low-skilled occupations exhibited strong ties to industry nodes. This pattern reflects the dominance of routine-based tasks in manufacturing activities at the time. Specifically, industrial production processes were largely standardized, and medium-skilled workers were primarily employed in operational roles requiring adherence to prescribed procedures (Autor et al., 2013; Fusillo et al., 2022). Moreover, the network statistics suggested that more nodes and edges were included in 2015, and the overall cross-relatedness network became more connected, which demonstrated that occupation diversification can be achieved through these newly formed industry-occupation connections. To supplement the analysis, a few linkages between industries and occupations are also illustrated in Figure 3 to explain the concept of cross-relatedness. These linkages are selfexplanatory. For example, petroleum engineering technicians work in the industries of petroleum processing and petroleum product manufacturing, while tailors and sewing workers engage in textile, shoe and hat manufacturing industries.

[INSERT FIGURE 3]

4.2 Econometric analysis

The relationship[3] between the relatedness density (or cross-density) of non-specialized occupations in a city and the probability of that city entering a new occupational specialization is explored in Figure 4. Blue and red bars represent relatedness and cross-relatedness density, respectively. Overall, the probability of specialization tends to increase with higher levels of density and cross-density, suggesting that occupations closely related to a region's existing industrial or occupational structure are more likely to emerge (Catalán et al., 2022; Xu et al., 2025). However, when relatedness density approaches a value of 1, the probability slightly declines, indicating that overly close cognitive proximity may reduce the likelihood of occupation diversification (Boschma, 2005). In contrast, the probability reaches its peak when cross-density is close to 1, highlighting the stronger effect of industrial–occupational proximity. Additionally, these patterns remain consistent across different skill-level subgroups, as shown in Figure B1.

[INSERT FIGURE 4]

A series of econometric models, including baseline regressions and heterogeneity

[3] It is important to note that the econometric analysis uses a harmonized classification of 63 2-digit occupations and a balanced sample of 244 prefecture-level cities consistently observed across the four census years (2000, 2005, 2010, and 2015). Although this reduces the regional sample from the 288 cities available in 2000 and 2015, it ensures consistency in both spatial coverage and occupational classification.

analyses, was conducted to empirically test the proposed framework (Tables 2 and 3). In Table 2, both occupation density and industry–occupation cross-density variables exhibit positive and statistically significant effects on the dependent variable, suggesting that both relatedness and cross-relatedness can contribute to regional occupational entries (supporting Hypothesis 1). In other words, occupational diversification is influenced not only by the existing occupational portfolio of a region but also by its underlying industrial structure *per se*. These findings are consistent with existing studies (Deegan et al., 2024; Zhou et al., 2025) and reinforce the importance of path-dependent skill and knowledge transfer as well as industry-driven task reconfiguration in shaping regional occupational dynamics. Regarding the magnitude of these two variables in Model 5—which includes both control variables and fixed effects—a one-unit increase in occupation density is associated with an approximate 21% increase in the probability of a city developing a new RCA in a given occupation. By comparison, a one-unit increase in cross-density corresponds to a 23% increase. The slightly stronger effect of industry–occupation cross-density implies that the evolving demands of local industrial structures may more directly shape the emergence of new occupations than existing occupational patterns alone, particularly in regions undergoing structural transformation.

[INSERT TABLE 2]

Regarding the interrelation between relatedness and cross-relatedness, Model 5 reveals a significant and negative coefficient for their interaction term, indicating a substitutive relationship and lending support to Hypothesis 2b. This finding stands in contrast to Deegan et al. (2024), who identified a complementary effect in Norwegian regions between 2009 and

2014. The discrepancy likely stems from contextual differences. While Deegan et al. focused on a short-term period within stable economies, this study spans a 15-year horizon (2000–2015) in China, a transitional economy experiencing profound industrial restructuring (He et al., 2018; Zhu et al., 2017). Over a longer period, industrial upgrading may increasingly outpace occupational adaptation, leading to functional substitution. This dynamic reflects the asymmetric pace of structural change: industrial systems evolve rapidly in response to market and technological changes, whereas occupational systems—embedded in education, training, and institutional structures—change more gradually. When new technologies like automation or digitalization emerge, they often demand skill sets that differ fundamentally from existing occupations, prompting firms to bypass traditional career pathways by hiring or retraining workers (Chen et al., 2025; Corradini et al., 2023). As the Chinese proverb aptly puts it, "*It takes ten years to grow trees, but a hundred years to cultivate people*." This structural lag also helps explain why cross-relatedness exerts a stronger influence than occupational relatedness, particularly in rapidly transforming regions, as evidenced in Model 4.

For the heterogeneity analysis, estimation results across occupational subgroups were compared by skill level. Specifically, the regression outcomes for high-, medium-, and low-skilled occupations are reported in Table 3, with corresponding coefficients and confidence intervals presented in Figure 5. In all cases, both occupation relatedness density and industry–occupation cross-relatedness density have positive and statistically significant effects in Table 3, confirming that both mechanisms contribute to occupational evolution. However, as illustrated in Figure 5, the magnitude of these effects varies considerably: medium-skilled occupations are consistently less responsive to both forms of relatedness than high- and low-skilled ones, providing strong empirical support for Hypotheses 3a and 3b. This limited responsiveness likely reflects the nature of medium-skilled jobs, which are typically defined by routine tasks (Consoli et al., 2021; Henning and Eriksson, 2021). These roles lack the adaptability of high-skilled occupations to respond to technological and industrial change, as well as the local embeddedness that allows low-skilled occupations to evolve through occupational linkages. Positioned structurally between the two, medium-skilled occupations often serve as transitional roles with limited anchoring in either knowledge-intensive industries or locally rooted service systems, thereby reducing their capacity to benefit from either path-dependent occupational spillovers or industry-driven task reconfiguration.

[INSERT TABLE 3]

[INSERT FIGURE 5]

Beyond that, the effect of industry–occupation cross-relatedness is most pronounced for high-skilled occupations, which lends support to Hypothesis 3c. Inversely, low-skilled occupations are more strongly influenced by occupational relatedness, which verifies Hypothesis 3d. These patterns suggest that the mechanisms of occupational evolution are skill-dependent and differentiated in structure: high-skilled workers are more capable of adapting to industry-driven transformations (Fusillo et al., 2022; Weinstein and Patrick, 2020), while low-skilled occupations remain embedded in task-based local occupational networks (Consoli and Sánchez-Barrioluengo, 2019). This highlights that it is important to consider task composition and human capital characteristics when analyzing the spatial dynamics of occupational change.

The robustness of the above estimation results was tested through several additional checks. First, consistent with Castaldi and Drivas (2023), baseline and heterogeneity models were re-estimated using logit and probit specifications. Second, the threshold values for determining RCA in occupations and industries were changed using alternative cutoffs of 0.8 and 1.2 according to Zhu et al. (2017) and Zhou et al. (2025). Third, a bootstrapping method was further applied to identify industrial and occupational specializations, as suggested by Tian (2013) and Xiao et al. (2018). Fourth, the industry–occupation cross-relatedness density variable was recalculated using three-digit rather than four-digit industry codes. As shown in Tables A2–A7, these alternative specifications yield results largely consistent with those presented in Tables 2 and 3.

5 Discussion and conclusions

In this study, the mechanisms shaping regional occupational evolution were explored by integrating occupational relatedness and industry–occupation cross-relatedness within a unified analytical framework. Based on recent studies of regional economic diversification and labor market polarization, three sets of hypotheses were tested, and consistent evidence supporting the co-evolution of labor structures and industrial dynamics among Chinese prefectural-level cities from 2000 to 2015 was uncovered. In light of the empirical results, three key findings with their discussion were obtained below.

First, the results confirmed that both occupational relatedness and industry–occupation cross-relatedness have a positive influence on the emergence of new occupational specializations. This suggests that regional occupational evolution is not exclusively driven by pre-existing occupational pathways alone but also by industrial structure. This finding advances the literature by linking micro-level occupational mobility with macro-level structural conditions (Deegan et al., 2024; Zhou et al., 2025). It also broadens the application of the relatedness principle beyond traditional product and industry spaces (Castaldi and Drivas, 2023; Catalán et al., 2022; Hidalgo et al., 2018), highlighting that labor dynamics are shaped by embeddedness across multiple structural dimensions.

