Regional static diversification and relatedness between industries

Østergaard, Christian Richter; Holm, Jacob Rubæk

Published in:
Industrial and Corporate Change

DOI (link to publication from Publisher):
10.1093/icc/dtac045

Publication date:
2023

Document Version
Accepted author manuscript, peer reviewed version

Regional static diversification and relatedness between industries

Christian Richter Østergaard

Aalborg University Business School, Aalborg University,

Fibigerstræde 11, 9220 Aalborg Ø, Denmark

Email: cro@business.aau.dk

ORCID 0000-0003-1782-202X

Jacob Rubæk Holm (corresponding)

Aalborg University Business School, Aalborg University,

Fibigerstræde 11, 9220 Aalborg Ø, Denmark

Email: jrh@business.aau.dk

ORCID 0000-0003-1110-1918

This article has been accepted for publication in Industrial and Corporate Change published by Oxford University Press.

https://doi.org/10.1093/icc/dtac045

Østergaard, C. R., & Holm, J. R. (2022). Regional static diversification and relatedness between industries. *Industrial and Corporate Change*.

Abstract

Relatedness has often been shown to have a central role in regional diversification. Knowledge flows between related industries are an important source of innovation leading to industry growth and to the establishment of new industries. Earlier studies have focused on dynamic diversification into new industries but in this study, we emphasize static diversification, that is, changes in balance between existing industries in a region. We use linked employer-employee data from Denmark and construct a range of measures at the level of regional industries. The role of relatedness in static diversification 2008-2013 is then analyzed based on industry characteristics in 2008. We find that relatedness plays a more important role in regions that are either spatially peripheral or economically non-peripheral, while relatedness has no relationship to static diversification on average. In addition to the empirical result, we contribute by comparing indices of relatedness, suggesting an alternative employment weighted index and conceptualizing relatedness as itself a disparity dimension of diversity. We finally ask how static and dynamic diversification affect disparity. We find that human capital intensity plays an important role in the link between diversification and the evolution of disparity.

Keywords

Regional diversification; Spatial periphery; Economic periphery; Human capital; Agglomeration externalities; Disparity;

1. Introduction

The process of regional diversification of economic activities has been subject to increased research interest in recent years, with a particular focus on dynamic diversification, that is, the emergence of new-to-the-region economic activities (Hidalgo et al., 2007; Neffke et al., 2011; Boschma et al., 2015; Essletzbichler, 2015; Boschma, 2017). Prior studies find that the process of regional diversification mainly consists of diversification into products, industries, or activities, which are related to the current products, industries, or activities in the region. The main argument is that new activities draw on and benefit from knowledge flows stemming from the existing portfolio of activities in the region. In addition, Neffke et al. (2011) show that not only are industries related to the existing structure more likely to enter into a region, but unrelated industries are also more likely to exit. Thus, the regional portfolio of related activities affects the possibilities for new activities.

These prior studies tend to overlook powerful forces that determine which industries grow in the global economy, and where they grow. This growth is important for dynamic diversification but also for the relative balance between the existing regional economic activities, i.e. for static diversification of regional industries. All regions are exposed to these forces. In advanced economies these include deindustrialization, job polarization, urbanization, globalization, demographic transition, increasing human capital (HC) intensity and climate change (Goos et al., 2014; Bernard et al., 2016; International Monetary Fund, 2018). These processes may strengthen or mitigate the tendency for diversification into related activities emphasized in prior studies and must be kept in mind when analyzing regional diversification. In addition to overlooking these forces, the prior literature largely neglects specific regional characteristics such as spatial and economic periphery. For example, a spatially peripheral region characterized by high relatedness among industries may not experience diversification because of the trend of urbanization. This means that the presence of related industries does not automatically lead to diversification. However, this also means that relatedness among industries can mitigate the negative effects of urbanization and create resilience to shocks (Xiao et al., 2018).

The purpose of this paper is to analyze the relative importance of relatedness in regional static diversification, with specific emphasis on the moderating effects of HC intensity and of locating in the spatial or economic periphery. The paper draws on a detailed matched employer-employee dataset for Denmark, which allows for the in-depth analysis of the processes affecting diversification. This analysis has two steps. First, we analyze the importance of relatedness in regional static diversification in a regression analysis. Secondly, we analyze how static diversification shapes the evolution of relatedness.

This paper contributes to the literature concerning regional diversification in several ways. First, it shows that there is no general positive effect of relatedness on static diversification. The central role of relatedness proposed by earlier studies is mainly found in the economically non-peripheral regions and in the spatially peripheral regions. One implication of this result is that relatedness in a spatially peripheral region could make the region more resilient to the negative effects of urbanization. Second, the paper demonstrates that relatedness is not more important for HC intensive industries, which is contrary to the expectations developed in the paper. However, HC intensive industries generally have both high growth and high relatedness, which implies that HC intensive industries benefit from relatedness in ways that do not necessarily lead to growth.

In addition to these empirical contributions the paper contributes conceptually by integrating relatedness conceptually within diversity as the disparity dimension of diversity in addition to the two other dimensions: variety and balance. The paper also contributes methodologically by suggesting an employment weighted index of relatedness.

2. Theories of regional diversification

2.1. Static diversification versus dynamic diversification

A more diversified region is a region where employment or output is more spread out over economic activities. Therefore, diversification is the process of spreading out economic activities across industries. According to Stirling (2007), diversity has three properties: 1) variety, 2) balance and 3) disparity.

1) Variety is the number of categories in a system. For a regional economy variety is the number of different economic activities as captured by for example industry codes. Increasing variety in the local economy is dynamic diversification, since it introduces new-to-the-region economic activities (Neffke et al. 2018).

2) Balance is the balance between the different categories. For a regional economy balance is the relative size of the different economic activities as captured by for example employment in industries. For a given number of different industries diversity is higher when employment is less concentrated. Increasing balance in the local economy is static diversification, since it spreads out the overall employment more evenly on different economic activities (Siegel et al., 1995; Neffke et al. 2018).

3) Disparity is the degree of difference between categories. For a regional economy disparity is a measure of how different the various economic activities of the region are relative to each other. Disparity between economic activities could be measured along different dimensions. For example, how different is the high-tech industry from agriculture or low-tech services. Relatedness is arguably an empirical measure of disparity since it captures the flow of skilled people and knowledge among industries. If two industries are more related, then they have less disparity.

Studies focusing on regional diversification in the dynamic sense focus on variety: the emergence of an industry, a product or a patent classification that is new-to-the-region (Hidalgo et al., 2018). This dynamic diversification will add new categories of activities to a region and increase the diversity, but it might only change balance of categories slightly, since these new activities might only employ few people. Thus, the regional industrial structure is largely unchanged. Focusing on dynamic diversification is intuitively appealing as diversification in the long run must entail the emergence of genuinely new activities, but these activities are not easily captured by the abovementioned methods as genuinely new activities by definition have no place in the applied classification system. Instead, studies of the emergence of new-to-region industries primarily capture the diffusion process of economic activities across regions. In the short term, regional static diversification is affected by such diffusion, but activities new-to-the region often have little impact in terms of employment.

The static diversity of a regional economy changes as industries grow and decline. The changes in regional static diversification can also be seen as dynamic at the firm level since firm entry and exit within existing industries also represent regional diversification of activities, as do significant changes in the size of incumbent firms (Klepper and Thompson, 2006). These changes affect the balance property of diversity. Thus, an analysis of changes in regional diversification with emphasis on static diversification is a study of relative growth and decline of industries.

In recent contributions on regional industrial dynamics, the models of regional diversification have emphasized how the existing industry structure affects the future regional diversification pattern. The presence of related industries in a region generates knowledge flows taking the form of both spillovers as well as labor mobility, which support further growth of these industries as well as increasing the likelihood of entry of related industries (see e.g. Hidalgo et al., 2007; Eriksson et al., 2008; Hausmann and Hidalgo, 2011; Neffke et al., 2011; Hartog et al., 2012; Boschma et al., 2015; Cortinovis and van Oort, 2015; Delgado et al., 2015; Essletzbichler, 2015; Caragliu et al., 2016; Boschma, 2017; Xiao et al., 2018). These studies of regional diversification show that industries benefit from knowledge flows stemming from the presence of related industries in the region. Therefore, the current regional industry structure moderates diversification by providing both opportunities and limitations for growth (Boschma, 2017). This leads to the following hypothesis:

Hypothesis H1a: Relatedness shapes the direction of static diversification. This means that higher relatedness of a regional industry towards the local industry structure is positively associated with growth of the industry.

