# The roles of diversity, complexity, and relatedness in regional development – What does the occupational perspective add?

Tom Broekel, Rune Dahl Fitjar & Silje Haus-Reve

Papers in Evolutionary Economic Geography

# 21.35

Utrecht University Human Geography and Planning

# The roles of diversity, complexity, and relatedness in regional development – What does the occupational perspective add?

Tom Broekel[*], Rune Dahl Fitjar and Silje Haus-Reve

UiS Business School, University of Stavanger

**Abstract**

Contemporary research highlights the importance of relatedness, diversity, and complexity for regional economic development. However, few empirical studies simultaneously test the relevance of all these dimensions or examine how their importance varies across distinct spatial contexts. The literature also concentrates on explaining regional diversification, whereas we know less about how they affect economic and employment growth. In addition, most studies have examined industrial relatedness at the expense of the at least similarly crucial occupational dimension when studying knowledge-based regional development. The chapter discusses these issues and presents a study on how occupational diversity, complexity and relatedness shape employment growth in Norway to illustrate how an occupational perspective on regional industries can add to the understanding of evolutionary economic development.

**JEL:** R11, O31, O33, J24

**Keywords:** relatedness, diversity, complexity, occupation, region, Norway

## INTRODUCTION

Economic geography has long been interested in how the economic structure of a place shapes the opportunities for innovation and growth there (Glaeser et al., 1992; Boschma and Iammarino, 2009; Capello and Lenzi, 2016; Eriksson, Hansen, and Winther, 2017). By now it is well-established that the degree to which a region's actors have specialized in a diversity of knowledge domains, how complex these domains are, and how closely they are interrelated matter for economic growth. A wide range of studies present empirical support for the importance of the dimensions of diversity, relatedness, and complexity for regional development (Frenken et al., 2007; Balland and Rigby, 2017; Balland et al., 2019).

Although previous research has convincingly demonstrated the importance of these dimensions, several limitations persist. This chapter discusses five such limitations of the contemporary literature: First, few empirical studies consider multiple dimensions simultaneously. Thus, the unique contribution of each dimension to a region's development remains somewhat unclear. Second, we see a surge in the number of studies that, inspired by the work on related diversification, seek to explain regions' probabilities of expanding their industrial portfolio (Neffke et al., 2011; Boschma, 2017). Yet, ultimately, diversification is merely a means to an end. It is economic growth that matters, and for which we need to know the relevance of relatedness, diversity, and complexity. Third, most traditional studies of regional development examine expansions of employment as the dependent variable. However, economic development is but one of many dimensions to development overall, and employment growth is in turn but one of many dimensions to economic development. In particular, complexity's contribution to growth may be less reflected in employment growth. Consequently, it is imperative that we expand our perspectives on development to consider additional dependent variables, including economic variables such as wage growth and value added, as well as non-economic variables such as sustainability. Fourth, to obtain generalizable results, it is important to confirm the

[*] Corresponding author: tom.brokel@uis.no, University of Stavanger Business School, Stavanger, Norway

contribution of relatedness, diversity, and complexity across regions. However, the relevance and impact of these three dimensions on economic development can be expected to vary across regions. So far, we know little about the heterogeneity of these dimensions' impacts. Last—and certainly not least—we argue that so far, the literature still tends to pay too little attention to the particular skills and tasks of employees required for regions to develop new economic activities. Following Wixe and Andersson (2017), Jara-Figueroa et al. (2018), Farinha et al., (2019), and Hane-Weijman et al. (2020), we argue in this chapter that adopting an occupation-centered perspective will give a more precise and informative picture in this context, compared to the traditional industry focus.

In addition to discussing these five issues, we empirically illustrate some of these aspects based on Norwegian data. Norway is an interesting case for several reasons. First, a large share of its regions currently prosper due to a (still) flourishing oil and gas sector. However, these regions need to develop new sectors and businesses that reduce their dependency on oil and gas and that simultaneously promise long-term and sustainable economic growth. To support this transition with suitable policies, we must better understand how existing knowledge and competence impact transition possibilities and processes. Second, we demonstrate the type of empirical analyses needed using high-quality data covering all Norwegian labor market regions and industries. Crucially, this data features information on occupations, which makes it possible to explicitly consider this dimension.

Contradicting our expectations based on the theoretical discussion and existing empirical insights, the empirical analyses suggest that employment growth between 2009 and 2014 in Norway was more strongly shaped by regional industrial relatedness than by occupational relatedness, complexity, and regions' occupational compositions. Consequently, the impact of these dimensions appears to be context- (or country-) specific, which calls for more research in this direction in the future.

The chapter starts by reviewing the literature on diversity, complexity and relatedness as drivers of economic development, and discusses how to extend the existing debate to the analysis of the occupational composition of industries. We then present an empirical analysis illustrating some of the theoretical arguments. In doing so, we first discuss some common measures of industry-regions' occupational diversity, complexity and relatedness. Second, we test their relationships with employment growth in Norwegian industry-regions. The conclusion discusses some of the literature's limitations against the backdrop of our empirical study.

## DIVERSITY, COMPLEXITY OR RELATEDNESS: THE THEORY AND THE REAL WORLD

How the industry structure of a region shapes regional economic performance is a classic question in economic geography. Marshall (1890) laid the foundation for thinking about how the location of an industry had consequences ranging from the emergence and growth of subsidiary industries to the transmission of skills in the local labor market (Belussi and Caldari, 2009). A similarly well-discussed extension of this research is whether regions benefit mainly from specializing in a small number of industries, or from diversifying into many different activities. In the tradition of Marshall, leading economists such as Arrow (1962) and Romer (1986) argued for the benefits of specialization by emphasizing that agglomerations of firms in the same industry generate external economies of scale (Glaeser et al., 1992). Such economies arise from positive externalities of a firm's activities in a place for those of other firms in the same industry and place. One firm's activities may benefit other firms by expanding the labor market, enabling the use of shared infrastructures and suppliers, or by knowledge spillover across firm boundaries.

A different argument is closely associated with Jane Jacobs's influential 1969 work, in which she posited that places benefit from a range of different industries being located there because knowledge spillovers across industries carry greater potential for the cross-fertilization of ideas than spillovers within industries. Benefits arising from the combination of heterogeneous knowledge are commonly referred to as ’Jacobs externalities’ Various studies have examined whether Marshall–Arrow–Romer externalities or Jacobs externalities are more important for development, often focusing on the role of knowledge spillovers and innovation (e.g., van der Panne 2004; Ejermo, 2005; Beaudry and Schiffauerova, 2009; Desrochers and Leppälä, 2011; Caragliu et al., 2016).

The emergence of evolutionary economic geography switched the discussion from how the specialization or diversity of regional industry structures shape development to how the industry structure changes over time through the entry of new industries and the loss of old ones (Boschma and Frenken, 2006). Regions are more likely to diversify into industries that are related to those in which they are currently specialized. Due to the path-dependent nature of regional development, the existence of related industries and competences in a region increases the likelihood of new (and related) industries emerging there. Moreover, these industries are more likely to succeed and grow when utilizing existing (related) competences and resources. This process has been conceptualized as “regional branching” (Boschma and Frenken, 2011) or “related diversification” (Frenken et al., 2007). Relatedness gives insights into the probable ease of a region’s diversification into new sectors, and therefore into a region’s path-dependent growth potential.

In addition to emphasizing the path-dependent nature of regional economic development, this literature also overcame the quasi-binary perspectives of Marshall–Arrow–Romer and Jacobs externalities. The optimal context for knowledge spillovers is neither specialization in a single industry nor diversification into a vast number of unrelated industries. Rather, a variety of related industries offers the best conditions for knowledge spillover, innovation, and growth (Frenken, van Oort and Verburg, 2007; Boschma and Iammarino, 2009). Within an industry, firms often have overlapping competences and thus limited potential for learning from one another. Across unrelated industries, knowledge may be difficult to integrate and to apply and, consequently, is often of limited value. Conversely, related industries provide a context in which the competences are similar, knowledge is complementary and the potential for learning is substantial. In this sense, related knowledge provides the optimal cognitive distance between domains, maximizing effectiveness of learning and novelty potential (Nooteboom, 2000; Boschma and Frenken, 2011). As a result, in addition to its importance for diversification processes, relatedness is associated with a supportive regional environment that helps industries become more innovative, successful, and resilient (Martin and Sunley, 2015).

Yet, relatedness primarily fuels path-dependent development that may ultimately narrow regions’ competence bases (Grillitsch et al., 2018). This tendency is frequently related to lock-in scenarios, i.e., situations in which regions cannot escape their development paths because they have run out of further growth potential. Regional development therefore requires the presence (and sustaining) of sufficient unrelated diversity. Combining diverse and unrelated competences and knowledge is difficult and therefore rare. Nevertheless, it also carries potential for radical innovations and knowledge combinations with great novelty (Nooteboom, 2000). Thus, diversity increases the potential for such radical innovations and the substantial economic growth frequently associated with them.