Second, together with existing literature, the relationship between two structural channels—occupational and industrial—can be either complementary or substitutive, depending on regional context. In some cases, these mechanisms reinforce each other, which indicates that new occupations tend to emerge when existing skills align with evolving industrial demands (Deegan et al., 2024). Nevertheless, in the case of this study, firms may bypass existing occupational structures altogether by hiring or restructuring internal tasks in the context of rapid technological or industrial transformation (Chen et al., 2025; Corradini et al., 2023). This dual pattern is novel to literature and supports the existence of both path-dependent and industry-driven trajectories in occupational evolution.

Third, the heterogeneity analysis revealed clear differences across skill levels. Medium-skilled occupations are least responsive to both relatedness channels. This result aligns with theories of labor market polarization, in which routine-based, automatable jobs are lacking in flexibility and more vulnerable to displacement (Consoli et al., 2021; Henning and Eriksson, 2021). Furthermore, it was discovered that high-skilled occupations are most responsive to industry–occupation cross-relatedness, while low-skilled occupations are more strongly shaped by occupational relatedness. These patterns mirror distinct mechanisms of adaptability: high-skilled workers usually possess transferable expertise suited to industrial change, whereas low-skilled occupations remain tied to localized, task-based job networks.

Based on the empirical results and theoretical propositions, this study offers the following policy implications. In the first place, the finding that both occupational and industry–occupation relatedness shape occupational evolution to a great degree implies that labor and industrial policies should not be designed in isolation. Instead, regions should adopt integrated strategies that align talent development with industrial path formation. Generic talent-attraction policies like offering tax incentives or subsidized housing may have limited effects in cities where existing industrial and occupational structures cannot absorb and retain high-skilled workers (Yang and Pan, 2020). In peripheral or structurally constrained regions, such policies may result in skill mismatch or human capital loss. Therefore, talent policies must be matched with regional absorption capacity, which ensures that incoming skills align with local economic trajectories (Gu and Jie, 2024).

Second, the complementary and substitutive relationship between occupational and industrial mechanisms highlights the need for context-sensitive workforce interventions. In relatively stable regions, diversification can be reinforced by strengthening coordination between labor supply and industrial expansion (Deegan et al., 2024). Contrarily, in regions undergoing rapid industrial restructuring, anticipatory workforce planning is essential. This includes predictive analytics for skill obsolescence, reskilling programs aligned with emerging industries, and mobility incentives across industrial sectors and occupational groups (Chen et al., 2025; Corradini et al., 2023). These measures can avoid structural mismatch during industrial upgrading and ensure labor force adaptability.

Third, the distinct responsiveness across skill groups underscores the need for skill-level-differentiated labor policies. While high-skilled labor plays a central role in regional innovation strategies (Wu et al., 2023), low-skilled occupations remain indispensable in supporting industrial activities and occupational systems (Fusillo et al., 2022). For instance, they provide essential functions such as cleaning, logistics and food services that enable higher-skilled workers to concentrate on cognitive and technical tasks, which thereby enhances overall productivity. As such, labor market interventions should avoid over-focusing on elite talents and instead acknowledge the complementary role of low-skilled occupations. For medium-skilled workers, special attention is required to circumvent marginalization owing to routinization, which may involve upskilling pathways or transitions to more resilient roles.

This study on occupation evolution still has several limitations that may be considered for future research directions. First, longitudinal panel data were used at the prefecture level, but the analysis was limited to discrete time intervals (2000, 2005, 2010, 2015), which may not fully capture the continuous and dynamic nature of occupational upgrading. Future research could employ higher-frequency or micro-level datasets like matched employer-employee records and individual occupation trajectories (Deegan et al., 2024; Hernández-Rodríguez et al., 2025) to better reflect the timing, sequencing and spatial dynamics of occupational transitions.

Second, the measurement framework focuses on relatedness structures but fails to explicitly incorporate the complexity dimension of occupations. Given that recent work increasingly integrates relatedness and complexity to explain structural change (Balland et al., 2019; Chen et al., 2025; Hane-Weijman et al., 2022), future research should take how these two dimensions jointly shape occupational evolution into consideration. For instance, it is worth investigating whether occupational relatedness or industry–occupation cross-relatedness plays a more decisive role when regions diversify into more complex occupations. By this means, the understanding of how knowledge recombination interacts with capability thresholds in the diversification process could be advanced.

Finally, while this study assumes that relatedness effects are broadly uniform across regions, the conditions under which relatedness translates into actual diversification outcomes may differ depending on structural and institutional contexts. Based on Hu et al. (2022), Liu et al. (2025) and Fu et al. (2024), future research could examine how structure–agency interactions, such as local governance capacity and policy instruments, influence the strength of relatedness effects. One promising approach is to test interaction terms between relatedness variables and contextual moderators to assess whether the effects of relatedness and cross-relatedness are amplified or constrained by local conditions.

6 References

Alabdulkareem A, Frank M R, Sun L, AlShebli B, Hidalgo C, Rahwan I, 2018, “Unpacking the polarization of workplace skills” *Science Advances* **4**(7) eaao6030

Autor D H, Dorn D, Hanson G H, 2013, “The China syndrome: Local labor market effects of import competition in the United States” *American Economic Review* **103**(6) 2121–2168

Balland P-A, Boschma R, Crespo J, Rigby D L, 2019, “Smart specialization policy in the European union: Relatedness, knowledge complexity and regional diversification” *Regional Studies* **53**(9) 1252–1268

Barbour E, Markusen A, 2007, “Regional occupational and industrial structure: does one imply the other?” *International Regional Science Review* **30**(1) 72–90

Bastian M, Heymann S, Jacomy M, 2009, “Gephi: An open source software for exploring and manipulating networks”, in *Proceedings of the International AAAI Conference on Web and Social Media* Proceedings of the International AAAI Conference on Web and Social Media, pp 361–362

Boschma R, 2005, “Proximity and innovation: a critical assessment” *Regional Studies* **39**(1) 61–74

Boschma R, 2017, “Relatedness as driver of regional diversification: a research agenda” *Regional Studies* **51**(3) 351–364

Boschma R, Eriksson R H, Lindgren U, 2014, “Labour market externalities and regional growth in Sweden: the importance of labour mobility between skill-related industries” *Regional Studies* **48**(10) 1669–1690

Buyukyazici D, 2023, “Skills for smart specialisation: relatedness, complexity and evaluation of priorities” *Papers in Regional Science* **102**(5) 1007–1031

Cappelli R, Boschma R, Weterings A, 2019, “Labour mobility, skill-relatedness and new plant survival across different development stages of an industry” *Environment and Planning A: Economy and Space* **51**(4) 869–890

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” *Economic Geography* **99**(3) 253–284

Catalán P, Navarrete C, Figueroa F, 2022, “The scientific and technological cross-space: is technological diversification driven by scientific endogenous capacity?” *Research Policy* **51**(8) 104016

Chen J, Li X, Zhou X, Ao R, 2025, “Skill capabilities behind the scenes: the role of occupational portfolio in regional industrial evolution” *Regional Studies* **59**(1) 2416535

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" *Annals of Regional Science* **72**(3) 1015–1045

Chen Y, Wu K, 2025, "Integrating artificial intelligence into regional technological domains: the role of intra- and extra-regional AI relatedness" *Cambridge Journal of Regions Economy and Society* **18**(1) 111–130

Chrisinger C K, Fowler C S, Kleit R G, 2012, "Shared skills: Occupation clusters for poverty alleviation and economic development in the US" *Urban Studies* **49**(15) 3403–3425

Consoli D, Fusillo F, Orsatti G, Quatraro F, 2021, "Skill endowment, routinisation and digital technologies: evidence from U.S. metropolitan areas" *Industry and Innovation* **28**(8) 1017–1045

Consoli D, Sánchez-Barrioluengo M, 2019, "Polarization and the growth of low-skill service jobs in Spanish local labor markets" *Journal of Regional Science* **59**(1) 145–162

Corradini C, Morris D, Vanino E, 2023, "Towards a regional approach for skills policy" *Regional Studies* **57**(6) 1043–1054

Deegan J, Broekel T, Haus-Reve S, Fitjar R D, 2024, "How regions diversify into new jobs: from related industries or related occupations?" *Regional Studies* **58**(11) 1965–1980

Fan J, Liu N, Tang W, 2025, "Skills and the city in China" *Regional Science and Urban Economics* **111** 104082

Farinha T, Balland P-A, Morrison A, Boschma R, 2019, "What drives the geography of jobs in the US? Unpacking relatedness" *Industry and Innovation* **26**(9) 988–1022

Fitjar R D, Timmermans B, 2017, "Regional skill relatedness: towards a new measure of regional related diversification" *European Planning Studies* **25**(3) 516–538

Fitjar R D, Timmermans B, 2019, "Relatedness and the resource curse: is there a liability of relatedness?" *Economic Geography* **95**(3) 231–255

