Related Variety and Regional Development

Harald Bathelt & Michael Storper

Papers in Evolutionary Economic Geography

22.14

Utrecht University
Human Geography and Planning

Draft: July 20, 2022
Word count: ca. 8,200 words (main text)

Related Variety and Regional Development

by

Harald Bathelt
University of Toronto
(harald.bathelt@utoronto.ca)

and

Michael Storper
LSE, UCLA
(m.storper@lse.ac.uk, storper@ucla.edu)

Acknowledgements

This paper was initially presented at the 2022 Global Conference on Economic Geography in Dublin. The authors would like to thank the audience, especially Tom Broekel, for supportive comments, as well as Adams Aghimien, Max Buchholz and Shawn Shuang for superb research assistance.

Related Variety and Regional Development

Abstract. Evolutionary approaches have contributed substantially to the growing knowledge body about regional development processes and their underlying mechanisms. They have advanced our understanding particularly by going beyond case study methods, using empirical, mostly regression-based statistical analyses. One key concept that underlies evolutionary economic geography (EEG) is that of "related variety". In EEG studies, regional industry structure is represented through its level of related variety, which in turn is found to be positively associated with favorable types of regional economic development. In this paper, we raise questions regarding the internal logic of the concept, its spatial expressions, measurement specifics, empirical regularities and biases, and the short- and long-term effects of related variety on regional development. Based on this examination, we make suggestions for future research.

Keywords. Economic geographies of places; evolutionary economic geography (EEG); regional development; regional specialization; related variety.

JEL Codes. L23; R11.

1. Introduction

There is a long tradition of thinking about how regional economies evolve gradually along a given pathway that is defined by their existing stock of economic activities, or, conversely, how they may break away from that pathway and reinvent themselves into something radically different (Chinitz 1961; Storper & Walker 1989; Boschma & Lamboy 1999; Bathelt & Glückler 2000). The former has been the central concern of a recent school of evolutionary economic geography (henceforth “EEG”) that can be traced back to the work of Boschma & Frenken (2006), and this school has become prominent in economic geography research (Martin & Sunley 2006; Essletzbichler & Rigby 2007; Boschma & Martin 2010a). As described by Boschma & Martin (2010b: 6f.), the goal of this approach is to investigate “the processes by which the economic landscape – the spatial organisation of economic production, circulation, exchange, distribution and consumption – is transformed from within over time”. A crucial feature of EEG has been a shift away from individual cases studies of regions, clusters and innovation systems toward comparative statistical and analytical work that aims to generalize about processes and mechanisms of regional economic development. This work has been very successful within the discipline and produced a large body of empirical studies that have shed new light on several important aspects of regional development processes. Our paper examines the development and deployment of one of its central concepts, that of “related variety” in technologies, industries and activities; it does not pretend to evaluate the entire corpus of EEG research.

A starting point for considering this literature is the pathbreaking study by Frenken et al. (2007). It has been the inspiration for a stream of publications that, broadly speaking, suggest that regions perform particularly well over time when their economic structure is characterized by industries that are technologically related but diverse within their related fields. A myriad of empirical studies conducted in the past 15 years find that related variety can lead to positive regional performance outcomes with respect to productivity, employment, innovation and patenting. Many such studies also imply that relatedness is a causal factor in tracing out the possible pathway of development of a region, its “roads taken and not taken”. Recent studies by Content & Frenken (2016) and Boschma (2017), among others, have summarized the main accomplishments of this body of research, identified missing aspects and suggested directions for future work.

EEG has also developed a distinctive and increasingly influential position on regional development policy, holding that it is best for a region to build on what is there and expand the portfolio of regional activities – their variety – within related areas. Thus, policy should promote related variety. In this view, expanding related variety is quite different from the older notion that a region should be widely diversified in what it does as a hedge against downturns. In a recent assessment, Martin & Sunley (2022, p. 70) note that related variety has become an obligatory concept in EEG that prioritizes this specific analytical approach to policy formation. They warn, however, that "[r]elatedness between a region's sectors of activity may facilitate recombination, branching, innovation and new path creation. But it may equally encourage structural lock-in."