Although the literature has, until recently, focused mainly on relatedness and diversity, recent studies add a third dimension to the mix. Inspired by Adam Smith, Hidalgo and Hausmann (2009) argue that the ever-increasing specialization of labor, capital, and knowledge make production systems more and more complex, as specialization implies growing heterogeneity and higher levels of interconnectedness. Put simply, more advanced products and services require a greater set of diverse expertise to be brought together, which in turn demands sophisticated coordination capabilities that are developed over time in a cumulative, path-dependent, and consecutive manner (Hidalgo and Hausmann 2009). Thus, those economies that are already active in complex activities also have the best opportunities and prerequisites of obtaining further competences in other and still more complex areas. By studying the evolution of countries’ export portfolios, Hidalgo and Hausmann (2009) provide solid empirical evidence in support of this argument.

Complementing the work of Hidalgo and Hausmann at the national level, Balland and Rigby (2017), Lo Turco and Maggioni (2020), Davies and Maré (2020), and Mewes and Broekel (2020) show that competences to produce and manage complex activities are unequally distributed at the subnational level as well. Few regions can specialize in the most complex production processes, and many regions focus on less complex or even simple activities. That is, whereas simple tasks can be performed almost anywhere, more complex ones have specific requirements that few places offer. Loosely speaking, any region can develop hunting and gathering, but only a few can specialize in advanced nanotechnology. The most striking differences in the ability to build competences in highly complex fields exist between rural and urban regions, with the latter being much more likely to be the location of firms and organizations operating at the complexity frontier. Prominently, Balland et al. (2020) show that the concentration of complex occupations, economic sectors, scientific fields, and technologies scales with urbanization.

The urban–rural divide and the relative scarcity of places that excel in complex activities are not merely interesting spatial patterns. In addition to providing the basis for future diversification, competences in complex activities also serve as attractive competitive advantages; the regions that are perceived to be the "most developed" also tend to be those in which the most complex economic activities of the moment are concentrated. For instance, during the period when car manufacturing was the pinnacle of human achievement, Detroit experienced outstanding economic growth. Before the Second World War, Berlin was one of the fastest-growing and wealthiest cities in the world; its nickname back then, 'Electropolis'. illustrates why: Berlin was home to firms such as Siemens AG and AEG, which were active in electrics and electronics—the most dynamic and complex technologies of that time (Killen, 2006). Today, California's Silicon Valley is the prime example for the co-concentration of competences in complex activities alongside economic prosperity.

At the national level, Hidalgo and Hausmann (2009), Ferrarini and Scaramozzino (2016), and Stojkoski et al. (2016) identify complexity as a driver of economic growth. Recently, Mewes and Broekel (2020) confirmed this at the regional level. Balland and Rigby (2017) show that these competences do not diffuse easily in space, implying that this type of advantage is geographically "sticky" and persistent. Consequently, regions with competences in complex activities enjoy competitive advantages that are rare and hard for others to imitate, and therefore economically valuable (Mewes and Broekel, 2020). Thus, complexity represents a crucial dimension that informs on regions' prospects for further upgrading of economic structures into activities of higher economic value generation and contemporary presence of highly profitable activities.

Based on these arguments, regional development processes are more likely to be sustainable and successful when they happen in an environment that is supportive in terms of relatedness, diversity, and complexity. These three dimensions of regions' industry structures are central to the contemporary understanding of how regions develop, and each plays a distinctive role: Diversity represents the potential for long-term development as well as novel and unrelated knowledge re-combinations. Relatedness captures the degree to which regions offer diversification potential that builds on existing strengths and is consequently both comparatively easier and potentially quicker to realize. Lastly, complexity approximates a regional economy's presence in more economically valuable activities and its capacity to develop such activities in the future.

**Limitations of the literature**

There is substantial empirical evidence indicating that each of the three dimensions supports regional diversification and economic growth. However, the extant literature remains underdeveloped in at least five aspects, which we will address in this contribution.

First, most of the empirical literature focuses on one or two of these three dimensions. For instance, much of the classical literature on urban growth examines the relationship between diversity and growth (Combes, 2000; Quigley, 1998). In their seminal contribution, Frenken et al. (2007) extend this view to include relatedness; their analysis reveals that regions characterized by related industrial structures experience better growth performance than others. This contribution of relatedness to regional growth is assessed by a range of subsequent studies, including Boschma and Iammarino (2009) and Morkutė et al. (2017). Mewes and Broekel (2020a) are one of the few that simultaneously consider relatedness and (technological) complexity. They find the latter to be more conducive to regional growth but leave the relationship with diversity unexplored. Notably, there is some evidence to suggest that relatedness and complexity are not independent of each other (Juhász et al., 2020) and that they are in fact complementary in promoting employment and GDP growth (Davies and Maré, 2021). However, the empirical literature on such interdependencies is just emerging and clearly more research is needed before we fully understand the interplay of relatedness, complexity and economic growth.

Second, many contemporary studies employ a related diversification framework. That is, they seek to explain the likelihood that regions will enter new economic and technological domains, using relatedness to existing strengths and competences in the region as the main explanatory variable (Boschma and Frenken, 2011; Mewes and Broekel, 2020b; Xiao et al., 2018). Alternatively, studies examine whether regions that are specialized in activities with high levels of relatedness (related variety) have greater economic growth than regions where core activities are less related (see e.g., Frenken et al., 2007; Ejdemo and Örtqvist, 2020; Barbieri et al., 2020). However, few studies investigate whether industries located in regions with related competences outperform industries in less supportive places.

Third, in addition to the lack of studies on industry-specific economic growth, the literature tends to focus on a limited set of dimensions of development. Most commonly, diversification success is approximated via growth in employment or significant jumps in regional employment shares; few studies look at other growth dimensions. For instance, Mewes and Broekel (2020a) assess complexity's impact on regional GDP growth and Hartmann et al. (2014) look at its relationship with income inequality. Although these studies broaden the view, there are still very few studies considering non-growth-related dimensions of development. Van den Berge et al. (2020) and Li et al. (2020) are notable examples of such studies that examine the role of relatedness in the transition to more sustainable industries. Clearly, much remains to be explored in terms of alternative dimensions of development.

Fourth, although the general importance of relatedness, diversity, and complexity is well established, their respective impacts on economic growth are likely dependent on the general conditions of the region and the industry in question. For instance, innovation in more complex industries requires more advanced learning and, consequently, activities in these industries are likely to benefit more from (local) knowledge spillover (Balland et al., 2020). The presence of diverse and heterogenous knowledge in a geographic vicinity is also of greater relevance in peripheral regions, as urban regions are almost universally characterized by higher levels thereof. In general, competition in urban regions for scarce resources is more intense, as reflected by higher rent prices and wages. These costs are less burdensome for complex industries with higher levels of added value than for simple industries. Consequently, the latter may face more growth obstacles in such places. Although geographers frequently highlight the place-dependence and place-sensitivity of economic development, little research in economic geography has examined such conditionalities in the effects of relatedness, diversity, and complexity.

Fifth, although these concepts are already well established in the literature, there is no consensus on how to capture them empirically or whether the employed empirical approximation fits the underlying theoretical arguments. In fact, empirical research typically relies on indicators that approximate multiple underlying processes and compress them into a single index. In many instances, researchers are particularly interested in the relatedness, diversity, and complexity—both in and of—regions' knowledge bases, as well as the knowledge-based processes underlying these dimensions. However, as we will argue in more detail in the next section, most studies seek to capture this by applying an industrial perspective, i.e., studying regions in terms of their composition of industries. This approach does capture a great deal of heterogeneity in regions' knowledge bases and is also rooted in data limitations. Nevertheless, it tends to overlook the fact that industries often comprise a range of heterogeneous activities; firms specialize in different activities and employ workers with distinct skill sets and abilities. Even though firms and workers may belong to the same industry, what they demand from their location may differ. Thus, the same regional conditions may be more or less valuable to different groups of workers within an industry. Consequently, to better understand the capacity of regions and their industries to develop and grow, we must focus much more on what their workers actually do, rather than what the industry produces as a whole. In short, following recent studies on occupational relatedness (Wixe and Andersson, 2017; Jara-Figueroa et al., 2018; Hane-Weijman et al., 2020; Dordmond et al., 2021), we propose to shift the analytical frame from the products that industries produce to the various activities in which they engage.

We argue in favor of a skill-oriented perspective on regional development that complements the more traditional, industry-based perspective. This argument is based on the recognition that skills transcend industries, for skills acquired in one industry are often used in others as well (Neffke and Henning, 2013). For instance, although the oil and gas and bio-technical industries produce completely different products, engineers in the two sectors frequently perform similar tasks and are, to some extent, interchangeable. Because skills and abilities are often associated with occupations, people working in the same occupation are more likely to perform similar tasks and therefore can more easily move between industries with minimal retraining. Thus, looking at industries' occupational composition can capture crucial heterogeneity in terms of competences, skills, and qualifications. This is particularly relevant when looking at the same industry in different locations, as each location may feature a distinct set of occupations. Therefore, it is often useful to examine spatial variations in the composition of occupations within industries (e.g., Fernández-Macías, 2012). Complementarly to this view is the analysis of regions' general occupational portfolios without considering the industrial dimension (see, e.g., Shutters et al., 2018). Although arguments can be made for each perspective's relevance to economic development, only empirical research can determine which view is the more insightful of the two.