The consequences for static diversification of the relationship in hypothesis H1a depends on whether the relationship is strongest for relatively small industries or for large industries. If the relationship is strongest for large industries, then the result will be increased concentration rather than increased static diversity. Our analysis incorporates this concern.

2.2 Other factors affecting regional diversification

Research on regional diversification has been highly empirically driven (Henning et al., 2013).Different data and methods are employed to operationalize the pattern of relatedness using names such as the product space (Hidalgo et al., 2007), the tree of industrial life (Andersen, 2003), the industry space or R-matrix (Neffke et al., 2011; Neffke et al., 2017), and the technology/knowledge space (Boschma et al., 2015; Kogler et al., 2017). The mechanisms identified as driving the regional diversification include knowledge flows through spillovers, spinoffs and local labor mobility. In addition, studies have shown that entrepreneurship is an important mechanism for regional diversification (Klepper, 2010; Noseleit, 2015).

Studies comparing core regions and peripheral regions often find that the diversification processes are different (Isaksen, 2015) and that the ability to leverage relatedness varies between core and peripheral regions (Kuusk, 2021). Knowledge from outside the region (Isaksen 2015) and knowledge from related industries within the region (Kuusk, 2021) are particularly important in peripheral regions. It is important to note that periphery is not only geographical periphery. It also includes economic peripheral regions without the growth, resources, and dynamism of the core regions (Rodríguez-Pose, 2018).

Core regions have large pools of highly educated labor that increase the absorptive capacity of the region (Cohen and Levinthal, 1990), and there is an increasing focus on the importance of access to non-local knowledge in order to achieve regional diversification (Neffke et al., 2018). Mechanisms facilitating knowledge flows are therefore important for diversification. In economic geography such mechanisms are the Marshall-Arrow-Romer (MAR) externalities: specialized suppliers, thick labor markets and informal interaction (Delgado et al., 2015). Caragliu et al. (2016) finds that MAR externalities are more important for industry employment growth in less population dense European regions than for the denser regions. In addition, Eriksson et al. (2008) finds that MAR externalities increase interregional job mobility in small peripheral regions, thus mitigating some of the disadvantages of thin labor markets in these regions. MAR externalities are conceptually close to relatedness as both models include a focus on knowledge flows through labor flows and informal interaction. However, while MAR externalities arise through scale and specialization of an industry, relatedness emphasizes how industries are linked through flows of people between employers both within and between industries. Thus, the presence of related industries could be a source of knowledge flows in a similar fashion as MAR externalities. However, relatedness is also different and broader than MAR externalities, since a flow of people with related skills from different industries could be interpreted as Jacobs externalities stemming from regional industry diversity (Jacobs, 1969). This leads to our second hypothesis:

Hypothesis H1b: Relatedness is more important for static diversification in peripheral regions. This means that the marginal effect of relatedness on growth is larger for industries located in such peripheries.

The importance of relatedness for static diversification entails that the growth of more HC intensive industries is relatively more dependent on the potential for knowledge flows through labor mobility in the region. HC intensive firms have a higher absorptive capacity than less HC intensive firms (Cohen and Levinthal, 1990), which makes them better at identifying and integrating the new knowledge from different industries flowing through labor mobility (Breschi and Lissoni, 2001). The integration of more distant knowledge could provide more opportunities for innovation processes (Boschma et al., 2009). Similarly, HC intensive firms are more dependent on high levels of HC for growth, while less HC intensive firms are less dependent on the knowledge transfer though labor mobility of low skilled employees. Boschma et al. (2009) finds a positive effect on firm productivity on the inflow of highly educated employees with related skills from related industries. This leads to the third hypothesis:

Hypothesis H1c: Relatedness is more important for static diversification for HC intensive industries. This means that the marginal effect of relatedness on growth is larger for HC intensive industries.

2.3. Feedbacks from static diversification to relatedness

Above, diversity was presented as having the three dimensions: variety, balance, and disparity, which correspond to dynamic diversity, static diversity, and relatedness respectively in empirical research. Earlier studies and the present study have focused on static and dynamic diversity using relatedness as an explanatory variable. Our final set of hypotheses concern changes in relatedness itself. In this analysis we treat relatedness as a population level concept that evolves with the population. From an evolutionary perspective, any trait that is positively associated with growth of an industry in a region will become more frequent in the region over time (Metcalfe, 1998). This leads to hypothesis H2a.

Hypothesis H2a: As relatedness has a positive effect for growth (hypothesis H1a), average relatedness will increase over time.

However, an industry level trait may be positively associated with growth of industries and yet become less frequent in the region if this trait also correlates with a secondary trait that is negatively associated with growth (Metcalfe, 1998; Holm et al., 2016). Such secondary traits potentially include the trends mentioned earlier. In particular, many jobs are created in HC intensive industries and urban regions meaning that the trends of increasing HC intensity and urbanization affect the evolution of relatedness (International Monetary Fund, 2018; Kemeny and Storper, 2020). This leads to hypotheses H2b and H2c building on the relations suggested by hypotheses H1b and H1c respectively:

Hypothesis H2b: As relatedness is more important in peripheral regions (hypothesis H1b), the trend of urbanization entails that average relatedness decreases over time.

Hypothesis H2c: As relatedness is more important for HC intensive industries (hypothesis H1c), increasing HC intensity entails that average relatedness increases over time.

Our six hypotheses are interrelated in the sense that hypotheses H1b and H1c qualifies hypothesis H1a on the relationship between relatedness and static diversification, and hypotheses H2a, H2b, and H2c concern changes over time in relatedness as derived from hypotheses H1a, H1b, and H1c, respectively.

3. Data and variables

Regional static diversification is studied by analyzing the growth of regional industries in Denmark from 2008 to 2013. Each observation of a regional industry is thus an industry-region combination that we observe in both 2008 and 2013. The data used are linked employer-employee census data from Statistics Denmark containing highly detailed information about the employees, the employment relations, the workplaces, and the firms. Regional industries are described by aggregating all the employees at workplaces in the region in that industry. We use the four-digit Nomenclature of Economic Activities (NACE) codes to delimit the industries and exclude health, education, and public sector administration. In order to comply with confidentiality requirements, all regional industries with less than four plants or less than ten full-time equivalent (FTE) employees are excluded. This removes about a quarter of all regional industries accounting for only one percent of employment in the sample. Thus, an industry only emerges in or disappears from a region when it consists of at least four plants and ten FTE employees. We also exclude the top and bottom five percent of regional industries by employment growth 2008-2013 to remove cases of rapid growth and decline during the financial crisis.

The Danish economy consists of very different regions, each with its own path of development. We distinguish between seven regions defined at the NUTS3 level, with the exception of the NUTS2 regions of Copenhagen and Zealand, since the corresponding NUTS3 regions are highly economically integrated.[i] A map of the resulting seven regions can be seen in the appendix.

The data contain 2311 observations of regional industries that exist in both 2008 and 2013 across the seven regions with approximately 1.4 million employees in 2008. These 2311 regional industries are distributed across 430 different four-digit industry codes. The dependent variable is the growth of these industries over the period, cf. Equation 1 where $x_{ir,t}$ denotes the regional employment of industry i in region r in year t.

$$Growth_{ir} = \frac{x_{ir,2013}}{x_{ir,2008}} \tag{1}$$

3.1. Relatedness

The concept of relatedness can be operationalized at three different levels: I) a distance or disparity metric quantifying relatedness between two industries, II) the average of the distance metric (I) between one industry and all other industries in the region. This describes the fit to the current regional industrial structure of this industry and is referred to as relatedness in this paper. III) the regional average of relatedness (II) measured over all industries in the region

For the distance metric (I) we apply the method used by Neffke et al. (2017) to compute a "skill relatedness index", which measures the distance between a pair of industries based on labor flows, where labor flows are people moving between industries when switching jobs. If labor flows exceed expected labor flows from industry j to industry i, then i is relatively related to j. The fact that labor flows more frequently between the industries than should be expected arguably reflects the fact that the skills acquired in industry j are applicable in industry i; hence, i is skill related to j, although in principle the opposite need not be true. The skill relatedness index has values ranging from 1 to -1, where 0 means that the expected flows and the observed flows are equal, while positive values mean that the observed flows are higher than the expected flows. The index is computed by pooling all the moves in consecutive years between industries in Denmark from 2008–2013, which results in a relatedness matrix. This matrix shows the distance between all the possible pairs of industries. The diagonal of the matrix is not defined by the methodology, but we place "1s" along the diagonal, reflecting the fact that the relatedness between an industry and itself is the maximum of the index. In order to emphasize that we do not impose the restrictions of symmetric relatedness, we refer to the variable as the skill *inflow* relatedness of i to j, SIR_{ij}.

$$SIR_{ij} = \frac{r_{ij}-1}{r_{ij}+1}, \text{ where}$$

$$r_{ij} = \frac{F_{ij}F_{..}}{F_{i.}F_{.j}} \qquad (2)$$

In Equation 2, F_{ij} is the number of people who leave a job in industry j and find a job in industry i, $F_{i.}$ and $F_{.j}$ are the total number of people leaving and entering the two industries, respectively, and $F_{..}$ is the total number of people moving between industries in the economy (Neffke et al., 2017).