Frenken K, Van Oort F, Verburg T, 2007, "Related variety, unrelated variety and regional economic growth" *Regional Studies* **41**(5) 685–697

Fu T, Liao K, Chen J, Cai Y, 2024, "Divergent multiple actor-driven path creation and spatial evolution of emerging digital economy industry in guangzhou city, china" *Applied Geography* **169** 103341

Fusillo F, Consoli D, Quatraro F, 2022, "Resilience, skill endowment, and diversity: evidence from US metropolitan areas" *Economic Geography* **98**(2) 170–196

Galetti J R B, Tessarin M S, Morceiro P C, 2021, "Skill relatedness, structural change and heterogeneous regions: evidence from a developing country" *Papers in Regional Science*

100(6) 1355–1377

Gong H, Hansen T, 2023, "The rise of China's new energy vehicle lithium-ion battery industry: the coevolution of battery technological innovation systems and policies" *Environmental Innovation and Societal Transitions* **46** 100689

Gu H, Jie Y, 2024, "Escaping from 'dream city'? Housing price, talent, and urban innovation in China" *Habitat International* **145** 103015

Guo Q, Zhou Y, Feng H, Hewings G, 2025, "Chinese firms' resilience after financial crisis in the context of local economic growth targets" *Regional Studies* **59**(1) 2354367

Hane-Weijman E, Eriksson R H, Rigby D, 2022, "How do occupational relatedness and complexity condition employment dynamics in periods of growth and recession?" *Regional Studies* **56**(7) 1176–1189

He C, Yan Y, Rigby D, 2018, "Regional industrial evolution in China" *Papers in Regional Science* **97**(2) 173–199

Henning M, Eriksson R H, 2021, "Labour market polarisation as a localised process: Evidence from Sweden" *Cambridge Journal of Regions, Economy and Society* **14**(1) 69–91

Hernández-Rodríguez E, Boschma R, Morrison A, Ye X, 2025, "Functional upgrading and downgrading in global value chains: Evidence from EU regions using a relatedness/complexity framework" *Papers in Regional Science* **104**(1) 100072

Hidalgo C A, 2023, "The policy implications of economic complexity" *Research Policy* **52**(9) 104863

Hidalgo C A, Balland P-A, Boschma R, Delgado M, Feldman M, Frenken K, Glaeser E, He C, Kogler D F, Morrison A, Neffke F, Rigby D, Stern S, Zheng S, Zhu S, 2018, "The principle of relatedness", in *Unifying Themes in Complex Systems IX* Eds A J Morales, C Gershenson, D Braha, A A Minai, and Y Bar-Yam Springer Proceedings in Complexity (Springer International Publishing, Cham), pp 451–457, https://link.springer.com/10.1007/978-3-319-96661-8_46

Hidalgo C A, Klinger B, Barabási A, Hausmann R, 2007, "The product space conditions the development of nations" *Science* **317**(5837) 482–487

Hu X, Li L, Dong K, 2022, "What matters for regional economic resilience amid COVID-19? Evidence from cities in northeast China" *Cities* **120** 103440

Jara-Figueroa C, Jun B, Glaeser E L, Hidalgo C A, 2018, "The role of industry-specific, occupation-specific, and location-specific knowledge in the growth and survival of new firms" *Proceedings of the National Academy of Sciences* **115**(50) 12646–12653

Jin W, Zhu S, Li D, He C, 2025, "Transcending the geographical stickiness of complex economic activities via strategic unrelated diversification” *Structural Change and Economic Dynamics* **73** 65–76

Klepper S, 2007, “Disagreements, Spinoffs, and the Evolution of Detroit as the Capital of the U.S. Automobile Industry” *Management Science* **53**(4) 616–631

Kogler D F, Whittle A, Kim K, Lengyel B, 2023, “Understanding regional branching: Knowledge diversification via inventor and firm collaboration networks” *Economic Geography* **99**(5) 471–498

Liu Z, Zhu S, He C, 2025, “Intercity personnel exchange is more effective than policy transplantation at reducing water pollution” *Nature Cities* **2**(3) 210–222

Martin R, Sunley P, 1998, “Slow convergence? The new endogenous growth theory and regional development” *Economic Geography* **74**(3) 201–227

Muneepeerakul R, Lobo J, Shutters S T, Goméz-Liévano A, Qubbaj M R, 2013, “Urban economies and occupation space: can they get ‘there’ from ‘here’?” *PLOS One* **8**(9) e73676

Neffke F, Henning M, 2013, “Skill relatedness and firm diversification” *Strategic Management Journal* **34**(3) 297–316

Pugliese E, Cimini G, Patelli A, Zaccaria A, Pietronero L, Gabrielli A, 2019, “Unfolding the innovation system for the development of countries: coevolution of Science, Technology and Production” *Scientific Reports* **9**(1) 16440

Qiao Y, Ascani A, Morrison A, 2024, “External linkages and regional diversification in China: The role of foreign multinational enterprises” *Environment and Planning A: Economy and Space* **56**(4) 1077–1101

Qin X, Hui E C-M, Shen J, 2024, “The emergence of industry 4.0 technologies across Chinese cities: the roles of technological relatedness/cross-relatedness and industrial policy” *Applied Geography* **173** 103455

Romer P, 1990, “Endogenous technological change” *Journal of Political Economy* **98**(5) 71–102

Rørheim J-E, Boschma R, 2022, “Skill-relatedness and employment growth of firms in times of prosperity and crisis in an oil-dependent region” *Environment and Planning A: Economy and Space* **54**(4) 676–692

Sunley P, Martin R, Gardiner B, Pike A, 2020, “In search of the skilled city: Skills and the occupational evolution of British cities” *Urban Studies* **57**(1) 109–133

Teece D J, Rumelt R, Dosi G, Winter S, 1994, “Understanding corporate coherence: Theory and evidence” *Journal of Economic Behavior & Organization* **23**(1) 1–30

Thompson W R, Thompson P R, 1987, “National Industries and Local Occupational Strengths: the Cross-Hairs of Targeting” *Urban Studies* **24**(6) 547–560

Tian Z, 2013, "Measuring agglomeration using the standardized location quotient with a bootstrap method" *Journal of Regional Analysis & Policy* **43**(2) 186–197

Timmermans B, Boschma R, 2014, "The effect of intra- and inter-regional labour mobility on plant performance in Denmark: the significance of related labour inflows" *Journal of Economic Geography* **14**(2) 289–311

Weinstein A, Patrick C, 2020, "Recession-proof skills, cities, and resilience in economic downturns" *Journal of Regional Science* **60**(2) 348–373

Wixe S, Andersson M, 2017, "Which types of relatedness matter in regional growth? Industry, occupation and education" *Regional Studies* **51**(4) 523–536

Wu K, Chen Y, Zhang H, Liu Y, Wang M, Ye Y, Gong W, 2023, "ICTs capability and strategic emerging technologies: evidence from Pearl River Delta" *Applied Geography* **157** 103019

Xiao J, Boschma R, Andersson M, 2018, "Industrial diversification in Europe: the differentiated role of relatedness" *Economic Geography* **94**(5) 514–549

Xu J, Li D, Yeh A G, Zhu S, 2025, "Degree of technological relatedness, firm heterogeneity, and survival condition: evidence from emerging high-tech firms in the Pearl River Delta, China" *Habitat International* **159** 103378

Xu W, Qin X, Li X, Chen H, Frank M, Rutherford A, Reeson A, Rahwan I, 2021, "Developing China's workforce skill taxonomy reveals extent of labor market polarization" *Humanities and Social Sciences Communications* **8**(1) 1–10

Yang Z, Pan Y, 2020, "Human capital, housing prices, and regional economic development: Will 'vying for talent' through policy succeed?" *Cities* **98** 102577

Zhang J, Mao X, Zhang R, Li Y, Yuan Q, 2025, "From path creation to path transformation: the geography of new energy vehicle production in China" *Applied Geography* **179** 103614

Zhou X, Ao R, Aihemaitijiang Y, Chen J, Tang H, 2023, "Influence of skill relatedness on the location choice of heterogeneous labor force in Chinese prefecture-level cities" *PLOS ONE* **18**(8) e0289803

Zhou X, Ao R, Li X, Chen J, 2025, "Self-relatedness or cross-relatedness: the co-evolution of industries and occupations among chinese cities" *Tijdschrift voor Economische en Sociale Geografie* **116**(1) 109–128

Zhu S, He C, Zhou Y, 2017, "How to jump further and catch up? Path-breaking in an uneven industry space" *Journal of Economic Geography* lbw047

7 Appendices

[INSERT TABLES A1-A7]

[INSERT FIGURES B1-B3]

8 Tables and figures

Table 1. Descriptive statistics

Variables	Definition	Observations	Mean	SD.	Min	Max	VIF
$Entry_{c,j,t}$	1 if the RCA of occupation j is below 1 at time t−1 but above 1 at time t in city c and zero otherwise.	45,549	0.137	0.344	0	1	
Density	Occupation relatedness density at time t-1.	45,549	0.347	0.188	0	0.891	1.319
CrossDensity	Industry-occupation cross-relatedness density at time t-1.	45,549	0.194	0.084	0.006	0.644	2.245
Population density	Logged population density at time t-1.	45,549	0.033	0.032	0.001	0.253	1.621
GDP	Logged gross domestic product of city c at time t-1.	45,549	1.561	1.100	-1.752	5.145	3.899
Stock	Logged employment stock of a city at time t-1.	45,549	0.047	0.068	0.005	0.878	1.577
Size	Logged employment size of an occupation at time t-1.	45,549	0.081	0.022	0.026	0.158	1.106

Notes: SD stands for standard deviation; VIF stands for variance inflation factor.