**Opening the black box of relatedness, diversity, and complexity with an occupational perspective**

Much of the empirical evidence on the relevance of relatedness, diversity and complexity is based on empirical approximations that may reflect certain processes to larger degrees than others. By taking an occupation-based perspective, it is possible to design new indicators and modify existing ones to better reflect the knowledge- and competence-based arguments on which the theoretical frameworks in this literature typically rely. There is an emerging strand of literature on occupational relatedness which does precisely that (e.g., Wixe and Andersson 2017; Jara-Figueroa et al., 2018; Farinha et al., 2019), and a recent contribution by Hane-Weijman et al. (2020) also includes occupational complexity. However, the concept of diversity has so far not been examined through an occupational lens.

The literature on regional development has overwhelmingly been preoccupied with the composition of regions' activity profiles from an industrial perspective. Studies typically examine a region's diversity or specialization of industries, its diversification from specializing in certain industries to specializing in others, and the relatedness across different industries in a region. Although there is an associated body of literature examining other activities such as exports and patenting, such research has rarely looked beyond industries to consider the breadth of activities occurring within them.

No doubt, industries differ substantially in their dominant knowledge (domains), labor and capital requirements, as well as infrastructure. However, when describing a region's composition of competences and knowledge, another dimension—that of occupations—will paint a different picture. Industries are composed of various occupations, reflecting activities that are jointly necessary to produce the outputs that characterize the industry. For instance, the restaurant industry employs chefs, waiters, dishwashers—and in some regions, valets—as well as managers and accountants. The hotel industry also employs all of these, with the addition of housekeepers, bellhops, concierges, bartenders and even aerobics instructors. Looking at occupations (alongside or instead of industries) has two implications.

First, this perspective allows us to look inside industries to consider their *internal diversity* in terms of knowledge and competences. Occupations are a better indicator of what individuals actually do as well as what types of competences and knowledge they possess. From the example above, occupational diversity is generally higher in the hotel industry than in the restaurant industry. This is because hotels employ all of the same occupations as restaurants—as well as additional occupations that are not typically associated with restaurants. This reflects a larger variety of tasks and knowledge involved in operating a hotel compared to a restaurant. Occupational diversity will also vary among regions, even within a single industry. For instance, the casino-adjacent hotel industry in Las Vegas employs croupiers, stand-up comedians, exotic dancers and even gondoliers, whereas the hotel industry in Salt Lake City—where gambling is illegal—probably does not. Such diversity exists in other industries as well, and must be taken into account—because if innovation is the result of new combinations of different types of knowledge, the diversity of knowledge within industries also has important implications for regional innovation. The potential for new combinations of knowledge emerging through within-industry interactions in the legal services industry, which mainly employs lawyers, is quite different compared to oil extraction, which typically involves large integrated companies employing geologists, chemists, drillers, welders, accountants, lawyers, communicators, salespeople, etc. We can therefore predict that industries in regions that encompass a diversity of different occupations will benefit from Jacobs externalities to a greater extent than industries in regions with less diversity. Therefore, we suggest complementing the industry level with a perspective on occupational diversity within industries. Notwithstanding this suggestion, it can also be interesting to consider the general occupational diversity in a region. However, regional occupational diversity may be of varying importance depending on the economic activity in focus and, in some instances, may show overlap with the dimension of relatedness. Accordingly, some caution is recommended when including more than one such dimension in empirical studies.

Second, using occupational information also implies a somewhat different perspective on relatedness. Occupations within an industry can be *related* as well as unrelated in terms of the skills which they require. In the previous example, the oil extraction industry involves various types of engineers: geological engineers, chemical engineers, physical engineers, construction engineers and so on. However, it also involves occupations which are unrelated to engineering, such as accounting, law and communications. Crucially, occupational relatedness may differ substantially from industrial relatedness. While the IT industry may generally have little in common with oil extraction, the oil industry does employ a significant number of IT workers, e.g., in internal IT departments or as programmers or modelers in exploration or operations departments. Thus, when measured at the industry level, IT sectors in an oil and gas-dominant region may appear to be surrounded by little related variety. However, when considering the occupational composition of the region, one might find related activity related to the IT industry occurring within other industries in that region.

The previous point underlines why the occupational perspective is complementary to the industrial perspective. Although such a skill- and occupation-centric perspective has been applied in previous works (Neffke et al., 2011; Shutters et al., 2018), few studies have considered it complementary to an industrial one (Wixe and Andersson, 2017; Jara-Figueroa et al., 2018). Yet, looking at both dimensions jointly is one way of opening up the so-called "black box" of relatedness to a certain degree. In the example of the oil extraction and IT industries, the two industries encompass some occupations that are highly related in terms of their knowledge requirements and work tasks. However, the rest of the industries' occupations are clearly not. Besides very different technology domains and distinct historical development paths, their differences also show in specific locational requirements. The oil extraction industry (apart from its management) will tend to prosper in regions with sufficient production facilities, easy transportation, access to harbors, the presence of supplier industries, etc. In addition, the labor market must supply specialized human capital with skills in engineering and specific manual labor tasks. In contrast, the IT industry will be supported by urban locations with sufficient amenities, industrial diversity, and excellent ICT infrastructures. From these perspectives, there would appear to be few reasons for the two industries to be co-located. Thus, commonly employed measures of relatedness based on labor mobility and co-location of employment will conclude that these industries are unrelated and will fail to detect the relatedness between some sub-sets of their activities. Consequently, comparing the two dimensions in empirical investigations will inform us about which of these dimensions of relatedness (knowledge or input/infrastructure/location) is more conducive to the industries' development.

In contrast to relatedness, there is not yet a consensus on the best approach for empirical measurement of economic and technological *complexity*. A number of approaches rely on patent data (Balland and Rigby, 2017; Broekel, 2019; Fleming and Sorenson, 2001), thereby favoring the technological dimension, which is interesting for many investigations. However, this approach does not capture the full breadth of economic activity in an industry. For this reason, when it comes to the complexity of sectors and industries, the literature primarily relies upon Hidalgo and Hausmann's (2009) economic complexity index, which is based on the spatial co-concentration of industries' employees and the method of reflection. Although it is commonly applied, this indicator has limitations. For example, it is based on the very same structures (regions' industrial portfolios) that it seeks to explain (growth and diversification). Furthermore, the approach also measures complexity in an indirect fashion (Broekel, 2019). Here, too, the occupational perspective offers an interesting alternative. Drawing upon an extensive body of literature that examines the tasks associated with occupations, researchers quantify the complexity of occupations based on the sophistication of those tasks. In particular, activities that require complex problem-solving, continuous learning, information retrieval, communication skills and management responsibilities—and that are nonroutine—are seen as complex (see, e.g., Ederer et al., 2015; Caines et al., 2017; Lo Turco and Maggioni, 2020). In the context of the present paper, such a perspective allows for assessing the *complexity of industries* by aggregating the complexity of their occupations.[1] Some industries are predominated by simple jobs with low requirements for complex problem-solving. For instance, supermarket retail involves a diversity of fairly simple jobs with low knowledge requirements, such as warehouse workers, shelf stockers and cashiers. Although supermarkets also involve more complex occupations, such as data analysts and marketers, these make up a small proportion of the total staff. In most cases, these jobs are also concentrated in certain regions with headquarters functions. Conversely, the legal services industry may not be very diverse, but it is driven by complex tasks and is therefore attractive for regions to specialize in.

The complexity of industries matters for economic growth because more complex tasks that are usually reflected in occupations require more intensively trained specialists, who are relatively rare. These occupations, and thus the regions in which they are located, will be able to extract rents by facing lower levels of competition. Furthermore, the greater knowledge intensity of these occupations

[1] Note that this differs from the definition of occupational complexity as used in Hane-Weijman et al., (2020). These authors use a measure based on the method of reflection (Hidalgo and Hausmann, 2009) applied to occupational data.

implies that they are more potent sources of knowledge; as a result, the presence of complex industries fuels cross-fertilization and knowledge externalities that in turn spur further economic growth. These industries' specialized and complex knowledge also contributes to their respective regions' future diversification into other— and potentially even more complex—activities. Thus, these industries represent a crucial basis for sustained long-term growth (Hidalgo and Hausmann, 2009). This discussion shows that the economic relevance of economic complexity is largely (but not exclusively) based on human capital arguments. This, in turn, is expressed in the sophistication and complexity of the tasks that individuals are able to perform. Although complexity may be quantified through indirect approaches based on industry co-concentration such as the one proposed by Hidalgo and Hausmann (2009), there are more direct and precise measures available when taking an occupational perspective.

Although a region's diversity, relatedness and complexity of occupations individually matter for its development, there are important interdependencies between them (Juhász et al., 2020). Building competences in more complex tasks is a long and demanding process that typically requires gradual diversification from less complex into more complex tasks. This process is easier when related occupations are present. Regions with a large variety of occupations are also more likely to contain unrelated knowledge combinations, which may occasionally spur the development of competence in unexpected areas.