As we pool the data, the distance metric SIR_{ij} remains constant across regions and over time, but at any point in time for any given region, industry i may be more or less related to the regional industrial structure depending on which other industries are present within the region. In Neffke et al. (2011) and in Essletzbichler (2015) this is called standardized closeness and is computed by first constructing a binary variable which is 1 for values of SIR_{ij} above the 90$^{\text{th}}$ percentile and 0 otherwise. Standardized closeness is then defined as the average of this new binary variable.

A problem with standardized closeness is that the measure weighs all industries equally, independent of their size. We propose to add weights for industry employment to standardized closeness to reflect that a larger industry represents a larger reservoir of skills. Let s_{irt} denote the regional employment share of industry i, in region r at time t, and the employment weighted average can be computed as in Equation 3. This is our preferred measure, and it is referred to as employment weighted relatedness (ER) to emphasize that contrary to standardized closeness, this measure is weighted by employment.

$$ER_{irt} = \sum_{i \in RPF_{rt}} I\left(SIR_{ij} > p90_t\right) s_{irt} \tag{3}$$

More formally, RPF_{rt} is the "regional portfolio," that is, the set of industries that are present in *r*, while $I(\cdot)$ is an indicator function that takes the value of 1 if the expression is true and 0 otherwise. $p90_t$ is the threshold that distinguishes related industries from other industries. This is defined as the 90th percentile of the cumulative distribution of SIR_{ij}. Cases where $SIR_{ij} = 1$ (i.e., cases where $i = j$) were excluded when determining the 90th percentile. For comparison, Equation 4 shows this formula for standardized closeness.

$$SC_{irt} = \sum_{j \in RPF_{rt}} I\left(SIR_{ij} > p90_t\right)/N_{rt} \tag{4}$$

Whereas a value of standardized closeness of, for example, 0.1 means that 10 percent of the industries present in region r in year t are related to industry i, a value of employment weighted relatedness of 0.1 means that 10 percent of jobs in region r in year t are related to industry i. An intermediate approach was suggested by Fitjar and Timmermans (2017). They suggest a regional skill relatedness index, which when rewritten to an industry level measure becomes a broader index of relatedness. It uses the square root of employment shares as weight so that larger industries have larger weight, though the weight is less than proportional to the size of the industry. In addition, the index considers labor flows in either direction equally important. This means that industry i benefits from the presence of industry j if labor moves disproportionality from j to i, but i also benefits if labor moves disproportionately from i to j. The index thus captures relatedness in a broader sense than just through labor flows. The formula for the broad relatedness index is shown in Equation 5.

$$BR_{irt} = \sum_{j \in RP\ \ rt} 0.5\left[I\left(SIR_{ij} > p90_t\right) + I\left(SIR_{ji} > p90_t\right)\right]\sqrt{s_{jrt}}/\sum_{k \in RP\ \ rt}\sqrt{s_{krt}} \tag{5}$$

In the analyses presented in this paper we focus on the employment weighted relatedness index, ER_{irt}, but also present results for standardized closeness, SC_{irt}, and for the broad relatedness index, BR_{irt}, for comparison.

3.2. Other variables

In addition to the index for relatedness (the main variable for hypotheses H1a and H2a) we include variables for periphery (for hypotheses H1b and H2b) and for HC intensity (for hypotheses H1c and H2c), and a number of controls for factors that affect the development of regional industries over time. For these variables we are only interested in the initial value in 2008 and hence we do not include a t subscript. Table 1 provides an overview of the variables included in the analysis.

Effect	Measure	Variable name
Relatedness	Relatedness indices of equations 3-5	Relatedness (ER, SC, BR)
Spatial periphery	Share of jobs in peripheral municipalities	Spatial
Economic periphery	Danish Kroner of active industrial policy per job	Economic
Human capital intensity	Share of employees with tertiary education	HC
Job polarization	Share of jobs in routine occupations	Routine
Absolute size of regional industry	Log of regional industry employment	Size
Entrepreneurship	Share of employees in plants founded less than five years ago	Start-ups
Relative regional specialization (Location Quotient)	Industry share of regional employment relative to industry share of national employment	RRS

Table 1. Overview of the variables included in the analysis

Periphery: We employ two different categories of periphery: The spatial peripheries are locations removed from urban centers, while the economic peripheries are locations where economic development is comparatively low. Following the Danish Economic Councils (DORS 2015) we define spatially peripheral municipalities as municipalities where there is no city with at least 45,000 inhabitants and where the median commute to such a city exceeds 30 minutes. The variable $Spatial_{ri}$ is the share of jobs for industry i in region r where the address for the workplace is within a spatially peripheral municipality.

The variable for economic peripheries is based on the observation that municipalities with weak job growth tend to spend funds on various forms of industrial policy and be the recipients of EU and national structural funds. Statistics Denmark provides information about the municipalities' yearly spending on active industrial policy.[ii] This is divided into seven categories, namely miscellaneous incomes and expenses, growth fora, tourism, human resource development, innovation and new technology, business services and entrepreneurship, and the development of peripheral and rural areas. There is some indication that municipalities have discretion in terms of labelling the expenses and there are significant variations from one year to the next, so we use the three-year average of the sum of all seven categories.[iii] $Economic_{ri}$ is the average expenditure per job in industry i in region r. As the municipalities within the region differ in relation to their expenditure per job, the industries will have different averages depending on whether they tend to be located in high expenditure municipalities or not. Maps showing both the spatial and economic periphery can be found in the appendix.

Human capital intensity: The HC intensity variable is as the share of employees with a tertiary level education, HC_{ri}.

Control variables: We control for occupational shifts away from routine jobs, agglomeration effects, and entrepreneurship. We also create industry fixed effects following the Eurostat definitions for knowledge-intensive services[iv] and high-tech classification of manufacturing[v]. This entails that we also have two residual classes: "Other services" and "Primary, utilities and construction". Fixed effects for regions are included to capture unobserved regional specificities.

A large literature has documented systematic shifts in the occupational structure of Denmark and other developed countries. In particular, jobs within three broad categories are generally found to be decreasing in frequency, arguably because a large share of the tasks in these jobs are relatively routine (Goos et al.

2014). Following Goos et al. (2014) these three broad categories and their one-digit International Standard of Classification of Occupations (ISCO) are "4: Clerical Support Workers," "7: Craft and Related Trades Workers," and "8: Plant and Machine Operators and Assemblers". The variable $Routine_{ri}$ is the share of jobs in the regional industry with an ISCO classification as a routine job.

To account for agglomeration effects, we create two variables: $Size_{ri}$ is the log of employment in the regional industry and RRS_{ri} is regional relative specialization measured by the location quotient for the regional industry. Entrepreneurship is measured with a variable based on workplace age so that entrepreneurship becomes synonymous with young plants. The variable $Start-up_{ri}$ is the share of employment in plants founded less than five years ago.

3.3. Descriptive statistics

The database dates back to 1980, but there are breaks in both the occupation and industry codes around 2007–2008, as well as a change to the administrative boundaries in Denmark in 2007. Hence, we only use data from 2008 until 2013, which is the most recent comparable data at the time of this study. The chosen period coincides with the global financial crisis, which caused a recession in 2008–2009, followed by a recovery from 2010–2014 at the aggregate level. At the level of observation, however, the different regional industries experienced heterogenous paths of development. Some recovered faster than others, and yet others were still in decline at the end of our dataset in 2013. In the dataset for our analysis the spread of yearly employment growth is described in Table 2.

Group	2008	2009	2010	2011	2012	2013
			Growth relative to previous year			
Top 5 percent	.	1.26	1.05	1.77	1.61	1.32
Middle 90 percent	.	0.95	0.96	1.05	1.01	1.00
Bottom 5 percent	.	0.70	0.72	0.88	0.86	0.78
			Relatedness (ER)			
Top 5 percent	0.0542	0.0533	0.0527	0.0537	0.0549	0.0552
Middle 90 percent	0.0645	0.0616	0.0603	0.0614	0.0618	0.0616
Bottom 5 percent	0.0684	0.0632	0.0597	0.0615	0.0620	0.0624

Note: Data grouped by cumulative distribution of total growth 2008-2013.