Table 2. Baseline regressions

	(1)	(2)	(3)	(4)	(5)
Density		0.222***		0.216***	0.395***
		(0.010)		(0.010)	(0.021)
CrossDensity			0.432***	0.405***	0.721***
			(0.029)	(0.029)	(0.042)
Interaction					-0.836***
					(0.088)
GDP	-0.000	-0.013***	-0.024***	-0.035***	-0.036***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Population density	-0.197***	-0.243***	-0.330***	-0.367***	-0.327***
	(0.063)	(0.063)	(0.064)	(0.064)	(0.064)
Stock	-0.015	-0.092***	-0.051*	-0.124***	-0.027
	(0.028)	(0.028)	(0.028)	(0.028)	(0.030)
Size	1.601***	1.453***	1.332***	1.206***	1.107***
	(0.067)	(0.067)	(0.070)	(0.070)	(0.070)
Constant	0.062***	-0.016*	0.011	-0.062***	-0.124***
	(0.008)	(0.009)	(0.009)	(0.009)	(0.011)
Occupation F.E.	Yes	Yes	Yes	Yes	Yes
City F.E.	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes
N	45,549	45,549	45,549	45,549	45,549
Adjusted R^2	0.015	0.026	0.020	0.030	0.032

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

Table 3. Occupational heterogeneity

	(1)	(2)	(3)	(4)	(5)	(6)
	High	High	Medium	Medium	Low	Low
Density	0.215***	0.394***	0.129***	0.309***	0.279***	0.472***
	(0.018)	(0.036)	(0.018)	(0.037)	(0.015)	(0.035)
CrossDensity	0.511***	0.895***	0.277***	0.595***	0.415***	0.725***
	(0.062)	(0.093)	(0.056)	(0.079)	(0.040)	(0.062)
Interaction		-0.893***		-0.869***		-0.865***
		(0.162)		(0.162)		(0.146)
GDP	-0.045***	-0.046***	-0.021***	-0.024***	-0.039***	-0.040***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Population density	-1.153***	-0.888***	-0.038	0.027	-0.321***	-0.304***
	(0.203)	(0.211)	(0.122)	(0.123)	(0.088)	(0.088)
Stock	-0.052	0.005	-0.076	0.049	-0.170***	-0.060
	(0.049)	(0.050)	(0.057)	(0.062)	(0.049)	(0.053)
Size	0.948***	0.837***	0.927***	0.808***	1.385***	1.308***
	(0.118)	(0.120)	(0.136)	(0.138)	(0.114)	(0.115)
Constant	-0.011	-0.083***	0.197***	0.129***	-0.040**	-0.102***
	(0.016)	(0.020)	(0.041)	(0.042)	(0.019)	(0.020)
Occupation F.E.	Yes	Yes	Yes	Yes	Yes	Yes
City F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
N	13,737	13,737	11,568	11,568	20,244	20,244
Adjusted R^2	0.031	0.032	0.021	0.023	0.037	0.038

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

Table A1. Empirical studies on skill relatedness

Article	Study region	Skill relatedness measurement	Focus(es)
Neffke and Henning (2013)	Swedish labor markets	Cross-industry labor flows	Firm diversification
Muneepeerakul et al. (2013)	U.S. Metro Areas	Occupation co-occurrence	Occupation relatedness measurement
Boschma et al. (2014)	Swedish functional regions	Cross-industry labor flows	Regional growth
Timmermans and Boschma (2014)	Danish labor markets	Cross-industry labor flows	Plant performance
Fitjar and Timmermans (2017)	U.S. Metro Areas	Occupation co-occurrence	Occupation diversification
Wixe and Anderson (2017)	Swedish municipalities	Occupation classification code	Regional growth
Alabdulkareem et al. (2018)	U.S. Metro Areas	Skill co-occurrence	The evolution of high- and low-wage occupations
Jara-Figueroa et al. (2018)	Brazil micro-regions	Cross-occupation labor flows	Growth and survival of new firms
Fitjar and Timmermans (2019)	Stavanger, Norway	Cross-industry labor flows	Industry growth
Cappelli et al. (2019)	Dutch municipalities	Cross-industry labor flows	New plant survival
Farinha et al. (2019)	U.S. Metro Areas	Occupation co-occurrence	Occupation diversification
Galetti et al. (2021)	Brazil micro-regions	Skill content of occupations	Industry diversification and employment growth
Xu et al. (2021)	Chinese cities	Skill co-occurrence	Economic growth and labor immigration
Rørheim and Boschma (2022)	Stavanger, Norway	Cross-industry labor flows	Employment growth
Hane-Weijman et al. (2022)	Swedish labor markets	Cross-occupation labor flows	Employment growth
Shutters (2024)	U.S. Metro Areas	Occupation co-occurrence	Multidimensional relatedness networks
Deegan et al. (2024)	Norway labor markets	Cross-occupation labor flows	Industry-occupation diversification
Zhou et al. (2025)	Chinese cities	Occupation-industry co-occurrence	Industry-occupation co-evolution

Notes: Due to space limitation, working papers are not included. All the articles are listed in chronological order with the most recent at the bottom.

Table A2. Robustness checks with logit regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Full model	Interaction	High	High	Medium	Medium	Low	Low
Self-relatedness	1.852***	3.992***	1.671***	3.671***	1.319***	3.711***	2.386***	4.736***
	(0.079)	(0.199)	(0.134)	(0.327)	(0.175)	(0.413)	(0.126)	(0.335)
Cross-relatedness	3.607***	7.600***	4.040***	8.454***	3.022***	7.404***	3.705***	7.766***
	(0.247)	(0.423)	(0.480)	(0.833)	(0.561)	(0.877)	(0.345)	(0.646)
Interaction		-9.730***		-9.637***		-11.193***		-10.247***
		(0.839)		(1.445)		(1.807)		(1.366)
GDP	-0.312***	-0.313***	-0.358***	-0.363***	-0.230***	-0.255***	-0.341***	-0.349***
	(0.025)	(0.025)	(0.044)	(0.044)	(0.058)	(0.057)	(0.039)	(0.039)
Population density	-3.166***	-2.795***	-8.835***	-6.550***	-0.336	0.491	-2.543***	-2.373***
	(0.623)	(0.619)	(1.764)	(1.779)	(1.225)	(1.210)	(0.849)	(0.838)
Stock	-1.067***	-0.037	-0.477	0.039	-0.847	0.633	-1.626***	-0.332
	(0.271)	(0.271)	(0.431)	(0.442)	(0.563)	(0.517)	(0.531)	(0.488)
Size	10.816***	9.806***	8.444***	7.370***	9.686***	8.428***	11.861***	11.031***
	(0.618)	(0.629)	(1.053)	(1.077)	(1.231)	(1.256)	(0.980)	(0.993)
Constant	-3.718***	-4.538***	-3.164***	-4.032***	-1.271***	-2.206***	-3.536***	-4.397***
	(0.087)	(0.112)	(0.139)	(0.193)	(0.369)	(0.393)	(0.170)	(0.203)
Occupation F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	45,549	45,549	13,737	13,737	11,568	11,568	20,244	20,244
Pseudo R^2	0.040	0.043	0.038	0.041	0.033	0.037	0.048	0.051