At this point, it is important to acknowledge certain nuances involved when working with occupational data. Generally, using occupational information allows researchers to assess the diversity and relatedness of regions' knowledge bases with greater precision than what they can do by just looking at industries. Furthermore, in terms of relatedness, occupational data can help disentangle knowledge-related effects from more pecuniary ones related to the sharing of infrastructure and institutions, the presence of specialized suppliers, and labor market pooling effects. However, there are several shortcomings. First, the validity of many occupational indicators has not yet been widely established. Therefore, it remains unclear to what extent measures of diversity, relatedness and complexity developed for industries can also be applied to occupations within industries. Crucially, most measures utilize information of standard classifications of occupations. Consequently, such measures are only as good as the occupational classifications on which they are based. These classifications differ across locations, e.g., between Europe and the US. Consequently, using this data usually requires the application of concordances and crosswalks that may potentially introduce systematic biases. Second, occupational data are often hard to come by, especially at the industry-region level. Some of the concepts discussed above require detailed longitudinal data at the individual level, which is even less accessible. Finally, the study of occupational complexity depends on survey data, which has only been gathered from limited samples of employees in select countries. Thus, from the data available, it may be unclear, for example, to what extent the complexity of an occupation in a particular country context also applies in other contexts. Nevertheless, when available, the benefits of occupational information clearly outweigh these drawbacks.

Next, we will illustrate how occupational diversity, relatedness, and complexity can be empirically measured and how their relevance for economic development can be assessed, using labor market regions in Norway as examples.

## MEASURING THE OCCUPATIONAL COMPOSITION OF REGIONAL INDUSTRIES

The literature offers various approaches to empirically capture relatedness, diversity, and complexity at different levels. Below, for each dimension, we present one of the most promising (in the eyes of the authors) contemporary measurement approaches. We therefore focus on the approaches that leverage the availability of occupational data. We strongly suggest that future research thoroughly compare different measures from both theoretical and empirical perspectives.

**Measuring occupational diversity**

To capture occupational diversity, one approach is to use a fractionalization measure based on, e.g., Alesina et al. (2003). This index is commonly used in studies of diversity in other contexts, such as ethnic or birthplace diversity. Applied to the study of occupational diversity at the industry-region level, the construction of this index is as follows:

$$DIVERSE_{irt} = 1 - \sum_{o=1}^{O} s_{oirt}^{2} \qquad \text{(Eq. 1)}$$

In this measure, $s$ is the proportion of employees in industry $i$, in region $r$, at time $t$ that work in occupation $o$; $O$ is the number of different occupations represented in that industry-region in the same year. The index ranges between 0 and 1. A maximum diversity value nearing 1 reflects a situation where an industry-region consists of an equal number of people in each occupation. A value of 0 reflects a situation where all the employees have the same occupation.

**Measuring occupational complexity**

In contrast to a growing number of studies relying on an indirect and second-nature approximation of economic complexity using the approach of Hidalgo and Hausmann (2009), considering the occupational dimension gives access to a direct first-nature approach of measuring industries' complexity. That is, we propose to follow insights from labor economics that differentiate the difficulty of employees' everyday tasks to assess the complexity of economic activities (see, e.g., Ederer et al., 2015). This measure of task or skill complexity does not rely on information on industries' spatial distributions as in that of Hidalgo and Hausmann (2009). This is particularly relevant in small economies or those where industrial distribution is strongly shaped by natural resources and topography, as in Norway.

Lo Turco and Maggioni (2020) propose assessing the complexity of industries by means of the average complexity of the occupations an industry employs. This involves examining occupations' task complexity, which is well established in labor economics. Caines et al. (2017) have developed an index of the complexity of individual occupations, which represents this idea. Their index is based on US O*NET data and represents a weighted aggregation of 35 variables as part of the O*NET subsections "Abilities," "Skills," and "Generalized Work Activities." The variables specifically capture to what degree occupations involve solving complex problems, finding original solutions, applying critical thinking, analyzing data and information, etc. Using a principal components analysis, Caines et al. (2017) combine these variables into a single index, which is a continuous number theoretically ranging from 0 to 1.

Unfortunately, this complexity measure is based on a modification to the 1990 US Census occupational codes, which poses some challenges for its applications to European data. However, the US Bureau of Labor provides a crosswalk from the US SOC classification system to the European ISCO classification.[2] Because this is a one-to-many matching (967 SOC to 424 four-digit ISCO occupations), the corresponding SOC-based complexity values are aggregated by averaging across all SOC codes associated with one ISCO code.[3] These can be matched to the industry–occupation data. In a final step, complexity values at the four-digit NACE industry level can be estimated by weighting the occupational complexity values with the corresponding region- and year-specific shares of these occupations in the corresponding NACE industry. A drawback of this approach is that although the measure of complexity

[2] https://www.bls.gov/soc/soccrosswalks.htm.
[3] We have experimented with alternative approaches, such as using the minimum and maximum, but the differences are marginal.

varies over time due to changes in the occupational composition of industries in regions, the complexity values associated with the underlying occupations requires extensive survey data. Therefore, these values are measured at one moment in time—namely in the year 2019—in the empirical illustration below.

**Measuring occupational and industrial relatedness**

The occupational relatedness of industries can be approximated using an approach first developed to study industrial relatedness. By now, many studies of relatedness rely on the revealed skill relatedness approach based on labor mobility (e.g., Neffke and Henning, 2013; Fitjar and Timmermans, 2017). In this approach, two industries are related when labor mobility between them is higher than what would be statistically predicted. That is, two industries require related skills if workers often change from one industry to the other. A similar approach can be used to study occupational relatedness, wherein the difference is the underlying relatedness matrix. In this case, two occupations requiring related skills are indicated by workers frequently changing between them.

The first step of developing an occupational relatedness measure is to construct the occupational relatedness matrix using register data that identify individuals who change their occupation from one year to the next. We can count the number of individuals changing from occupation *o* to occupation *k* and compare this to the overall number of individuals starting to work in occupation *k* or leaving work in occupation *o*. When we observe greater mobility between any pair of occupations than what would be statistically predicted based on the overall tendency to take up or leave work in these occupations, we can consider the occupations to be related. This measure can further be projected to the industry-region level by calculating the related density for each industry-region (Hidalgo et al., 2007). Many studies focus on relatedness to activities in which the region has a revealed comparative advantage, i.e., where the location quotient is above one. However, a drawback of this approach is the rather arbitrary cut-off of observations with shares lower than the national average. Therefore, it can also be a good alternative to directly rely on occupations' employment shares.

This method requires a second step, where for each occupation *o* in region *r*, we weight the regional employment share of any other occupation *k* present in the same region with the corresponding relatedness measure in year *t*. Subsequently, all weighted employment shares are summed, yielding the related density of occupation *o* in region *r* and year *t*. In a final step, this occupation-specific measure is projected to the industry level by multiplying it by the region-specific occupation shares of industry *i*.

## EMPIRICAL ILLUSTRATION

To gain some understanding of these measures and to provide an initial indication of how various occupational characteristics are distributed across industries, we examine the distribution of occupational diversity, relatedness, and complexity in Norwegian industry-regions. We rely on individual-level registry data linked to establishments (linked employer–employee data) from Statistics Norway. The data contain detailed longitudinal information on workplace, industry and work location of all individuals employed in Norway's private sector for the period 2009–2016. The data cover all inhabitants over age 16 who are employed by private establishments and include a range of information about individual workers and establishments. From this registry, we first build a data set of the numbers of workers in each occupation per industry in each economic region of Norway; industries are identified at the four-digit NACE level using the SNI2007 industry classification system. Occupations are defined using the seven-digit level of the Norwegian SSYK98, which is consistent with the international ISCO-88 standard. For all estimations, we rely on the four-digit ISCO level.

From this data, we construct the measures defined above. We calculate all measures for industries in labor-market regions, which in Norway correspond to the statistical category of “economic regions.” This allows for the measures of occupational diversity, complexity and relatedness to vary across different regions within the same industry, reflecting different regional characteristics and positions within each industry. Economic regions are officially defined by Statistics Norway and represent NUTS-4 regions at the level between the county and municipality, which are the official political and administrative units. Following Gundersen and Juvkam (2013), we merge integrated labor markets, which gives 78 regions in total. Below, we present the three novel indices for occupational diversity, relatedness and complexity in Norwegian industry-regions.