Table 2. The evolution and spread of employment growth and relatedness

The top and bottom five percent in Table 2 are computed after excluding the initial top and bottom five percent of most extreme observations. Table 2 shows yearly employment growth for three different groups in our data: The five percent of regional industries with highest total employment growth 2008-2013, the five percent of regional industries with the lowest total employment growth 2008-2013, and the remaining 90 percent of regional industries. For the middle 90 percent the general business cycle is visible as there is decline until 2010 and then a recovery. But the five percent with highest growth over the period grow continuously while the five percent with lowest growth have continuous decline.

Table 2 also shows average relatedness for the three groups of regional industries over the period. It can be seen in Table 2 that the relatedness changes slowly over time. Average relatedness declined until 2010 in all three groups and then stabilized resulting in a small net decrease over the period. It is interesting to see that the high growth industries on average have low relatedness while the low growth industries have higher than average relatedness. This apparently negative relationship between relatedness and growth is likely to be caused by both relatedness and growth being lower for old manufacturing industries, and it should thus disappear when we control for these factors in the regression analysis.

The economic crisis starting in 2008 could moderate the effect of relatedness on diversification, but the crisis is difficult to include as there are no obvious region or industry specific measures of the impact of the crisis. The main effect is likely to have been the creation of a large supply of laid-off labor. However, the increased supply of labor would also lead to an increased demand, since it allows related industries to grow (Morkutė et al., 2017).

Table 3 shows correlations, standard deviations and means for the variables used in the regression analysis. The variables are normalized for the regressions and the correlations, while the standard deviations and means are computed prior to normalization. For all independent variables the 2008 values are used in the regression.

	Growth	Spatial	Economic	Routine	HC	Size	Start-up	RRS	ER
Spatial	-0.001								
Economic	-0.017	0.499†							
Routine	-0.208†	0.056†	0.132†						
HC	0.107†	-0.217†	-0.287†	-0.281†					
Size	-0.206†	-0.126†	-0.234†	0.139†	0.100†				
Start-up	0.193†	-0.040	-0.036	-0.284†	0.058†	-0.156†			
RRS	-0.113†	0.066†	0.072†	0.072†	-0.035	0.158†	-0.157†		
ER	-0.121†	0.019	0.038	0.389†	-0.007	0.152†	-0.214†	0.281†	
Mean	0.958	0.289	324.6	0.298	0.130	5.448	0.239	1.252	0.064
SD	0.444	0.293	146.9	0.203	0.132	1.325	0.210	1.201	0.034

SD: Standard deviation. †: significant at 5%

Table 3. Descriptive statistics

The dependent variable, $Growth$, has significant correlation with several of the independent variables. Notably, $Growth$ has negative correlation with the relatedness index (ER), as also discussed in relation to Table 1. The mean of the dependent variable is 0.958 meaning that the average regional industry declined by 4.2 percent over the period.

The positive correlation between $Economic$ and $Spatial$ (economic and spatial periphery) was to be expected but interestingly, neither has a significant correlation with $Growth$.

$Routine$ has negative correlation with $Growth$ while HC has a positive correlation with $Growth$. This means that HC intensive industries have higher growth and that industries with a large share of routine jobs have low growth. As expected, industries consisting of young plants have higher growth, while the variables for agglomeration effects have the opposite correlations compared to expectations: small industries without relative specialization tend to grow.

3.4. Empirical models and decomposition

Our analysis proceeds in two steps. First, we use linear regression analysis to analyze the relationship between relatedness and regional static diversification (hypotheses H1a, H1b, and H1c).

The regression analysis estimates the model in Equation 6. m is used as an index for the 2311 observations of regional industries. β is a vector of parameters and α_r is the parameter for region r. Z_m is a vector of the variables for spatial and economic periphery, HC intensity, the control variables for routine jobs, size, specialization and entrepreneurship as well as the industry fixed effects. Equation 6 shows that the variables in Z_m enter the model both as linear terms and as interactions with ER_m. The fixed effects for regions are outside the parenthesis as they are not interacted with ER_m.

$$Growth_m = \beta(1 + ER_m)(1 + Z_m) + \sum_r \alpha_r Region_{r,m} + \epsilon_m \qquad (6)$$

We use weighted effects coding for the industry and region fixed effects so that the estimated direct effect of relatedness is the average effect and not the effect for the reference category (Nieuwenhuis et al., 2017). The error term ϵ_m is clustered at the industry level. All variables are standardized to mean zero and standard deviation one.

The method for the second step analyzing hypotheses H2a, H2b and H2c entails decomposing the change in average employment weighted relatedness across all industries into a number of distinct effects, and it is similar to the decompositions undertaken by Kogler et al. (2017) and Essletzbichler (2015). The mathematical tautology used to decompose the change in average relatedness is standard and is repeated here as Equation 7. $\overline{ER}$ is the weighted average for 2008, $\overline{ER'}$ is the weighted average for 2013 and Δ indicates the difference between 2008 and 2013. m indexes the regional industries and s_m is the employment share of regional industry m. C, N, and X are the sets of continuing, entering, and exiting regional industries respectively.

$$\Delta\overline{ER} = \overline{ER'} - \overline{ER}$$

$$= \sum_{m\in C} \Delta s_m(ER_m - \overline{ER}) + \sum_{m\in N} s'_m(ER'_m - \overline{ER}) -$$

$$\sum_{m\in X} s_m(ER_m - \overline{ER}) + \sum_{m\in C} s'_m \Delta ER_m \quad (7)$$

The four elements of the right-hand side of Equation 7 are:

1. Between effect. If regional industries with high relatedness grow more than others, then average relatedness will increase. This is computed from the change in regional industries' employment shares together with the difference between the regional industries' relatedness and average relatedness.
2. Entry effect. The aggregate economy should diversify into related activities and, hence, entry should contribute positively to average relatedness. This is computed from the difference between the new regional industries' relatedness and the initial aggregate average.

3. Exit effect. Regional industries with low relatedness should disappear, which should increase average relatedness. This is computed from the difference between the disappearing regional industries' relatedness and the initial aggregate average.
4. Within effect. The relatedness of a given regional industry changes as other industries in the region grow, decline, enter and exit. The within effect is therefore computed from the change in relatedness index for all regional industries.

The method in Equation 7 is a bivariate method focusing on the relationship between the growth of regional industries and changes in relatedness. In other words, none of the other variables included in our regression analysis are held constant. Therefore, the second step in our analysis is the application of an extended decomposition method that considers confounding effects within the first term on the right-hand side of equation 7; the between effect (Holm et al., 2016).

The between effect can be rewritten as the sum of K products, where K is the number of confounding variables included, plus one for the variable in focus. This is shown in Equation 8.

$$\sum_{m \in C} \Delta s_m (ER_m - \overline{ER}) = \sum_{k=1}^{K} \gamma_k Cov(z_{k,m}, ER_m) \quad (8)$$

The between effect from Equation 7 is shown on the left of Equation 8. Equation 8 says that the between effect for the variable ER_m, that is, the degree to which aggregate changes in average ER_m depend on the covariance between ER_m and growth, depends on K other variables ($z_{k,m}$) that simultaneously covary with growth and with ER_m.

In our case there are 23 $z_{k,m}$ variables. These are the eight variables of Table 1, nine fixed effects for industries and six fixed effects for regions. The element arising from $z_{k,m} = ER_m$ is the direct effect of relatedness, while the remaining effects $(z_{k,m} \neq ER_m)$ are the indirect effects of the other variables. For example, the effect of HC intensity on the change in average relatedness is computed when $z_{k,m} = HC_m$. The K parameters γ_k are the estimated slope coefficients of a weighted least squares regression of relative industry growth on the K variables of the analysis with s_m as the weight. The γ_k parameters and the covariances are computed using the set of continuing industries, C.

Generally, it is prudent to apply the identity in Equation 7 prior to extending it with Equation 8, since the relevance of further decomposing the between effect (Equation 8) depends on the relative importance of the between effect in the overall decomposition of Equation 7.

4. Results

4.1 Relatedness and static diversification

For comparability we have first estimated the model without the control variables. These regressions describe the bivariate relationships between the relatedness indices and growth and are shown in Table 4.