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

Table A3 Robustness checks with probit regressions

	(1) Full model	(2) Interaction	(3) High	(4) High	(5) Medium	(6) Medium	(7) Low	(8) Low
Self-relatedness	1.036***	2.186***	0.958***	2.040***	0.736***	1.982***	1.326***	2.572***
	(0.044)	(0.108)	(0.076)	(0.180)	(0.094)	(0.220)	(0.070)	(0.182)
Cross-relatedness	2.003***	4.105***	2.277***	4.625***	1.613***	3.875***	2.076***	4.171***
	(0.134)	(0.227)	(0.263)	(0.451)	(0.294)	(0.462)	(0.188)	(0.344)
Interaction		-5.254***		-5.250***		-5.860***		-5.461***
		(0.458)		(0.797)		(0.958)		(0.745)
GDP	-0.170***	-0.173***	-0.200***	-0.203***	-0.121***	-0.137***	-0.187***	-0.192***
	(0.014)	(0.014)	(0.024)	(0.024)	(0.030)	(0.030)	(0.021)	(0.021)
Population density	-1.660***	-1.439***	-5.009***	-3.706***	-0.172	0.247	-1.406***	-1.298***
	(0.323)	(0.321)	(0.950)	(0.968)	(0.630)	(0.628)	(0.442)	(0.437)
Stock	-0.597***	-0.030	-0.257	0.040	-0.453	0.345	-0.870***	-0.192
	(0.143)	(0.146)	(0.231)	(0.237)	(0.289)	(0.283)	(0.271)	(0.262)
Size	6.005***	5.478***	4.801***	4.241***	5.311***	4.635***	6.538***	6.120***
	(0.339)	(0.345)	(0.585)	(0.596)	(0.685)	(0.697)	(0.533)	(0.539)
Constant	-3.718***	-4.538***	-3.164***	-4.032***	-1.271***	-2.206***	-3.536***	-4.397***
	(0.087)	(0.112)	(0.139)	(0.193)	(0.369)	(0.393)	(0.170)	(0.203)
Occupation F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	45,549	45,549	13,737	13,737	11,568	11,568	20,244	20,244
Pseudo R^2	0.035	0.039	0.034	0.037	0.030	0.034	0.044	0.047

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

Table A4. Robustness checks with an alternative LQ threshold value of 0.8

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Full model	Interaction	High	High	Medium	Medium	Low	Low
Self-relatedness	0.216***	0.381***	0.234***	0.404***	0.128***	0.264***	0.269***	0.451***
	(0.009)	(0.020)	(0.017)	(0.035)	(0.018)	(0.035)	(0.015)	(0.034)
Cross-relatedness	0.305***	0.614***	0.383***	0.769***	0.195***	0.452***	0.322***	0.634***
	(0.027)	(0.041)	(0.060)	(0.092)	(0.054)	(0.077)	(0.038)	(0.061)
Interaction		-0.708***		-0.792***		-0.604***		-0.745***
		(0.075)		(0.146)		(0.139)		(0.125)
GDP	-0.036***	-0.035***	-0.042***	-0.041***	-0.014*	-0.015*	-0.045***	-0.045***
	(0.004)	(0.004)	(0.006)	(0.006)	(0.008)	(0.008)	(0.007)	(0.007)
Population density	-0.348***	-0.298***	-1.172***	-0.862***	-0.006	0.039	-0.285***	-0.266***
	(0.063)	(0.063)	(0.176)	(0.190)	(0.123)	(0.123)	(0.088)	(0.087)
Stock	-0.002	-0.001	-0.001	-0.003	-0.011	-0.007	0.001	0.004
	(0.004)	(0.004)	(0.005)	(0.005)	(0.008)	(0.008)	(0.007)	(0.007)
Size	0.092***	0.085***	0.073***	0.065***	0.071***	0.064***	0.113***	0.106***
	(0.005)	(0.005)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)
Constant	0.053***	-0.014	0.066**	-0.021	0.243***	0.189***	0.120***	0.057*
	(0.016)	(0.018)	(0.026)	(0.031)	(0.048)	(0.050)	(0.031)	(0.033)
Occupation F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	45,549	45,549	13,737	13,737	11,568	11,568	20,244	20,244
Adjusted R^2	0.029	0.031	0.032	0.033	0.021	0.022	0.035	0.036

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

Table A5. Robustness checks with an alternative LQ threshold value of 1.2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Full model	Interaction	High	High	Medium	Medium	Low	Low
Self-relatedness	0.195***	0.428***	0.184***	0.399***	0.118***	0.316***	0.269***	0.539***
	(0.010)	(0.021)	(0.018)	(0.036)	(0.019)	(0.037)	(0.016)	(0.036)
Cross-relatedness	0.451***	0.816***	0.551***	0.977***	0.310***	0.625***	0.479***	0.855***
	(0.032)	(0.042)	(0.067)	(0.094)	(0.060)	(0.078)	(0.044)	(0.061)
Interaction		-1.215***		-1.220***		-1.054***		-1.370***
		(0.095)		(0.183)		(0.176)		(0.164)
GDP	-0.032***	-0.031***	-0.039***	-0.037***	-0.012	-0.014*	-0.040***	-0.039***
	(0.004)	(0.004)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)
Population density	-0.429***	-0.341***	-1.173***	-0.716***	-0.070	0.022	-0.380***	-0.339***
	(0.063)	(0.064)	(0.181)	(0.201)	(0.122)	(0.123)	(0.088)	(0.088)
Stock	-0.002	0.000	-0.001	-0.003	-0.011	-0.006	-0.000	0.005
	(0.004)	(0.004)	(0.005)	(0.005)	(0.008)	(0.008)	(0.007)	(0.007)
Size	0.094***	0.085***	0.077***	0.069***	0.069***	0.059***	0.115***	0.106***
	(0.005)	(0.005)	(0.007)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)
Constant	0.077***	0.014	0.116***	0.029	0.256***	0.209***	0.122***	0.069**
	(0.015)	(0.016)	(0.025)	(0.028)	(0.047)	(0.048)	(0.031)	(0.031)
Occupation F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	45,549	45,549	13,737	13,737	11,568	11,568	20,244	20,244
Adjusted R^2	0.029	0.032	0.029	0.031	0.021	0.024	0.036	0.038

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

Table A6. Robustness checks with standardized location quotients

	(1) Full model	(2) Interaction	(3) High	(4) High	(5) Medium	(6) Medium	(7) Low	(8) Low
Self-relatedness	0.164***	0.365***	0.166***	0.381***	0.091***	0.229***	0.228***	0.447***
	(0.010)	(0.018)	(0.019)	(0.032)	(0.019)	(0.031)	(0.016)	(0.029)
Cross-relatedness	0.349***	0.786***	0.408***	0.881***	0.277***	0.609***	0.331***	0.802***
	(0.032)	(0.045)	(0.064)	(0.090)	(0.059)	(0.086)	(0.046)	(0.067)
Interaction		-1.122***		-1.281***		-0.788***		-1.218***
		(0.083)		(0.162)		(0.150)		(0.132)
GDP	-0.034***	-0.034***	-0.037***	-0.035***	-0.017**	-0.021***	-0.042***	-0.044***
	(0.004)	(0.004)	(0.006)	(0.006)	(0.008)	(0.008)	(0.007)	(0.007)
Population density	-0.520***	-0.404***	-1.221***	-0.673***	-0.105	-0.041	-0.440***	-0.385***
	(0.065)	(0.065)	(0.184)	(0.201)	(0.124)	(0.125)	(0.090)	(0.090)
Stock	0.003	0.005	0.003	0.001	-0.006	-0.002	0.005	0.011*
	(0.004)	(0.004)	(0.005)	(0.005)	(0.008)	(0.008)	(0.007)	(0.007)
Size	0.110***	0.102***	0.087***	0.079***	0.081***	0.075***	0.135***	0.128***
	(0.005)	(0.005)	(0.007)	(0.007)	(0.009)	(0.009)	(0.009)	(0.009)
Constant	0.130***	0.064***	0.170***	0.077***	0.300***	0.256***	0.182***	0.127***
	(0.015)	(0.015)	(0.023)	(0.026)	(0.046)	(0.047)	(0.030)	(0.031)
Occupation F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	45,549	45,549	13,737	13,737	11,568	11,568	20,244	20,244
Adjusted R^2	0.027	0.030	0.027	0.030	0.020	0.023	0.033	0.036