Table 1 shows the top 10 industries for occupational diversity in Norwegian regions. Various manufacturing industries are at the top of the list, alongside some mining industries. This reflects the division of labor in many manufacturing industries, where each worker specializes in a distinct part of the production process.

| | NACE | INDUSTRY | DIVERSITY |
|---|---|---|---|
| **1** | 28.91 | Manufacturing of machinery for metallurgy | 0.928 |
| **2** | 07.29 | Mining of other non-ferrous metal ores | 0.906 |
| **3** | 28.95 | Manufacturing of machinery for paper/paperboard production | 0.901 |
| **4** | 64.20 | Activities of holding companies | 0.899 |
| **5** | 20.30 | Manufacturing of paints, varnishes and coatings | 0.895 |
| **6** | 26.60 | Manufacturing of electromedical/electrotherapeutical equipment | 0.884 |
| **7** | 21.10 | Manufacturing of basic pharmaceutical products | 0.880 |
| **8** | 23.51 | Manufacturing of cement | 0.868 |
| **9** | 08.91 | Mining of chemical/fertilizer minerals | 0.866 |
| **10** | 28.15 | Manufacturing of bearings, gears, gearing/driving elements | 0.858 |

Table 1: Occupational diversity in Norwegian industries

Table 2 shows the top 10 industries with respect to average occupational complexity. The most complex industries in Norway include those related to oil and gas, research and development, and information technology. Architectural activities is another industry characterized by high average occupational complexity. Overall, sorting by average occupational complexity provides similar results as other classifications of industries (such as those based on education levels or knowledge intensity), which is reassuring for the face validity of the indicator. However, occupational complexity also reveals the sophistication of—in this case—the Norwegian oil and gas sector. Whereas oil and gas production are relatively simple extractive activities in some countries, operations in Norway’s deep-water offshore fields pose major technological challenges that have led to the development of a highly complex knowledge-based industry (Thune et al., 2019). Such contextual differences may be masked by classifications based on industry codes (e.g., the OECD definitions of high-technology industries).

| | NACE | INDUSTRY | COMPLEXITY |
|---|---|---|---|
| **1** | 35.23 | Trade of gas through mains | 72.282 |
| **2** | 71.11 | Architectural activities | 69.945 |
| **3** | 06.10 | Extraction of crude petroleum | 64.370 |
| **4** | 72.11 | R&D on biotechnology | 63.574 |
| **5** | 72.19 | Other R&D on natural sciences etc. | 63.237 |
| **6** | 62.02 | Computer consultancy activities | 61.720 |
| **7** | 62.09 | Other information technology/computer services activities | 61.622 |
| **8** | 62.01 | Computer programming activities | 61.146 |
| **9** | 62.03 | Computer facilities management activities | 61.114 |
| **10** | 71.12 | Engineering activities/related technical consultancy | 59.757 |

Table 2: Occupational complexity in Norwegian industries

Table 3 shows the top 10 industries in terms of average occupation-related density in regions of Norway. This list covers a wide range of different industries, from those consisting of relatively simple low-end service occupations (mainly in retail) to those employing high-end specialists in highly complex occupations such as information technology, finance, and oil and gas. Several industries are also in various parts of the value-chain for food, from production (aquaculture) to processing and manufacturing to retail.

| | NACE | INDUSTRY | OCCUPATIONAL RELATEDNESS |
|---|---|---|---|
| **1** | 47.21 | Retail sale of fruit and vegetables | 0.046 |
| **2** | 47.26 | Retail sale of tobacco products | 0.038 |
| **3** | 06.10 | Extraction of crude petroleum | 0.037 |
| **4** | 47.63 | Retail sale of music and video recordings | 0.034 |
| **5** | 66.11 | Administration of financial markets | 0.033 |
| **6** | 10.12 | Processing and preservation of poultry and meat | 0.032 |
| **7** | 03.21 | Marine aquaculture | 0.031 |
| **8** | 62.02 | Computer consultancy activities | 0.030 |
| **9** | 10.82 | Manufacturing of cocoa, chocolate etc. | 0.030 |
| **10** | 62.03 | Computer facilities management activities | 0.030 |

Table 3: Industries with the highest average occupation-related density

For comparison, Table 4 shows the top 10 industries with the highest industry-related density when using the traditional industrial relatedness measure, based on skill relatedness calculated from labor mobility across industries. In this case, the list is heavily dominated by industries in the oil and gas value chain, including general oil service providers, specialist manufacturers of machinery and ships, and engineering services. However, food-related industries are present here as well, reflecting another large sector in the Norwegian economy.

There is some overlap between the industries with the highest levels of occupational and industrial relatedness. However, at the same time, there are also noticeable differences between the two lists. Although industries in the oil and gas supply chain clearly stand out as having the highest levels of industrial relatedness, this is not so for occupational relatedness. Consequently, although the oil and gas-related industries appear to be at the core of the labor mobility network across industries, this does not apply to all of its occupations. Thus, shifting the focus to occupational relatedness reveals that Norway's regional labor markets may be less shaped by the oil and gas industries than what a simple focus on industry relatedness would predict.

| | NACE | INDUSTRY | INDUSTRY RELATEDNESS |
|---|---|---|---|
| **1** | 09.10 | Supply for petroleum/natural gas extraction | 0.317 |
| **2** | 28.92 | Manufacturing of machinery for mining/quarrying/construction | 0.242 |
| **3** | 07.10 | Mining of iron ores | 0.220 |
| **4** | 25.40 | Manufacturing of weapons and ammunition | 0.110 |
| **5** | 30.11 | Building ships and floating structures | 0.105 |
| **6** | 71.12 | Engineering activities/related technical consultancy | 0.095 |
| **7** | 47.11 | Non-specialist stores with food, beverages, etc. | 0.094 |
| **8** | 69.20 | Accounting/bookkeeping/tax consultancy | 0.090 |
| **9** | 16.21 | Manufacturing of veneer sheets, wood-based materials | 0.083 |
| **10** | 10.20 | Processing and preservation of fish, etc. | 0.081 |

Table 4: Industry with the highest average industry-related density

**HOW ARE COMPLEXITY, RELATEDNESS AND DIVERSITY RELATED TO REGIONAL DEVELOPMENT?**

To further illustrate the use of the empirical measures, we assess their relation to long-term regional employment growth in Norway. We start from the initial conditions in 2009 and examine their associations with the subsequent development of employment in Norwegian industry-regions. We use the NACE four-digit level for industries and economic regions corresponding to functional labor markets. The dependent variable is the corresponding average annual growth rates of employment for the period 2009–2014. We fit an OLS regression model in which employment growth is explained by the initial values of employment ($EMPL_{ir}^{2009}$) as well as the set of independent variables with their values in the year 2009 ($\mathbf{X}_{ir}^{2009}$). The vector $\mathbf{X}_{ir}^{2009}$ contains the variables of interest—diversity, relatedness, and complexity—as well as relevant controls such as population density ($POP.DEN_r^{2009}$) and the location quotient of the focal industry ($LQ_{ir}^{2009}$). We use multiway clustered standard errors to account for potential clustering in the industrial and regional dimensions. To further reduce potential biases induced by industry-wide shocks or specificities, we add industry-fixed effects at the two-digit NACE level. All variables are log-transformed (indicated by an 'l' preceding the variables' names). We increase the robustness of the estimations by excluding industries with fewer than five employees and focusing on the private sector. Table 5 shows descriptive statistics for the variables included in the analysis. Figure 2 shows the pairwise correlations between all variables included in the empirical models for the analytical sample.

As discussed in the theory section, we expect that the importance of relatedness, complexity, and diversity will vary between simple and complex industries. To explore this, we split the data (industry-regions) into three samples based on industries' levels of complexity. The first sample contains all industry-regions in the lower third of the complexity distribution; the second sample contains those in the middle third, and the last sample contains the upper third.

| VARIABLE | N | MEAN | ST. DEV. | MIN | PCTL(25) | PCTL(75) | MAX |
|---|---|---|---|---|---|---|---|
| **EMPL** | 6,276 | 192.295 | 592.261 | 8.000 | 30.000 | 144.2 | 12,890 |
| **POP.DEN** | 6,276 | 44.136 | 54.808 | 2.211 | 8.165 | 63.990 | 206.304 |
| **LQ** | 6,276 | 4.165 | 12.991 | 1.002 | 1.748 | 3.030 | 424.011 |
| **IND.REL** | 6,276 | 0.079 | 0.064 | 0.000 | 0.030 | 0.114 | 0.810 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **DIVERS** | 6,276 | 0.574 | 0.161 | 0.000 | 0.462 | 0.696 | 0.938 |
| **COMPLEX** | 6,276 | 40.618 | 8.765 | 7.086 | 36.124 | 45.161 | 71.806 |
| **OCC.REL** | 6,276 | 0.016 | 0.010 | 0.000 | 0.008 | 0.022 | 0.073 |