Relatedness index:	ER		BR		SC	
	Estimate	SE	Estimate	SE	Estimate	SE
Slope coefficient	-0.121***	(0.018)	-0.195***	(0.019)	-0.127***	(0.043)
Note: *p<0.1; **p<0.05; ***p<0.01						

Table 4. Slope coefficients from simple regressions of growth on respective relatedness indices

Table 4 shows that all three indices are similar in their negative bivariate relationship to the growth of regional industries. Industries with many related jobs in the region tend to have lower growth than others. When adding the control variables and interaction terms of Equation 6 the strength of this negative relationship decreases. The result is shown in Table 5. The second and third columns substitute our preferred measure of relatedness with the two alternatives discussed previously: the broad relatedness (BR) index in column two and standardized closeness (SC) in column three.

	1		2		3	
Relatedness index:	ER		BR		SC	
	Estimate	SE	Estimate	SE	Estimate	SE
Relatedness	-0.020	(0.019)	-0.141**	(0.057)	-0.074**	(0.037)
Spatial	0.003	(0.044)	0.005	(0.050)	-0.005	(0.042)
Economic	0.031	(0.037)	0.035	(0.043)	0.027	(0.039)
HC	0.095***	(0.019)	0.084***	(0.016)	0.119***	(0.033)
Routine	-0.053**	(0.026)	-0.068**	(0.029)	-0.052	(0.032)
Start-ups	0.095***	(0.034)	0.077***	(0.028)	0.102***	(0.036)
Size	-0.196***	(0.031)	-0.120***	(0.046)	-0.203***	(0.031)
RRS	-0.044	(0.040)	-0.047	(0.039)	-0.058	(0.043)
High-tech manuf.	-0.235***	(0.030)	-0.185***	(0.037)	-0.212***	(0.081)
Med-high-tech manuf.	-0.188***	(0.034)	-0.135**	(0.058)	-0.070	(0.053)
Med-low-tech manuf.	-0.059	(0.039)	-0.015	(0.056)	0.031	(0.047)
Low-tech manuf.	-0.149***	(0.045)	-0.166***	(0.050)	-0.098**	(0.048)
Primary Util. Constr.	0.182***	(0.045)	0.179***	(0.033)	0.200***	(0.044)
High-tech KIS	0.259***	(0.042)	0.240***	(0.044)	0.208***	(0.038)
Market KIS	-0.052	(0.045)	-0.045	(0.045)	-0.059	(0.088)
Financial KIS	0.052	(0.032)	-0.005	(0.052)	-0.012	(0.032)
Other KIS	-0.022	(0.040)	-0.033	(0.054)	-0.073*	(0.040)
(Omitted: other services)						
Copenhagen	0.121	(0.095)	0.097	(0.108)	0.122	(0.098)
N. Jutland	-0.021	(0.066)	-0.007	(0.067)	-0.019	(0.067)
E. Jutland	0.038	(0.036)	0.053	(0.033)	0.032	(0.032)
S. Jutland	-0.025	(0.045)	-0.050	(0.044)	-0.026	(0.050)
W. Jutland	-0.009	(0.045)	-0.003	(0.051)	0.019	(0.052)
Zealand	-0.027	(0.059)	-0.039	(0.071)	-0.044	(0.079)
(Omitted: Funen)						
Constant	-0.014	(0.009)	-0.083***	(0.024)	0.010	(0.013)
Interactions with Relatedness						
High-tech manuf.	-0.034	(0.078)	-0.081	(0.114)	0.315**	(0.131)
Med-high-tech manuf.	0.115***	(0.041)	0.002	(0.026)	0.181	(0.153)
Med-low-tech manuf.	-0.066	(0.040)	-0.048	(0.061)	0.048	(0.059)

Low-tech manuf.	-0.014	(0.048)	-0.044*	(0.024)	-0.123**	(0.063)
Primary Util. Constr.	-0.148	(0.123)	-0.034	(0.066)	-0.217***	(0.027)
High-tech KIS	-0.100	(0.097)	0.095	(0.108)	-0.276***	(0.066)
Market KIS	-0.206	(0.172)	-0.166*	(0.091)	-0.326***	(0.087)
Financial KIS	-0.051	(0.077)	0.014	(0.065)	0.117	(0.200)
Other KIS	0.028	(0.074)	0.240	(0.321)	-0.366**	(0.156)
Spatial	0.036**	(0.014)	0.017	(0.016)	0.046**	(0.020)
Economic	-0.038***	(0.014)	0.016	(0.030)	0.004	(0.011)
HC	0.023	(0.022)	-0.025	(0.035)	0.066	(0.060)
Routine	0.005	(0.021)	-0.025	(0.016)	-0.011	(0.021)
Start-ups	-0.053**	(0.023)	-0.063**	(0.030)	0.011	(0.029)
Size	-0.002	(0.018)	0.106***	(0.027)	0.025	(0.025)
RRS	0.007	(0.016)	0.007	(0.012)	0.038*	(0.022)
Observations	2311		2311		2311	
R2	0.117		0.120		0.128	
Adj. R2	0.102		0.105		0.113	
F-stat (df=39;2271)	7.697***		7.955***		8.518***	
VIF > 5	*None*		Relatedness (VIF=8.95)		*None*	

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5. Regression results for regional industry growth

As weighted effects coding is used for the categorical variables and all continuous variables have mean zero, the estimated effect of relatedness in Table 5 is the effect for the average industry when all variables interacted with relatedness are zero (i.e. at their mean). The estimates for regions and industry/technology-groups are deviations from this mean. Table 5 shows that there is no statistically significant effect on growth of relatedness (ER) for the average industry. We can thus reject hypothesis H1a. This questions whether the property of relatedness in general moderates regional diversification (Boschma, 2017) and the concentration of industries in space (Hidalgo et al., 2018), or relatedness primarily moderates dynamic diversification (Boschma, 2017; Hidalgo et al., 2018). However, the presence of skill related industries in a region might still support employment growth under specific circumstances cf. hypotheses H1b and H1c. This is explored by the interaction effects.

The estimated effects for the interactions of ER with three variables are statistically significant. The interaction effect for Spatial is positive indicating a stronger effect of relatedness in the spatial periphery. The interaction effect of Start-ups is negative indicating a weaker effect for relatedness when industry employment is concentrated in younger plants, and the interaction effect for Economic is negative indicating a weaker effect of relatedness when policy spending is higher, i.e. in the economic periphery. These results qualify and partially support hypothesis H1b: Relatedness is more important for static diversification in peripheral regions. However, this holds for spatially but not for economically peripheral regions. Industries predominantly located the economic periphery in 2008 were not able to exploit the potentially beneficial knowledge flows from the presence of related industries over the ensuing five years. There are many reasons for why regions become economically peripheral regions, but one of them could be the lacking ability of the regional industries to exploit benefits from relatedness. The result for spatial periphery is in line with Kuusk (2021) who finds that industries in peripheral regions rely more heavily on other related industries for employment growth.

Hypothesis H1c is rejected as the interaction term for HC intensity is not significant, thus the effect of relatedness is not greater for HC intensive industries.

As argued above, if relatedness has an effect on static diversification, it must have an effect on growth. But growth may either lead to increased balance or increased concentration depending on whether the small or large industries grow more than others. There is no statistically significant interaction term for Size meaning that relatedness leads to neither balance nor concentration systematically. As no general effect of relatedness was found (hypothesis H1a) this result is not surprising.

Columns two and three of Table 5 show that the alternative relatedness indices perform somewhat different. The broad index in column two captures skill relatedness, as our preferred index in column one, but it also captures relatedness in a broader sense where industries also benefit from the local presence of other industries to which they more often act as sources of skills. This could be the explanation for the negative direct effect of relatedness in column two: competition from other industries for employment hinders growth. The estimated direct effect in the third column, where relatedness is measures as standardized closeness, is also negative. This underlines the importance of focusing on related jobs (cf column 1) and not just related industries and supports other studies questioning whether relatedness is always positive (Fitjar and Timmermans, 2019).

To graphically illustrate the marginal effects of relatedness considering the interaction terms of Equation 6 we show the marginal effect at ± 2 standard deviations of the variable interacted with Relatedness. This is shown in Figure 1.