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

Table A7. Robustness checks with 3-digit industries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Full model	Interaction	High	High	Medium	Medium	Low	Low
Self-relatedness	0.217***	0.419***	0.223***	0.380***	0.130***	0.321***	0.274***	0.525***
	(0.009)	(0.024)	(0.017)	(0.042)	(0.018)	(0.045)	(0.015)	(0.040)
Cross-relatedness	0.293***	0.577***	0.344***	0.605***	0.186***	0.445***	0.321***	0.649***
	(0.024)	(0.037)	(0.048)	(0.081)	(0.044)	(0.068)	(0.034)	(0.056)
Interaction		-0.760***		-0.622***		-0.740***		-0.917***
		(0.085)		(0.161)		(0.164)		(0.140)
GDP per capita	-0.027***	-0.026***	-0.030***	-0.029***	-0.008	-0.008	-0.037***	-0.038***
	(0.003)	(0.003)	(0.005)	(0.005)	(0.007)	(0.007)	(0.006)	(0.006)
Population density	-0.322***	-0.281***	-1.094***	-0.882***	0.014	0.054	-0.250***	-0.232***
	(0.062)	(0.062)	(0.176)	(0.191)	(0.120)	(0.120)	(0.086)	(0.086)
Stock	-0.003	-0.002	-0.002	-0.003	-0.012	-0.009	-0.000	0.004
	(0.004)	(0.004)	(0.005)	(0.005)	(0.008)	(0.008)	(0.007)	(0.007)
Size	0.093***	0.086***	0.078***	0.072***	0.071***	0.062***	0.112***	0.104***
	(0.005)	(0.005)	(0.007)	(0.007)	(0.009)	(0.009)	(0.009)	(0.009)
Constant	0.047***	-0.025	0.075***	0.000	0.238***	0.176***	0.098***	0.024
	(0.016)	(0.017)	(0.026)	(0.032)	(0.048)	(0.049)	(0.031)	(0.032)
Occupation F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	45,549	45,549	13,737	13,737	11,568	11,568	20,244	20,244
Adjusted R^2	0.030	0.031	0.031	0.032	0.021	0.022	0.036	0.038

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

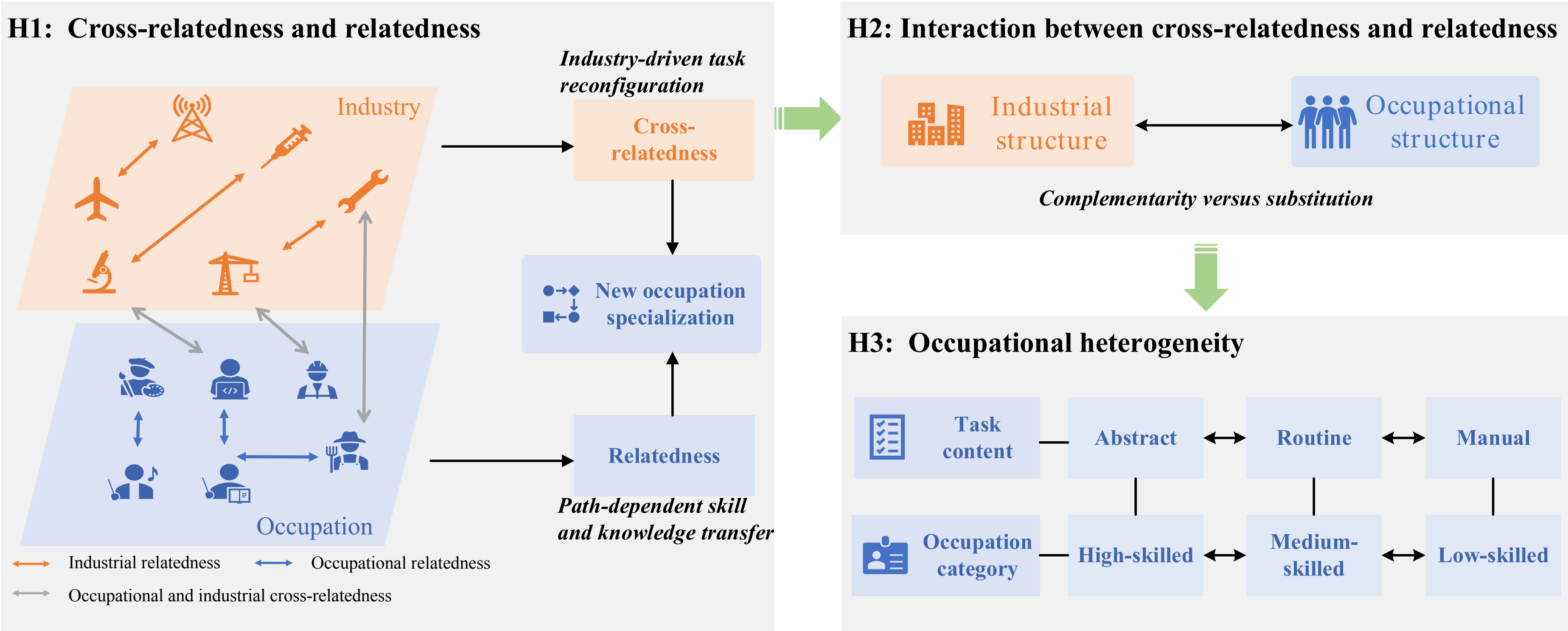

Figure 1. Theoretical framework of relatedness and cross-relatedness in co-shaping regional occupational evolution across occupations of different skill levels