Table 5: Descriptive statistics

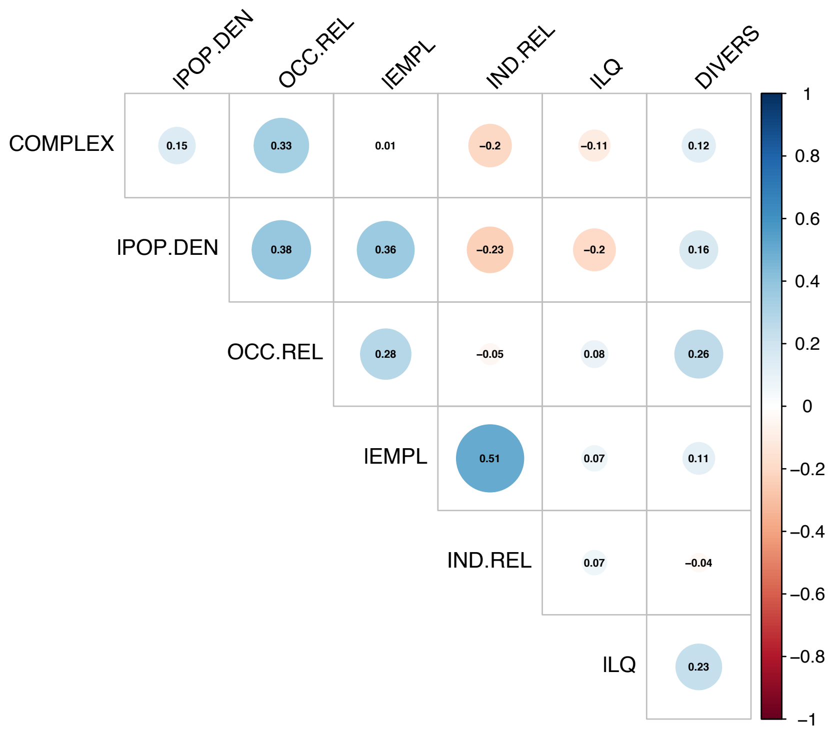


Figure 1: Correlations

Table 6 presents the results of the regressions. In the basic model (column 1), the initial levels of the dependent variable *IEMPL* are significantly negatively related to growth rates, following a commonly observed pattern in the literature. Moreover, the estimates confirm positive development in more urbanized regions, as population density (*POP.DENS*) is associated with higher employment growth rates. In the basic model, we do not find any significant relationship between location quotient and the dependent variables. This is quite unexpected, given the large attention that specialization and agglomeration have received in the traditional literature on regional growth (e.g., Combes, 2000). The explanatory power of the variable is also not accounted for by other variables, because it is not correlated with the absolute employment or with any of the relatedness variables (Figure 1).

With respect to the variables of interest, we observe a significantly positive coefficient of industrial relatedness (*IND.REL*). This suggests that industries co-located with other related industries outgrow those located in less supportive surroundings. This confirms a main proposition of the literature on relatedness and supports the concept of regions' path-dependent development (Boschma, 2017). We find no significant relationship between employment growth and industries' diversity of occupations nor their level of complexity. The absence of such a relationship may be attributable to their influence being conditional on the levels of complexity. By splitting the data into different subsamples (see previous section), we shed on light on the potential explanations for this absence.

We have argued that *IND.REL* should reflect pecuniary externalities related to labor market pooling, specialized suppliers, and sharing of infrastructure to a greater extent than *OCC.REL*. In contrast, *OCC.REL* can be expected to capture knowledge spillover effects more directly than *IND.REL*. Therefore, *IND.REL* can be expected to be more relevant to industrial growth in complex industries (Balland et al., 2020).

In contrast, less complex industries (e.g., manufacturing) often prefer surroundings with supportive infrastructures and relatively cheap industrial sites (Keeble et al., 1983). In Norway, for many manufacturing industries, proximity to hydropower facilities is crucial because it ensures access to cheap energy. For such industries, competitive advantages may arise from pecuniary externalities. Although this prediction appears plausible, it is not supported by the analysis. Table 7 gives insights into the relationships between relatedness and employment growth when splitting the sample into simple, average and complex industries, respectively. Generally, manufacturing industries are more likely to fall into the first or second category, whereas knowledge-intensive activities are most often equated with complex tasks. Neither *OCC.REL* nor *IND.REL* have significant coefficients in any models for this sub-sample analysis, which casts some doubt on complexity being the decisive dimension here.

Although it does not have a significant relationship with employment growth for the full sample (Table 6), occupational diversity (*DIVERSE*) does have a significantly positive coefficient when splitting the observations by their level of complexity. In this case, diversity appears to support growth for industries with an average level of complexity (Table 7), which suggests that such industries relying on a diversity of skills, competences, and knowledge (as expressed in the diversity of their occupations) grow more than those with less diversity in their occupations. Industries of medium complexity are largely manufacturing- or wholesale-based. Therefore, we interpret diversity as a reflection of industries concentrating many of their activities in the same places. Concentrating many activities within the same industry will translate into a diverse set of activities that have, on average, a medium level of complexity. The observed positive effects are likely to represent the benefits associated with the co-location of headquarters, R&D labs, and production units in the same places. Especially in manufacturing sectors, co-location smoothens management processes, facilitates knowledge spillover, and supports easier communication (Ivarsson and Alvstam, 2017).

In contrast to relatedness and diversity, we do not observe a significant coefficient for complexity (*COMPLEX*) in any of the models. Accordingly, this dimension appears to be of little relevance to employment growth in Norwegian regions during this period.

| | MODEL 1 | MODEL 2 | MODEL 3 | MODEL 4 |
|---|---|---|---|---|
| **LEMPL** | -0.012*** | -0.012*** | -0.012*** | -0.013*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| **LPOP.DEN** | 0.007*** | 0.007** | 0.007*** | 0.006** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| **LLQ** | -0.006 | -0.006 | -0.006 | -0.006 |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| **IND.REL** | 0.085* | 0.085* | 0.090* | 0.092* |
| | (0.042) | (0.042) | (0.041) | (0.042) |
| **DIVERS** | | -0.002 | -0.004 | -0.003 |

| | | | | |
|---|---|---|---|---|
| | | (0.013) | (0.013) | (0.014) |
| **COMPLEX** | | | 0.000 | 0.000 |
| | | | (0.000) | (0.000) |
| **OCC.REL** | | | | 0.280 |
| | | | | (0.248) |
| **NUM. OBS.** | 6226 | 6226 | 6226 | 6226 |
| **$R^2$ (FULL MODEL)** | 0.052 | 0.052 | 0.052 | 0.052 |
| **$R^2$ (PROJ MODEL)** | 0.010 | 0.010 | 0.011 | 0.011 |
| **ADJ. $R^2$ (FUL MODEL)** | 0.041 | 0.041 | 0.041 | 0.041 |
| **ADJ. $R^2$ (PROJ MODEL)** | -0.000 | -0.001 | -0.001 | -0.000 |
| **NUM. GROUPS: NACE2** | 65 | 65 | 65 | 65 |
| **$^{***}P < 0.001$; $^{**}P < 0.01$; $^{*}P < 0.05$** | | | | |

Table 6: Estimated results – main model

**STATISTICAL MODELS**

| | Simple | Simple | Simple | Simple | Average | Average | Average | Average | Complex | Complex | Complex | Complex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **LEMPL** | -0.018$^{***}$ | -0.018$^{***}$ | -0.018$^{***}$ | -0.018$^{***}$ | -0.005 | -0.006 | -0.005 | -0.005 | -0.015$^{***}$ | -0.015$^{**}$ | -0.016$^{**}$ | -0.016$^{**}$ |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.004) | (0.004) | (0.003) | (0.003) | (0.005) | (0.005) | (0.005) | (0.005) |
| **LPOP.DEN** | 0.010$^{**}$ | 0.010$^{**}$ | 0.010$^{*}$ | 0.010$^{*}$ | 0.005 | 0.004 | 0.005 | 0.005 | 0.007 | 0.007 | 0.006 | 0.006 |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.003) | (0.003) | (0.003) | (0.003) | (0.004) | (0.005) | (0.004) | (0.004) |
| **LLQ** | -0.005 | -0.004 | -0.005 | -0.005 | -0.012$^{*}$ | -0.013$^{*}$ | -0.013$^{*}$ | -0.013$^{*}$ | -0.003 | -0.002 | -0.002 | -0.002 |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.006) | (0.006) | (0.005) | (0.006) |
| **IND.REL** | 0.095 | 0.098 | 0.100 | 0.106 | 0.027 | 0.024 | 0.023 | 0.021 | 0.173 | 0.177 | 0.180 | 0.182 |
| | (0.056) | (0.057) | (0.055) | (0.056) | (0.048) | (0.048) | (0.047) | (0.048) | (0.104) | (0.105) | (0.107) | (0.111) |
| **DIVERS** | | -0.018 | -0.018 | -0.021 | | 0.044$^{*}$ | 0.044$^{*}$ | 0.045$^{*}$ | | -0.026 | -0.024 | -0.024 |
| | | (0.020) | (0.020) | (0.020) | | (0.018) | (0.018) | (0.018) | | (0.025) | (0.025) | (0.025) |
| **OCC.REL** | | | 0.177 | 0.138 | | | -0.184 | -0.178 | | | 0.546 | 0.538 |
| | | | (0.515) | (0.495) | | | (0.551) | (0.556) | | | (0.363) | (0.383) |
| **COMPLEX** | | | | 0.001 | | | | -0.001 | | | | 0.000 |
| | | | | (0.001) | | | | (0.002) | | | | (0.001) |
| **NACE-2 FE** | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| **CLUSTERED STD. ERR. NACE-4** | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| **CLUSTERED STD. ERR. REGION** | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| **NUM. OBS.** | 2074 | 2074 | 2074 | 2074 | 2077 | 2077 | 2077 | 2077 | 2075 | 2075 | 2075 | 2075 |
| **$R^2$ (FULL MODEL)** | 0.101 | 0.101 | 0.101 | 0.101 | 0.067 | 0.070 | 0.070 | 0.070 | 0.064 | 0.064 | 0.065 | 0.065 |
| **$R^2$ (PROJ MODEL)** | 0.019 | 0.019 | 0.019 | 0.019 | 0.010 | 0.013 | 0.013 | 0.014 | 0.010 | 0.011 | 0.012 | 0.012 |
| **ADJ. $R^2$ (FULL MODEL)** | 0.075 | 0.075 | 0.074 | 0.074 | 0.041 | 0.043 | 0.043 | 0.042 | 0.035 | 0.035 | 0.035 | 0.035 |
| **ADJ. $R^2$ (PROJ MODEL)** | -0.010 | -0.010 | -0.010 | -0.010 | -0.017 | -0.015 | -0.015 | -0.016 | -0.021 | -0.021 | -0.020 | -0.021 |
| **NUM. GROUPS: NACE2** | 55 | 55 | 55 | 55 | 54 | 54 | 54 | 54 | 60 | 60 | 60 | 60 |