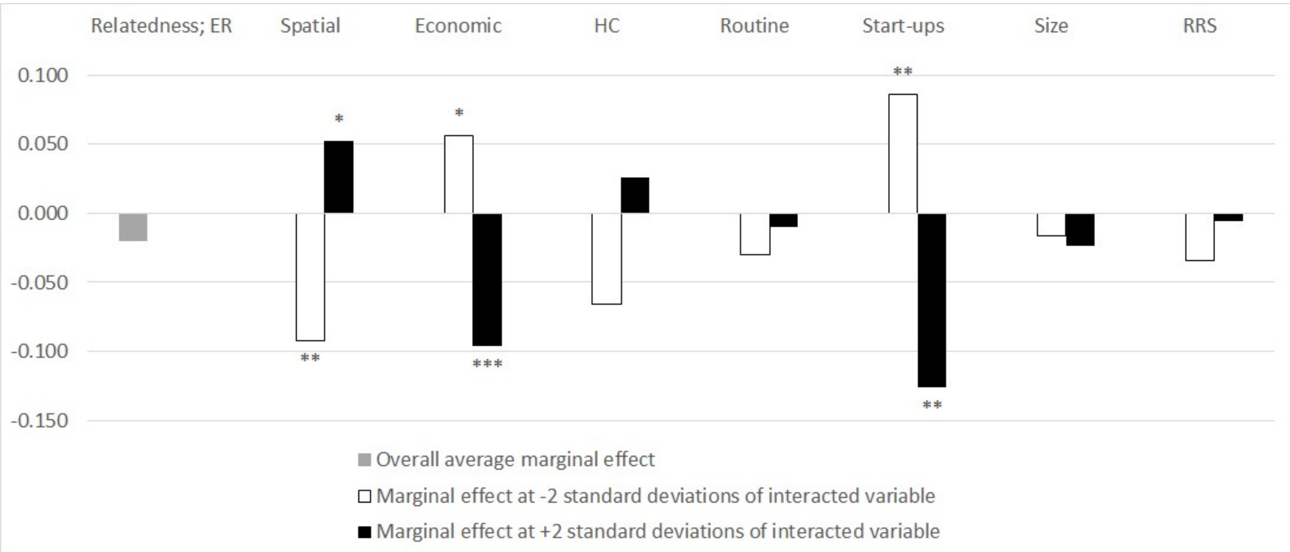

Figure 1. Marginal effects of Relatedness

Figure 1 shows that the interpretations of the results in Table 5 all hold at ± 2 standard deviations of the respective variables discussed above. To test the significance of the marginal effects in Figure 1 the implied restrictions were included in the model one at a time and tested with F-tests.

4.2 Feedback from static diversification to relatedness

Table 1 showed a net decrease in average relatedness in Denmark 2008-2013. As argued earlier, relatedness can also be seen as a dimension of diversity: disparity, and in this section, we analyze the role of static diversification, that is, differential growth of industries, in the decline of average relatedness. Table 6 shows how the decrease in average relatedness is decomposed when using the methodology of Equation 7. The change in ER at the aggregate level is -0.005 meaning that the number of jobs related to the average regional industry decreased by 0.5 percentage points. This is standardized so the effects sum to -1 in Table 6.

Between effect	Entry effect	Exit effect	Within effect
-0.338	-0.036	-0.022	-0.605

Table 6. Decomposition step 1

The effect of dynamic diversification in regions (entry effect and the exit effect) on the change in average relatedness is negligible. The main driver is the within industry effect which accounts for 60.5 percent of the decrease in average relatedness while the between industry effect accounts for 33.8 percent. The between effects reflect the importance of relative growth. Regional industries with relatively high rates of growth had low relatedness, and mathematically this accounts for about one third of the national decrease in average relatedness. This result rejects hypothesis H2a: there is no positive relationship between employment growth and relatedness leading to increasing average relatedness at the aggregate level.

In step 2 of the decomposition (Equation 8) the between effect is split up into the direct relatedness effect on growth of regional industries, and the effects of other variables that correlate with both growth and relatedness.

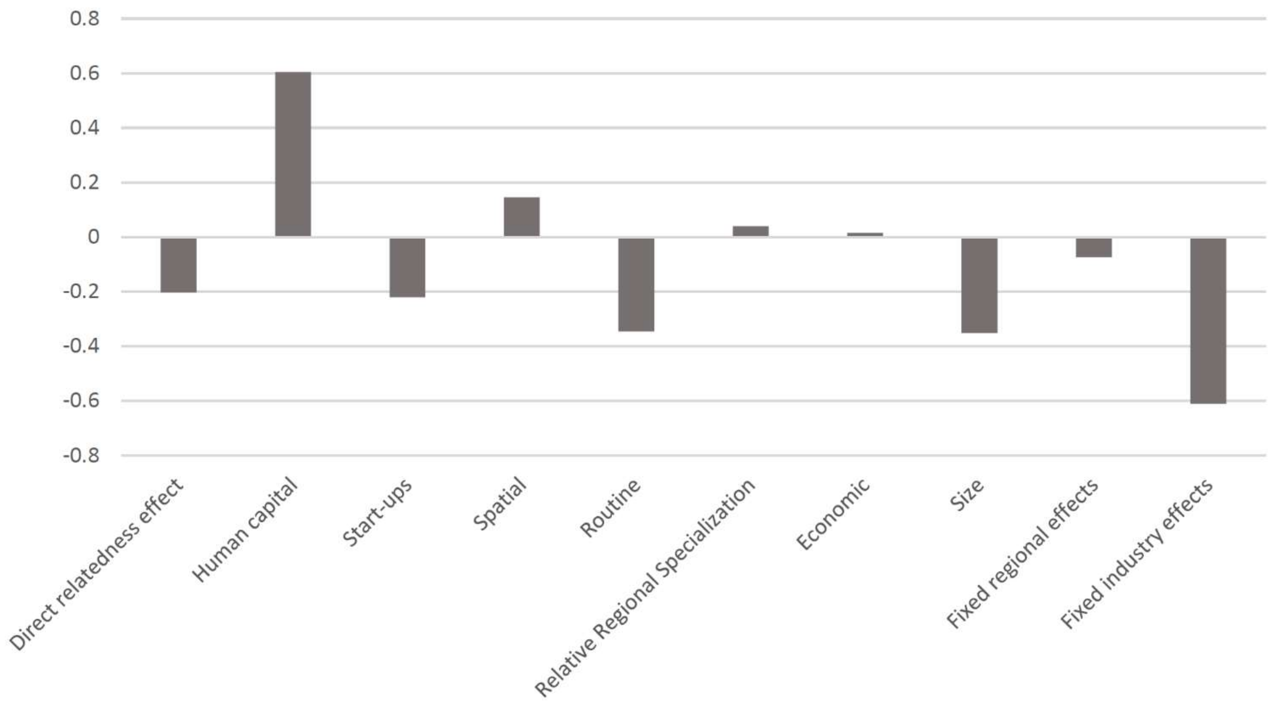

Figure 2. Decomposition step 2

In Figure 2 the effects again sum to -1 as the between effect in step 1 was negative. The direct effect attributed to relatedness is small and negative, and very close to the estimate from the regression analysis. In the regression analysis we did not find that relatedness is relatively more important for HC intensive regional industries, but Figure 2 shows that HC has a sizable positive contribution to the change in relatedness. In other words, there is a positive correlation between relatedness and HC intensity across regional industries and as HC is positively associated with growth, the positive correlation provides a positive contribution to the evolution of average relatedness. Conversely, in line with our hypothesis H1b, the variables for spatial and economic periphery strengthen the effect of relatedness on growth, but the lack of a direct relationship between the two variables and growth entails that they do not affect the evolution of average relatedness. Thus, we can reject hypothesis H2b but not H2c. However, the argument for H2c must be developed further. HC intensive industries do have both high growth and high relatedness, but this is not because there is a positive relationship between relatedness and growth. One alternative explanation could be that HC intensive industries co-locate with related industries not to exploit these industries as sources for skilled employees, but because they rely on other common local factors.

The main negative effect that cancels out the positive contribution from HC is the industry fixed effects. In other words, there is a strong tendency for structural change towards industries with low levels of relatedness, and this relationship between industry fixed effects and growth strengthens the negative correlation between relatedness and growth.

5. Conclusions

When discussing and analyzing diversity and diversification it is fruitful to see diversity as having three dimensions: variety, balance, and disparity. The literature on regional dynamic diversification by definition focusses on variety and finds that relatedness is an important factor shaping regional industrial diversification patterns. In this paper, we show that the central role of relatedness for dynamic diversification does not extend to static diversification in general, that is, it does not extend to the balance dimension of diversity. Relatedness does, however, appear central to static diversification in spatially peripheral regions and in economically non-peripheral regions. We show results for three different indices of relatedness for comparability. Two of the indices are taken from the existing literature while the third is an employment-weighted relatedness index developed in this paper. None of the indices show a positive relationship between relatedness and static diversification, which rejects the hypothesis that relatedness shapes the direction of static diversification. Furthermore, the relationship between relatedness and employment growth is found to be independent of industry size, which underlines that relatedness is not systematically related to change in the balance dimension of diversity.