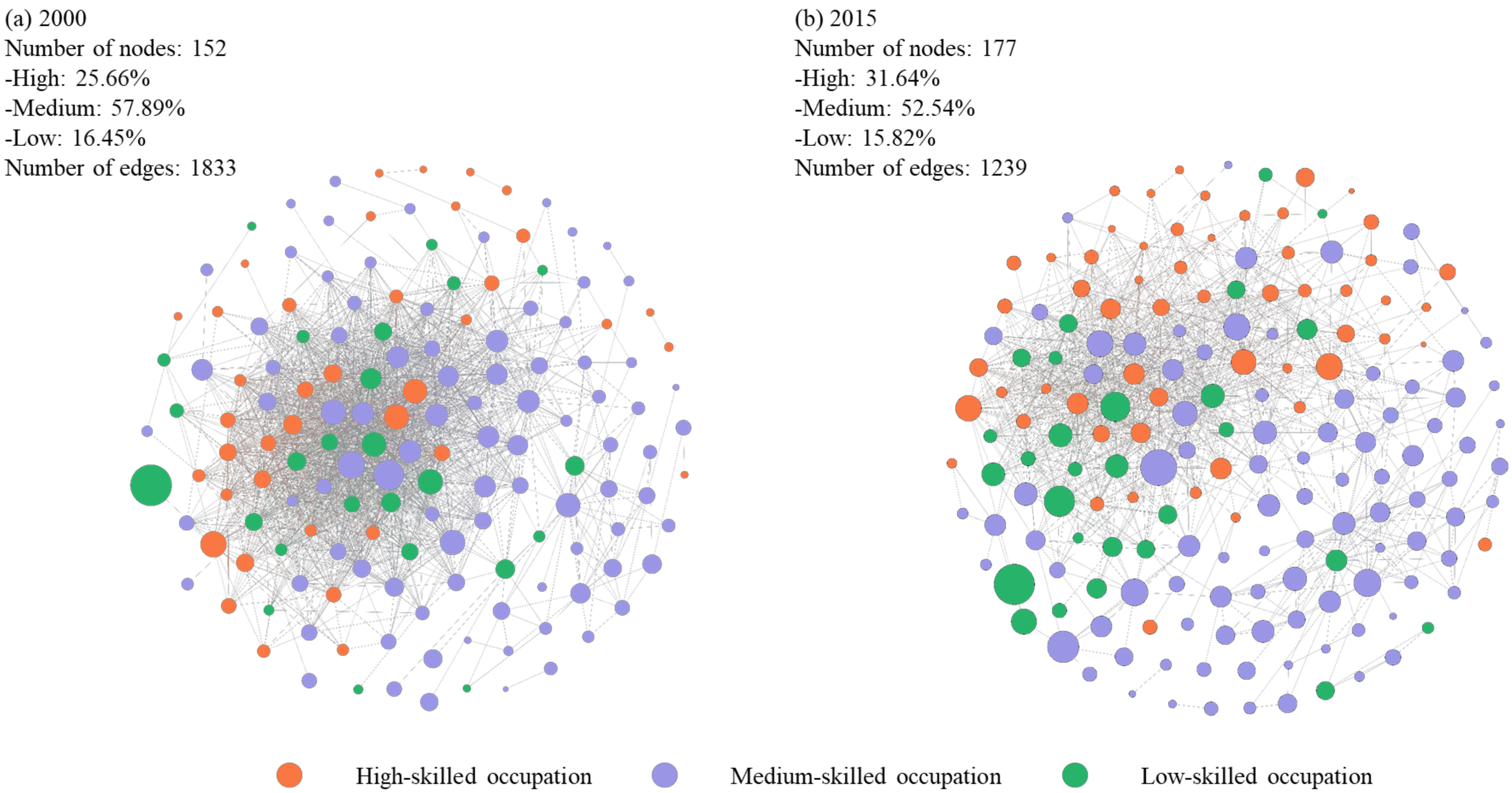

Figure 2. Occupation space, 2000 and 2015

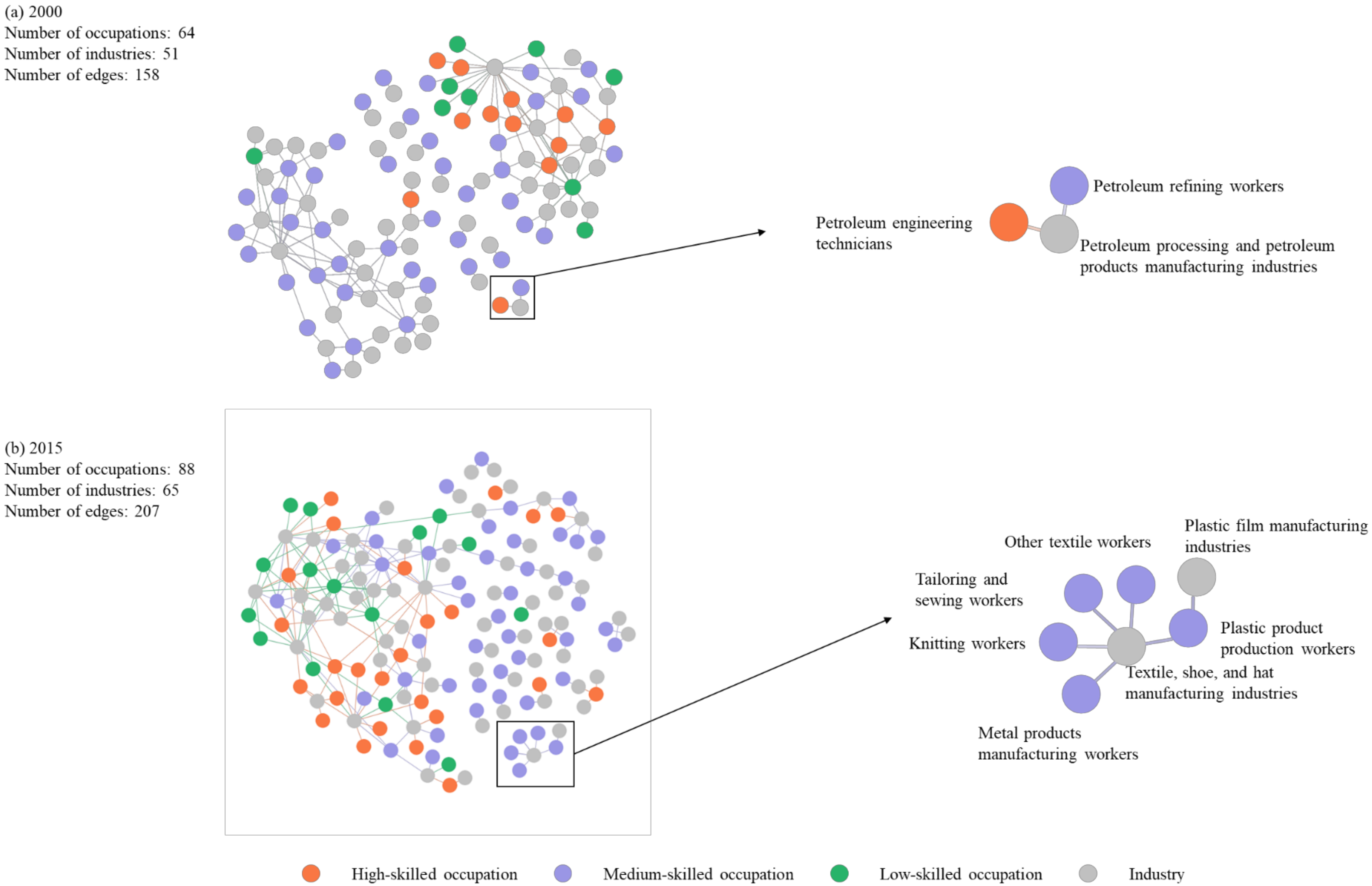

Figure 3. Industry-occupation cross-space, 2000 and 2015

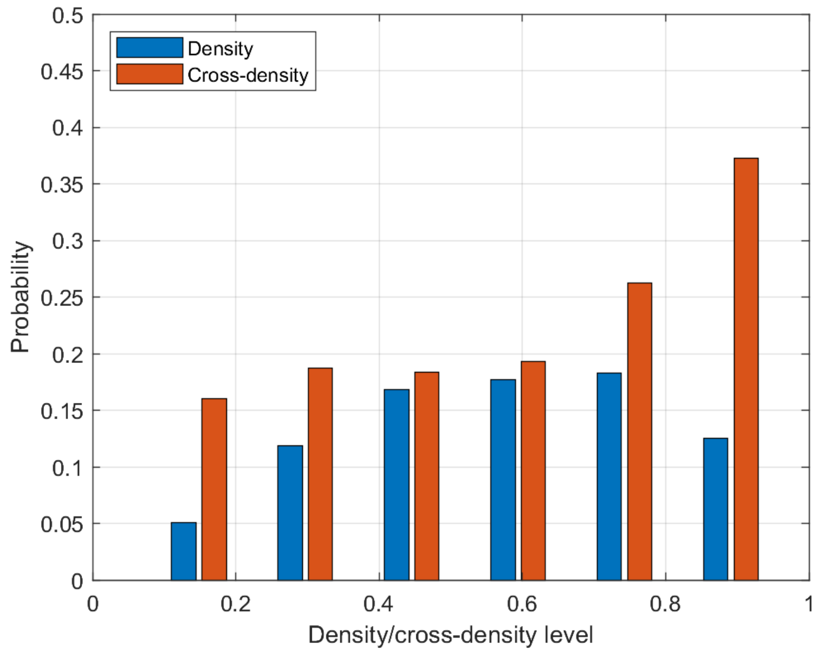

Figure 4. Probability of new occupation entry in a region as a function of relatedness density and cross-relatedness density

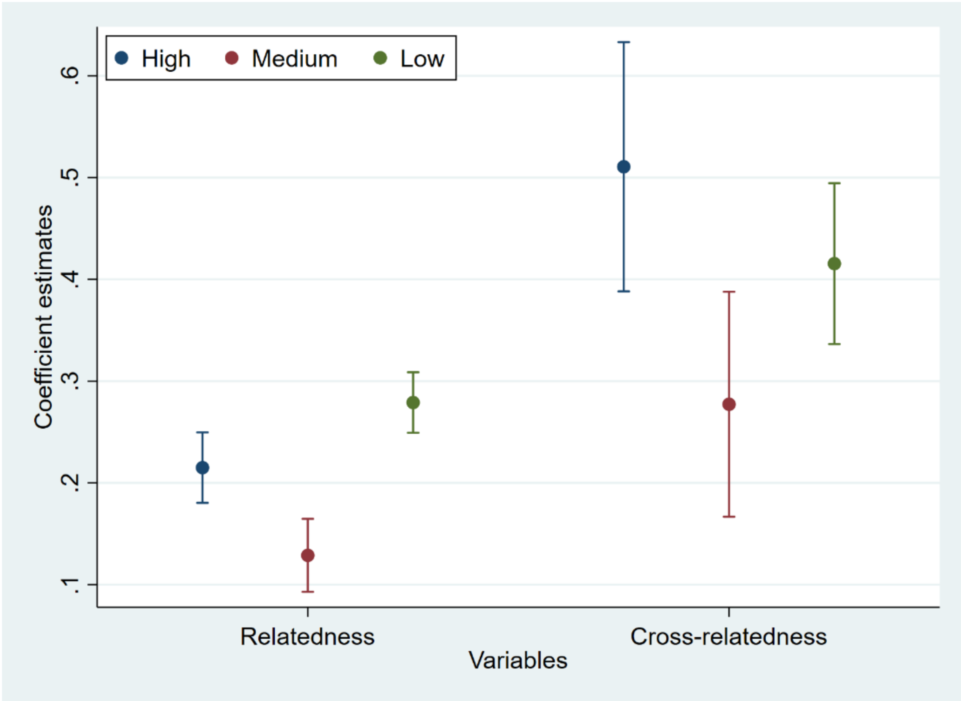

Figure 5. Regression coefficients with confidence intervals by occupation skill level