**$^{***}P < 0.001$; $^{**}P < 0.01$; $^{*}P < 0.05$**

Table 7: Determinants of employment growth in simple, average, and complex industries

## CONCLUSION

The literature in regional development is generally in agreement that knowledge is an important driver of the economic growth of industries and regions. It is also the consensus that the sheer existence and magnitude of knowledge (including skills), as well as the co-location of competent actors, are not sufficient to explain growth differentials among regions. In fact, it is increasingly argued that the composition and quality of knowledge in regions is at least as crucial as the magnitude of their availability in this regard. However, we still know little about this dimension of regions' knowledge and skill portfolios. The present chapter focuses on three dimensions, namely diversity, relatedness and complexity, which are argued to be crucial for explaining regional growth. Although each of these dimensions has already been featured in empirical studies, the literature still has important limitations in several aspects, which we discussed in this chapter:

First, most existing studies tend to focus on a single dimension of a region's knowledge composition or, at most, consider two dimensions at a time. Few studies include all three (or even more) dimensions simultaneously. However, because the different dimensions are intertwined, it is important both to control for other dimensions and to consider the possibility of interactions between them. Second, the literature has so far been predominantly concerned with regional diversification. Fewer studies target other economic development outcomes. This makes it difficult to answer whether more diverse, related, and complex regions actually experience better economic development in a broader sense. Third, the literature that actually addresses developments apart from diversification also tends to have a rather limited view on this. Existing studies mainly explore the drivers of employment growth, largely neglecting equally important aspects such as sustainability, prosperity, and inequality. Fourth, empirical research often does not consider whether—and if so, how—the impact of diversity, relatedness and complexity varies across different contexts. For instance, growth processes are likely to be distinctively heterogeneous for cities and rural regions, for different institutional contexts, for specific countries, and for various industries. Finally, there is still no consensus on how to empirically approximate any of these dimensions. In particular, the three dimensions seem to be particularly related to regions' and industries' occupational composition. However, occupational data have been used to a very limited extent (e.g., Wixe and Andersson, 2017; Jara-Figueroa et al., 2018; Farinha et al., 2019; Hane-Weijman et al., 2020).

To illustrate these points, we conducted an empirical study using data on Norwegian regions and the industries and occupations within them. The analyses consider diversity, relatedness and complexity simultaneously and assess their contributions to employment growth. In the course of the analyses, we presented promising ways of empirically approximating the different dimensions. We also investigated the extent to which their influence varies between different types of industries (simple and complex) and between distinct geographical contexts (urban and rural). Moreover, we follow Wixe and Andersson (2017), Farinha et al. (2019), and Whittle and Kogler (2020) and open the black box of relatedness. Specifically, we differentiate between industrial and occupational relatedness, which allows for disentangling mechanisms related to knowledge spillovers and pecuniary externalities. In sum, the analysis underlines and addresses crucial aspects of the previous theoretical discussion and presents ideas on how to address some of the existing limitations in empirical set-ups.

Yet, the empirical analysis suggests that in the Norwegian context, the three dimensions (diversity, relatedness, and complexity) and the two distinct types of relatedness (occupational and industrial) provide few explanations for regional employment growth. Greatly contrasting with the theoretical predictions, we find empirical evidence only for the idea that pecuniary externalities arise from the presence of related industries, which translates into a positive association with employment growth. The empirical exercises thereby convey two important messages. Firstly, they clearly demonstrate the possibilities that are opened by working with occupational information alongside the traditional focus on industries. This approach allows for considering novel indicators that describe the structure of regional knowledge bases in a more detailed manner and from alternative perspectives. Most importantly, it opens the door to assessments of the qualitative dimension of regions' knowledge portfolios. Secondly, context appears to be very important. The relevance of each of these dimensions has been confirmed in previous research, which, in some instances, also implied the use of similar indicators. And yet, even when using very detailed and reliable data, we fail to identify clear associations of most of these dimensions with regional employment growth in Norway. Consequently, employment growth in Norway appears to be driven by other factors. Norway is unique in many ways, most notably in the dominance of the oil and gas as well as seafood industries. These industries experienced a substantial demand boom during the observational period (2009 and 2014), as oil prices rose by about 25%. This translated into a strong employment boom in oil and gas industries (Fitjar and Timmermans, 2019). The qualitative dimensions of regions' knowledge portfolio may be less relevant. However, this is speculative, and we must leave it to future research to identify the precise contextual dimension underlying these rather unexpected results. Alternatively, like any other empirical study, the present study is not free of shortcomings, which may potentially be the root of the surprising findings.

To overcome the context (or country) specificity of the empirical analysis discussed above, we encourage investigations utilizing information on multiple countries. This will facilitate controlling for such distorting country specificities and thereby help to identify the general relevance of relatedness, diversity, and complexity. In addition, unique developments can be used to study the importance of these factors in extraordinary situations. For instance, exploring how these dimensions contribute to the resilience of regions in times of general economic shocks is of crucial relevance for policy makers. Some early empirical results underline this argument (see, e.g., Otto et al., 2014; Balland, Rigby and Boschma, 2015; Diodato and Weterings, 2015; Eriksson, Henning and Otto, 2016; Holm, Østergaard and Olesen, 2017).

Furthermore, employment growth is just one dimension to consider. Although it has received the most attention in the literature, regional development growth does not show only in employment gains. For instance, regional economies also expand economically by expanding employees' wages. Employment growth and wage increases may go hand in hand, as growth in employment signals an increased demand for labor, which in competitive labor markets will lead to rising wages. Consequently, one can expect the three dimensions of diversity, relatedness, and complexity to have similar effects on wage as on employment growth. However, there are also good reasons for why they may differ. For instance, employment growth can be difficult to achieve because of a lack of qualified job candidates. In such cases, and with increasing demand for an industry's products or services, already employed workers will have greater leverage in salary negotiations. When new workers cannot be hired, the only visible indicator of an industry's growth will be increasing wages, while employment numbers will remain stable. Similarly, the contribution of an industry's employees to growth may materialize in employment growth outside the focal region or country. For instance, an innovation developed by an R&D team in Oslo, Norway, may result in a new production facility in Aalborg, Denmark. The facility will contribute to the revenues and profits of the corresponding company in Oslo, which in turn may increase the wages of the R&D team. In sum, looking at employment growth alone may give an incomplete picture of the true growth effects of diversity, relatedness, and complexity. Other dimensions, such as wage growth, need to be considered in future studies. Clearly, future work will have to overcome these limitations to reveal the true picture of the importance of diversity, relatedness, and complexity on regional growth.

In sum, in addition to providing a theoretical discussion on the importance of the qualitative dimension of regions' knowledge portfolios, this study highlights and illustrates some of the empirical approaches as well as their challenges that future research must address and overcome. Among others, it shows that having better (very detailed and fine-grained) data is not always sufficient. It also draws attention to the context specificity of these types of empirical exercises and raises the question of how future research can address this.

**REFERENCES**


Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth, 8*(2), 155-194.

Arrow, K.J. (1962). The economic implications of learning by doing. *Review of Economic Studies*, 29: 155-173.

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.

Balland, P.-A., Rigby, D. L. (2017). The geography of complex knowledge*. Economic Geography, 93*(1), 1–23.

Balland, P.-A., Rigby, D., Boschma, R. (2015). The technological resilience of US cities. *Cambridge Journal of Regions, Economy and Society*, 8(2), 167-184.

Balland, P.A., Jara-Figueroa, C., Petralia, S.G., Steijn, M.P.A., Rigby, D.L., Hidalgo, C.A., 2020. Complex economic activities concentrate in large cities. Nat. Hum. Behav. 4, 248–254.

Barbieri, N., Perruchas, F., Consoli, D. (2020). Specialization, Diversification, and Environmental Technology Life Cycle. *Economic Geography*, 96(2), 161-186.

Beaudry, C., Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, 38(2): 318-337.

Belussi, F., Caldari, K. (2009). At the origin of the industrial district: Alfred Marshall and the Cambridge school. *Cambridge Journal of Economics*, 33: 335-355.

Boschma, R.A., Frenken, K. (2006). Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. *Journal of Economic Geography*, 6(3): 273-302.

Boschma, R., Iammarino, S. (2009). Related variety, trade linkages, and regional growth in Italy. *Economic Geography*, 85(3): 289-311.

Boschma, R., 2017. Relatedness as driver of regional diversification: a research agenda. Reg. Stud. 51, 351–364.

Boschma, R.A., Iammarino, S., 2009. Related variety, trade links, and regional growth in Italy. Econ. Geogr. 85, 289–311.