The results show that relatedness is important for industries in the spatial periphery and negative for industries in the economic periphery. This underlines that spatially peripheral regions and economically peripheral regions should be treated differently in research and in policy. Relatedness is not found to be more important for HC intensive industries, but it is more important for industries where a larger share of jobs is in older plants. In other words, while the creation of new plants is generally positively related to growth, the effect of relatedness is weaker for industries where many jobs are at new plants. This might be explained by increased competition for labor with specific skills which enhance the liability of newness (Fitjar and Timmermans, 2019). Taken together, the results show that the relationship between relatedness and diversification is context dependent (Kuusk, 2021), and thus the results qualify the "principle of relatedness" (Hidalgo et al., 2018). There is compelling evidence in the innovation literature that firm level innovation, i.e. new combinations of knowledge, are combinations of related knowledge (Laursen, 2012). This leads to new specializations and new firms in new industries that are related to existing specializations and industries. At the regional industry level, however, the explanatory power of relatedness for diversification is relatively limited. In particular, static diversification appears to be driven by factors that we cannot account for in our data. Future research could for example look more into the roles of innovation, and technological and institutional change for diversification at the regional industry level. There is a wide range of analyses of dynamic diversification and future studies of static diversification could be inspired by these analyses, for example by focusing on social capital and institutions (Antonietti and Boschma, 2021).

Relatedness can also be seen as a dimension of diversity in itself as it captures disparity between industries. When industries grow and decline the average relatedness of the regional industry structure will also change. As relatedness is not found to be related to growth it is no surprise that differential growth does not explain the changes in average relatedness. We do, however, find a relationship between both relatedness and HC intensity, and between HC intensity and growth, which is associated with increasing average relatedness over the period. In this study we have focused on growth over a five year period for comparability with earlier studies but it is possible that some effects such as HC intensity work over shorter timespans – for example growth over a two year period. The nexus of HC intensity, relatedness and growth deserves further attention.

The results have implications for policy making. Regional diversification into new industries is high on the policy agenda (Boschma, 2017). The literature on dynamic diversification suggests that policy makers should avoid promoting diversification into unrelated industries as this is unlikely to succeed (Neffke et al., 2011). This paper adds that a more nuanced approach is needed as static diversification does not rely on relatedness to the same extent as dynamic diversification. In other words, policy for dynamic diversification may not create more jobs in the region in the foreseeable future.

A limitation of this study is that the results may reflect the empirical context. The years after the 2008 financial crisis in Denmark were dominated by a slow recovery, and there is, to the best of our knowledge, so far no research on how and why relatedness between industries change over time and varies across economies. In addition, studies of diversification necessarily rely on existing industry classifications. Over time industries change in respect to outputs produced, technologies employed, and skills used but this is overlooked in studies of both dynamic and static diversity.

6. References

Andersen, E.S. (2003), 'The evolving tree of industrial life: an approach to the transformation of European industry.' Paper for the second workshop on the Transformation of Europe, 31 January – 2 February, Turin, Italy.

Antonietti, R., & Boschma, R. (2021). 'Social capital, resilience, and regional diversification in Italy,' *Industrial and Corporate Change*, **30**(3), 762-777.

Bernard, A.B., Smeets, V., and Warzynski, F. (2016), 'Rethinking deindustrialization,' *CEP Discussion Papers*, No. 1423. London: Center for Economic Performance, London School of Economics and Political Science.

Boschma, R., Eriksson, R., and Lindgren, U. (2009), 'How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity,' *Journal of Economic Geography*, **9**(2), 169-190.

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

Boschma, R., Balland, P-A., and Kogler, D. (2015), 'Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010,' *Industrial and Corporate Change*, **24**(1), 223–250.

Breschi, S., & Lissoni, F. (2001), ‘Knowledge spillovers and local innovation systems: a critical survey,’ *Industrial and Corporate Change*, **10**(4), 975-1005.

Caragliu. A., de Dominicis, L., and de Groot, H.L.F. (2016), ‘Both Marshall and Jacobs were right!,’ *Economic Geography,* **92**(1), 87-111.

Cortinovis, N., and van Oort, F. (2015), ‘Variety, economic growth and knowledge intensity of European regions: a spatial panel analysis,’ *Annals of Regional Science*, **55**(1), 7-32.

Cohen, W.M., and Levinthal, D.A. (1990), ‘Absorptive capacity: a new perspective of learning and innovation,’ *Administrative Science Quarterly,* **35**(1), 128–152.

Delgado, M., Porter, M.E., and Stern, S. (2015), ‘Clusters and the great recession,’ Paper presented at the DRUID conference, 15-17 June, 2015, Rome, Italy.

DORS. (2015), *Dansk Økonomi, forår 2015*. Copenhagen, Denmark: Danish Economic Councils.

Eriksson, R., Lindgren, U., and Malmberg, G. (2008), ‘Agglomeration mobility: effects of localization, urbanization and scale on job changes,’ *Environment and Planning A,* **40**(10), 2419-2434.

Essletzbichler, J. (2015), ‘Relatedness, industrial branching and technological cohesion in US metropolitan areas,’ *Regional Studies*, **49**(5), 752–766.

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

Fitjar, R.D., and Timmermans, B. (2019), ‘Relatedness and the resource curse: Is there a liability of relatedness?,’ *Economic Geography*, **95**(3), 231-255.

Goos, M., Manning, A., and Salomons, A. (2014), 'Explaining job polarization: routine-biased technological change and offshoring,' *American Economic Review,* **104**(8), 2509–2526.

Hartog, M., Boschma, R., and Sotarauta, M. (2012), 'The impact of related variety on regional employment growth in Finland 1993-2006: High-tech versus medium/low-tech,' *Industry and Innovation,* **19**(6), 459-476.

Hausmann, R., and Hidalgo, C.A. (2011), 'The network structure of economic output,' *Journal of Economic Growth*, **16**(4), 309–342.

Henning, M., Stam, E., and Wenting, R. (2013), 'Path dependence research in regional economic development: cacophony or knowledge accumulation?,' *Regional Studies*, **47**(8), 1348–1362.

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

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., and Zhu, S. (2018), 'The Principle of Relatedness,' In: Morales, A., Gershenson, C., Braha, D., Minai, A., and Bar-Yam, Y. (eds) *Unifying Themes in Complex Systems IX.* ICCS 2018. Springer Proceedings in Complexity. Springer, Cham.

Holm, J.R., Andersen, E.S., and Metcalfe, J.S. (2016), 'Confounded, augmented and constrained replicator dynamics: complex selection processes and their measurement,' *Journal of Evolutionary Economics*, **26**(4), 803–822.

International Monetary Fund. (2018), *World Economic Outlook: Cyclical Upswing, Structural Change*. Washington, DC: IMF.

Isaksen, A. (2015), ‘Industrial development in thin regions: trapped in path extension?,’ *Journal of Economic Geography,* **15**(3), 585–600.

Jacobs, J. (1969). ‘*The economy of cities’*. Random House, New York.

Kemeny, T. and Storper, M. (2020), ‘Superstar cities and left-behind places: disruptive innovation, labor demand, and interregional inequality,’ *International Inequalities Institute, London School of Economics and Political Science*. Working Paper (41). London, UK.

Klepper, S. (2010), ‘The origin and growth of industry clusters: the making of Silicon Valley and Detroit,’ *Journal of Urban Economics*, **67**(1), 15–32.

Klepper, S., and Thompson, P. (2006), ‘Submarkets and the evolution of market structure,’ *The RAND Journal of Economics*, **37**(4), 861–886.

Kogler, D.F., Essletzbichler, J., and Rigby, D. (2017), ‘The evolution of specialization in the EU15 knowledge space,’ *Journal of Economic Geography,* **17**(2), 345–373.

Kuusk, K. (2021). ‘Regional differences in how related variety ‘works’: the case of labour mobility,’ *European Planning Studies*, **29**(10), 1951-1973.

Laursen, K. (2012). ‘Keep searching and you’ll find: what do we know about variety creation through firms’ search activities for innovation?,’ *Industrial and Corporate Change*, **21**(5), 1181-1220.

Metcalfe, J.S. (1998), *Evolutionary Economics and Creative Destruction*. Routledge, London and New York.

Morkutė, G., Koster, S., and Van Dijk, J. (2017), ‘Employment growth and inter-industry job reallocation: spatial patterns and relatedness,’ *Regional Studies*, **51**(6), 958-971.

Neffke, F., Henning, M., and Boschma, R. (2011), 'How do regions diversify over time? Industry relatedness and the development of new growth paths in regions,' *Economic Geography*, **87**(3), 237–265.

Neffke, F., Otto, A., and Weyh, A. (2017), 'Inter-industry labour flows,' *Journal of Economic Behaviour and Organization*, **142**, 275–292.