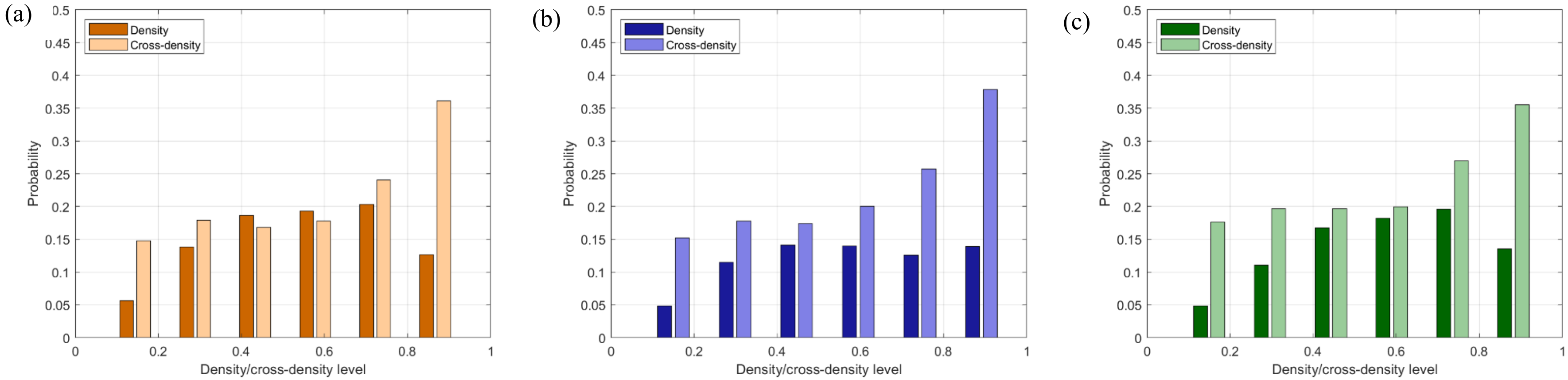

Figure B1. Probability of new occupation entry in a region as a function of relatedness density and cross-relatedness density: (a) high-skilled, (b) medium-skilled, and (c) low-skilled occupations

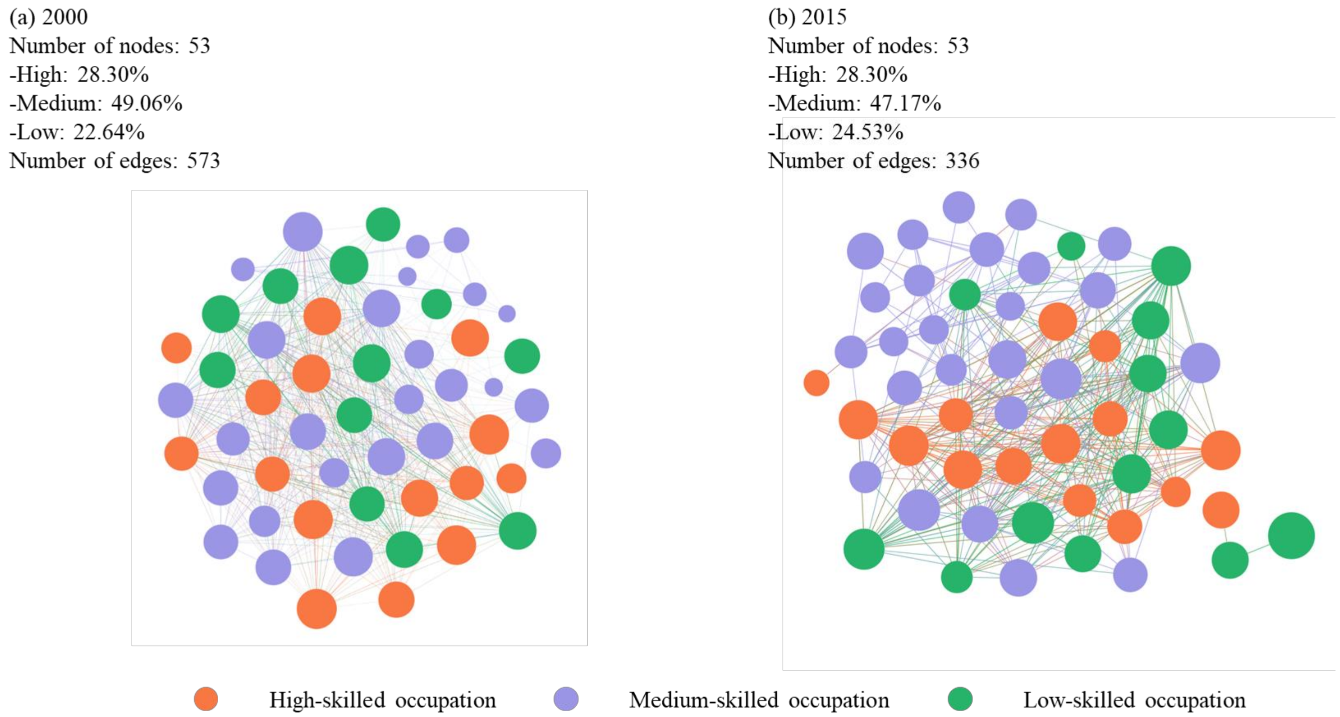

Figure B2. Occupation space constructed from 2-digit occupations in 2000 and 2015

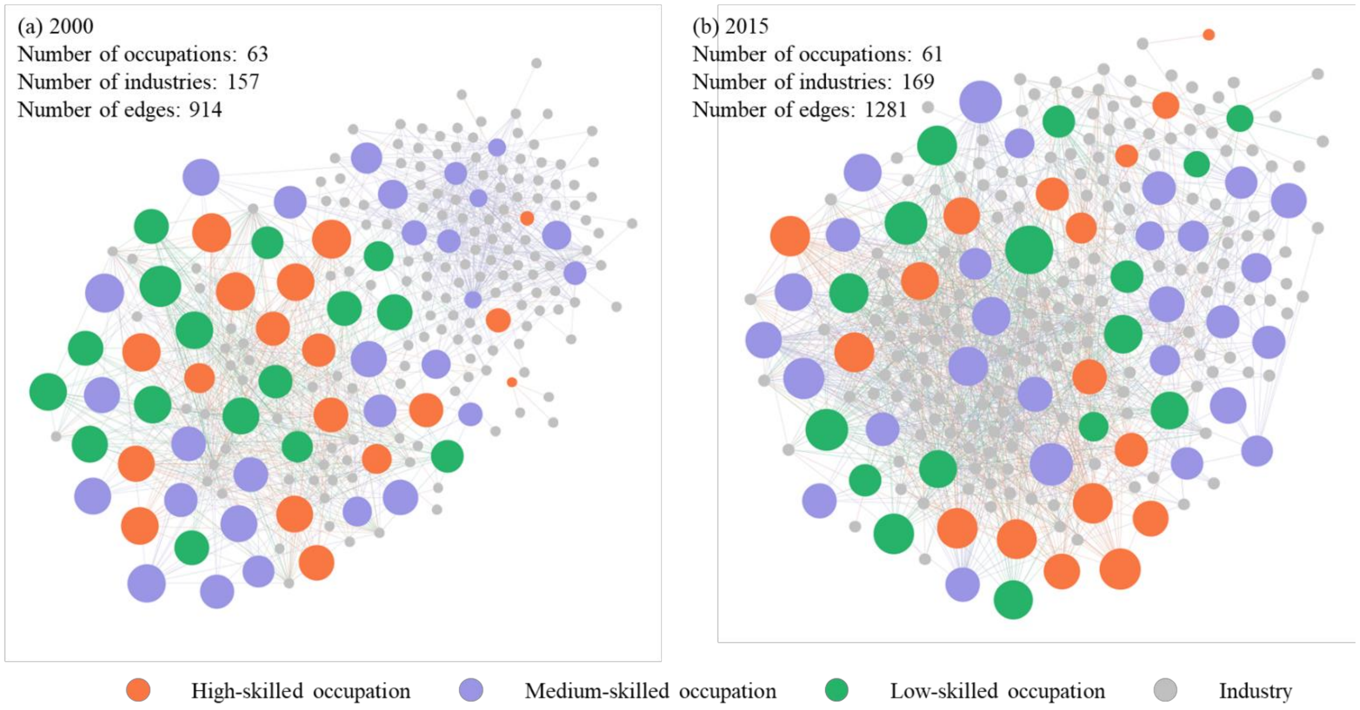

Figure B3. Industry-occupation cross-space constructed from 2-digit occupations in 2000 and 2015