Boschma, R., Frenken, K., 2011. Technological relatedness and regional branching, in: Kogler, D.F., Feldman, M.P., Bathelt, H. (Eds.), Beyond Territory. Dynamic Geographies of Knowledge Creation, Diffusion, and Innovation. Milton Park, New York, pp. 64–81.

Broekel, T., 2019. Using structural diversity to measure the complexity of technologies. PLoS One 14, 1–27. https://doi.org/10.1371/journal.pone.0216856

Caines, C., Hoffmann, F., Kambourov, G. (2017). Complex-task biased technological change and the labor market. *Review of Economic Dynamics, 25*, 298-319.

Capello, R., Lenzi, C. (2016). Innovation modes and entrepreneurial behavioral characteristics in regional growth. *Small Business Economics, 47*(4), 875-893.

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

Combes, P.-P., 2000. Economic Structure and Local Growth: France, 1984-1993. J. Urban Econ. 47, 329–355.

Davies, B., Maré, D. C. (2020). Relatedness, complexity and local growth. *Regional Studies*. DOI: 10.1080/00343404.2020.1802418

Deegan, J., Broekel, T., Fitjar, R.D. (2021). Searching through the haystack: The relatedness and complexity of priorities in smart specialization strategies. *Economic Geography*, forthcoming.

Desrochers, P., Leppälä, S. (2011). Opening up the 'Jacobs Spillovers' black box: local diversity, creativity and the processes underlying new combinations. *Journal of Economic Geography*, 11: 843-863.

Diodato, D., Weterings, A.B.R. (2015). The resilience of regional labour markets to economic shocks: Exploring the role of interactions among firms and workers. *Journal of Economic Geography*, 15(5), 723-742.

Dordmond, G., de Oliveira, H. C., Silva, I. R., & Swart, J. (2021). The complexity of green job creation: An analysis of green job development in Brazil. *Environment, Development and Sustainability*, *23*(1), 723–746.

Ederer, P., Nedelkoska, L., Patt, A., Castellazzi, S., 2015. What do employers pay for employees' complex problem-solving skills? Int. J. Lifelong Educ. 34, 430–447. https://doi.org/10.1080/02601370.2015.1060026

Ejdemo, T., Örtqvist, D. (2020). Related variety as a driver of regional innovation and entrepreneurship: A moderated and mediated model with non-linear effects. *Research Policy*, 49(7), 104073.

Ejermo, O. (2005). Technological diversity and Jacobs' externality hypothesis revisited. *Growth and Change*, 36(2): 167-195.

Eriksson, R. H., Hansen, H. K., Winther, L. (2017). Employment growth and regional development: industrial change and contextual differences between Denmark and Sweden. *European Planning Studies*, 25(10), 1756-1778.

Eriksson, R.H., Henning, M., Otto, A. (2016). Industrial and geographical mobility of workers during industry decline: The Swedish and German shipbuilding industries 1970-2000. *Geoforum*, 75, 87-98.

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.

Ferrarini, B., Scaramozzino, P., (2016). Production complexity, adaptability and economic growth. Struct. Change Econ. Dyn. 37, 52–61.

Fernández-Macías, E. (2012). Job polarization in Europe? Changes in the employment structure and job quality, 1995-2007. *Work and Occupations*, 39(2), 157-182.

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.

Fleming, L., Sorenson, O., (2001). Technology as a complex adaptive system: evidence from patent data. Res. Policy 30, 1019–1039.

Frenken, K., van Oort, F.G., Verburg, T., (2007). Related variety, unrelated variety and regional economic growth. Reg. Stud. 41, 685–697.

Glaeser, E.L., Kallal, H.D., Scheinkman, J.A., Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, 100(6): 1126-1152.

Grillitsch, M., Asheim, B., Trippl, M. (2018). Unrelated knowledge combinations: The unexplored potential for regional industrial path development. *Cambridge Journal of Regions, Economy and Society*, 11(2), 257-274.

Gundersen, F. Juvkam, D. (2013). Inndelinger i senterstruktur, sentralitet og BA regioner. NIBR report 2013-1.

Hane-Weijman, E., Eriksson, R.H., Rigby, D. (2020). How do occupational relatedness and complexity condition employment dynamics in periods of growth and recession? Papers in Evolutionary Economic Geography, 20.

Hidalgo, A., Hausmann, R., (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 10570–10575.

Hidalgo, C.A., Klinger, B., Barabasi, A.-L., Hausmann, R., (2007). The Product Space Conditions the Development of Nations. Science (80-. ). 317, 482–487.

Holm, J.R., Østergaard, C.R., Olesen, T.R. (2017). Destruction and reallocation of skills following large company closures. *Journal of Regional Science*, 57(2), 245-265.

Ivarsson, I., Alvstam, C.G., (2017). New technology development by Swedish MNEs in emerging markets: The role of co-location of R&D and production. Asian Bus. Manag. 16, 92–116.

Jacobs, J. (1969). *The Economy of Cities.* New York: Vintage Books.

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 of the United States of America*, 115(50), 12646-12653.

Juhász, S., Broekel, T., & Boschma, R. (2020). Explaining dynamics of relatedness: the role of co-location, complexity and collaboration. *Papers in Regional Science*, *September 2019*, 1–19. https://doi.org/10.1111/pirs.12567

Keeble, D., Owens, P.L., Thompson, C. (1983). The urban-rural manufacturing shift in the European Community. *Urban Studies*, 20(4), 405-418.

Killen, Andreas. *Berlin electropolis*. University of California Press, 2006.

Li, D., Heimeriks, G., Alkemade, F. (2020). The emergence of renewable energy technologies at country level: Relatedness, international knowledge spillovers and domestic energy markets. *Industry and Innovation*, 27(9), 991-1013.

Lo Turco, A., Maggioni, D. (2020). The knowledge and skill content of production complexity. *Research Policy*, 104059.

Marshall, A. (1890). *Principles of Economics.* London: Macmillan.

Martin, R., Sunley, P., (2015). On the notion of regional economic resilience: Conceptualization and explanation. J. Econ. Geogr. 15, 1–42.

Mewes, L., Broekel, T., (2020a). Technological complexity and economic growth of regions. Res. Policy. https://doi.org/https://doi.org/10.1016/j.respol.2020.104156

Mewes, L., Broekel, T., (2020b). Subsidized to change? The impact of R & D policy on regional technological diversification, The Annals of Regional Science. Springer Berlin Heidelberg.

Morkutė, G., Koster, S., Van Dijk, J., (2017). Employment growth and inter-industry job reallocation: spatial patterns and relatedness. Reg. Stud. 51, 958–971.

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

Neffke, F., Henning, M., 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.

Nooteboom, B. (2000). *Learning and innovation in organizations and economies*. Oxford: Oxford University Press.

Otto, A., Nedelkoska, L., Neffke, F., (2014). Skill-relatedness und Resilienz: Fallbeispiel Saarland. Raumforsch. Raumordn. 72, 133–151.

Quigley, J.M., (1998). Urban diversity and economic growth. J. Econ. Perspect. 12, 127–138.

Rigby, D.L., Roesler, C., Kogler, D., Boschma, R., Balland, P.A. 2019. Do EU regions benefit from smart specialization? Papers in Evolutionary Economic Geography, 19.31.

Romer, P. (1986). Increasing returns and long-run growth. *Journal of Political Economy* 98: 71-101.

Shearmur, R., Doloreux, D., (2008). Urban hierarchy or local buzz? high-order producer service and (or) knowledge-intensive business service location in Canada, 1991-2001. *Professional Geographer*, 60, 333–355.

Shutters, S. T., Lobo, J., Muneepeerakul, R., Strumsky, D., Mellander, C., Brachert, M., and Bettencourt, L. M. (2018). Urban occupational structures as information networks: The effect on network density of increasing number of occupations. *PLoS One*, 13(5), e0196915.

Stojkoski, V., Utkovski, Z., Kocarev, L., (2016). The impact of services on economic complexity: service sophistication as route for economic growth. PLoS ONE 11 (8), 1–29.

Thune, T., Engen, O.A., Wicken, O. (2019). Transformations in petroleum: Innovation, globalisation and diversification. Thune, T., Engen, O.A., Wicken, O. (eds). *Petroleum Industry Transformations: Lessons from Norway and Beyond.* London: Routledge, pp. 1-21.

Van den Berge, M., Weterings, A., Alkemade, F. (2020). Do existing regional specialisations stimulate or hinder diversification into cleantech? Environmental Innovation and Societal Transitions, 35, 185-201.

Van der Panne, G. (2004). Agglomeration externalities: Marshall versus Jacobs. Journal of Evolutionary Economics, 14(5), 593-604.

Whittle, A., Kogler, D.F. (2020). Related to what? Reviewing the literature on technological relatedness: Where are we now and where can we go? Papers in Regional Science, 99(1), 97-113.

Wixe, S., Andersson, A. (2017). Which types of relatedness matter in regional growth? Industry, occupation and education. Regional Studies, 51(4), 523-536.

Xiao, J., Boschma, R., Andersson, M., (2018). Industrial Diversification in Europe: The Differentiated Role of Relatedness. Econ. Geogr. 94, 514–549.