Neffke, F., Hartog, M., Boschma, R., and Henning, M. (2018), 'Agents of structural change: The role of firms and entrepreneurs in regional diversification,' *Economic Geography*, **94**(1), 24-48.

Nieuwenhuis, R., Grotenhuis, M., and Pelzer, B. (2017), 'Weighted effect coding for observational data with wec,' *The R Journal*, **7**(1), 477-485.

Noseleit, F. (2015), 'The role of entry and market selection for the dynamics of regional diversity and specialization,' *Regional Studies*, **49**(1), 76–94.

Siegel, P. B., Johnson, T. G., and Alwang, J. (1995), 'Regional economic diversity and diversification,' *Growth and Change*, **26**(2), 261-284.

Stirling, A. (2007), 'A general framework for analysing diversity in science, technology and society,' *Journal of the Royal Society Interface*, **4**(15), 707-719.

Xiao, J., Boschma, R., and Andersson, M. (2018), 'Resilience in the European Union: the effect of the 2008 crisis on the ability of regions in Europe to develop new industrial specializations,' *Industrial and Corporate Change*, **27**(1), 15-47.

Appendix

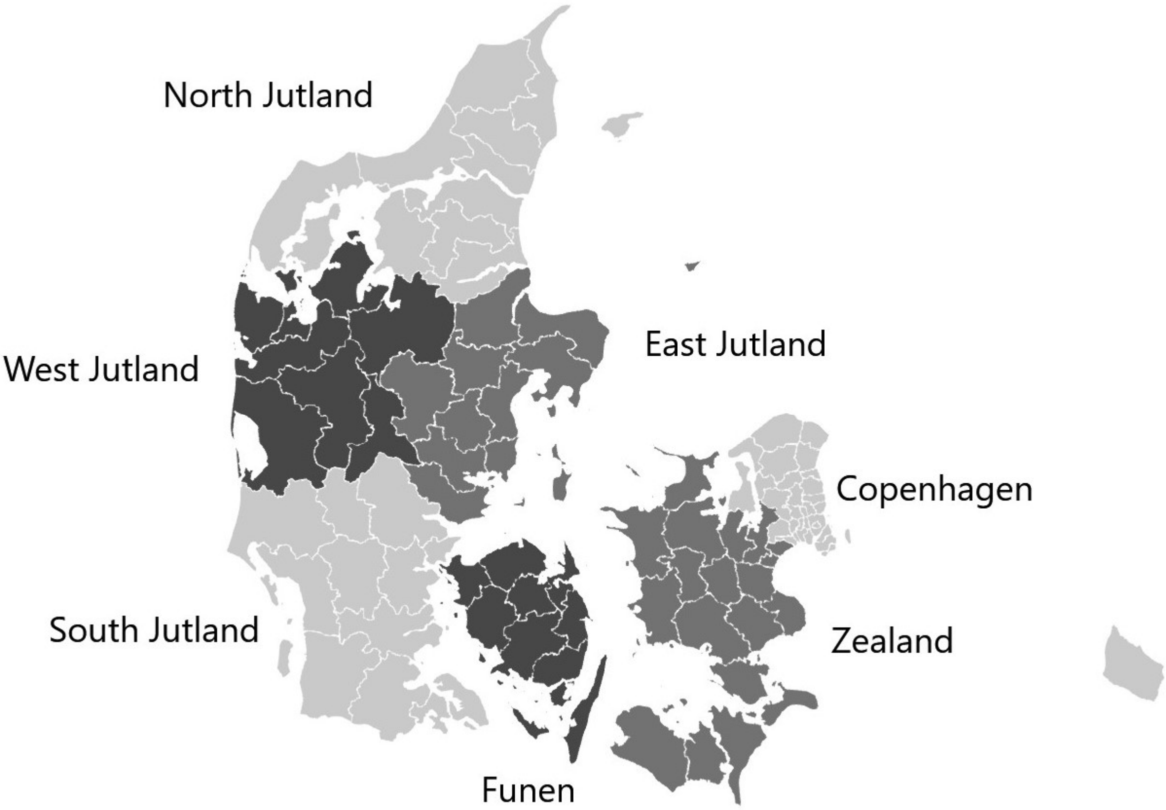

Figure 3. Danish NUTS 3 regions and the NUTS 2 regions Copenhagen and Zealand

Figure 4. Spatial periphery in Denmark

The variable for spatial periphery is taken from DORS (2015) and is relatively standard. By this definition spatial periphery are the darkest municipalities in Figure 4. They are neither city municipalities nor close-to-city municipalities.

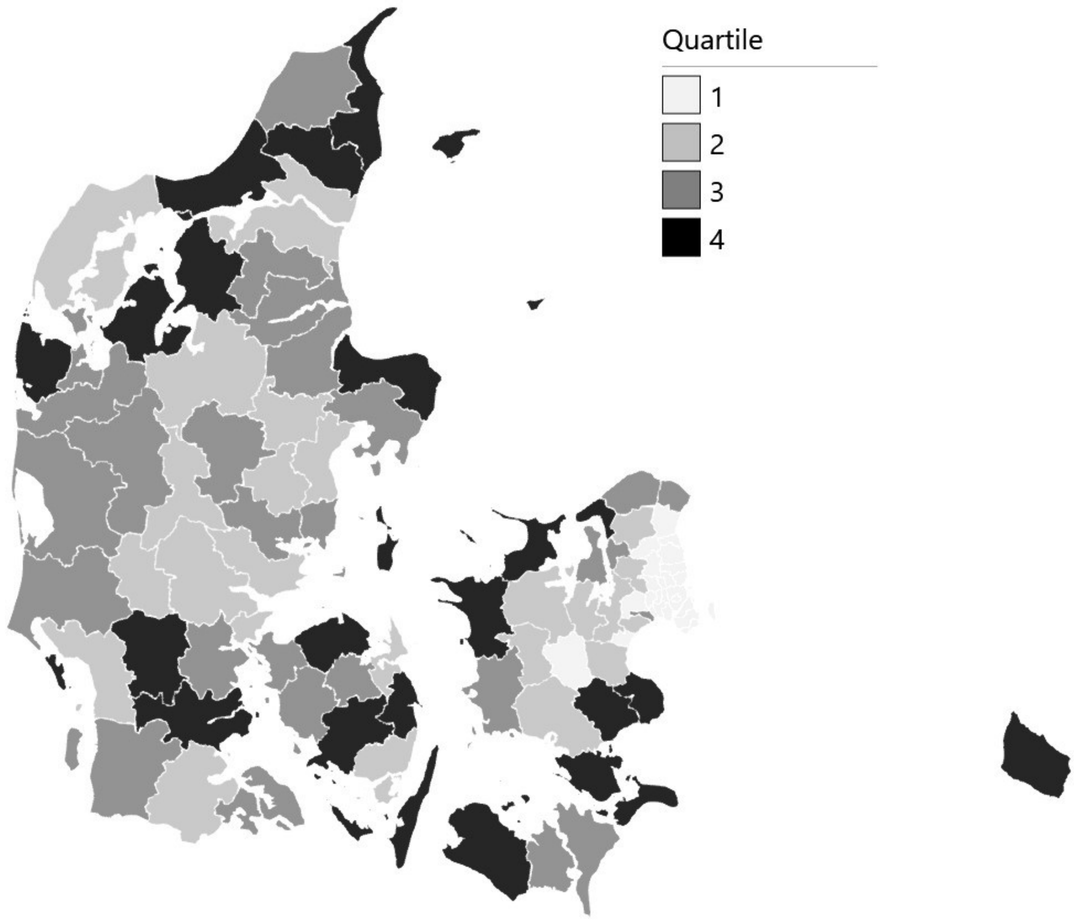

Figure 5. Economic periphery in Denmark

Economic periphery is defined by municipal spending on active industrial policy per full time equivalent job in the municipality. In Figure 5 the municipalities of Denmark are shaded depending on the quartile for the level of spending. As can be seen by comparing Figures 4 and 5 spatial and economic periphery to some degree overlap – i.e. no city municipalities are in the fourth quartile by spending – but there are also spatially peripheral areas that are economically central and economic periphery that is not spatially peripheral.

[i] NUTS: Nomenclature of Territorial Units for Statistics

[ii] Found on www.statistikbanken.dk, Table REGK31. The functions referred to here as active industrial policy are the functions numbered 6.48.xx. Last accessed 9 August 2021.

[iii] Miscellaneous is mostly a negative expense indicating that incomes are registered in this account. Some of the remaining accounts are used by very few municipalities, although they are used consistently each year by those municipalities.

[iv] https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Knowledge-intensive_services_(KIS)

[v] https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:High-tech_classification_of_manufacturing_industries