Notwithstanding this active research output, relatively little research has thus far thoroughly assessed veracity of the core assumption, i.e. that increasing related variety is a key foundation of successful regional economic development (Henning 2019; Martin & Sunley 2022). In this paper, we carefully examine the fundamental assumptions behind and characteristics of this approach. Though EEG is currently developing into some new and different directions, we focus on the empirical research based on related variety. These empirical studies rest on the ways that the concept of related variety is operationalized in the form of specific technical-statistical measures of how activities are related to one another, or not so, and how such relationships shape development over time. The key to the usefulness of these measures is the extent to which they capture the relatedness that they claim to capture and correctly measure its effects on development. This requires that we examine the assumptions behind operationalization of technological relatedness in the economy, its relationship to geographical proximity and knowledge spillovers, and both aspects in relation to dependent variables of economic development and performance.

Our paper begins by considering the logic and assumptions behind this concept. We then examine the measurement and potential geographical expressions of related variety. This is followed by a presentation of our own results on certain empirical regularities in the United States that call into question some of the conclusions of EEG research. We conclude by emphasizing the need for a fuller conceptualization and discussion of the pathways of regional economic change.

2. Related Variety and Regional Development: A Conceptual and Empirical Survey

Boschma & Frenken (2011a) present numerous arguments that technologically related industries are better for regional economic development than a collection of non-related industries (Boschma & Frenken 2006; Frenken et al. 2007). With respect to the older debate over whether specialization or diversification is better for regional development, they try to square the circle of that debate by merging the two concepts into the notion of related variety, a specific form of internally-diversified specialization or related diversification.[1] They reason that this situation generates the benefit of having activities that are closely-related, enough that they will naturally create knowledge spillovers among them and generate technological dynamism through the combination and recombination of such knowledge. Extending that argument from technologies to geographies, they hold that certain forms of geographical proximity and technological proximity are often linked, because technologically-related activities are likely to co-locate as they benefit from knowledge spillovers and then co-evolve from the learning that goes along with such spillovers. Interactive learning is said to require that the cognitive, social and geographical "distance" (a difference) between economic actors is not too large. A first observation is that this kind of thinking is prima facie logical and is in good company with economic historians and other social scientists who have thought about regional economic specialization and its dynamic dimensions (Mokyr, 1991).

In terms of intellectual lineage, EEG often refers to the work of Marshall (1920), Jacobs (1969) and Audretsch & Feldman (1996) among others. It thus places itself in the context of researchers who have explored various dimensions of regional economic specialization, including the "learning" dimension of clusters, agglomeration and regional innovation systems (Cooke et al. 1997; Malmberg & Maskell 2002; Duranton & Puga 2004). What is novel and different in EEG is a specific statistical operationalization of the technological structures that are said to positively affect knowledge spillovers, learning and regional innovation. The basic thrust of the work is that – at the level of 4- or 5-digit NAICS codes or equivalent granularity – the sub-sectors that are nested in a sector are strongly technologically related, in the sense referred to above. They are assumed to have a significantly higher potential to spill over knowledge to and to learn from one another so as to generate innovations and/or growth. These consequences of technological proximity are assumed to be enhanced when these related sub-sectors are geographically co-located. Whether or not the presumed knowledge spillovers among related activities are strengthened by co-location is said to depend on a number of factors. These can include whether the sectors are new or mature, routinized or non-routinized, what their underlying assumed innovation potential is, and how much inherent potential there is for the use of complementarities.

Table 1 gives an overview of the 30 most-cited publications on related variety according to SCOPUS citations (29 November 2021). The table characterizes each paper in terms of its geographical basis, number of observations, dependent variables, related and unrelated variety

[1] While EEG work often emphasizes the diversification aspect over specialization; by the same token this can be viewed as clustering.

indicators, some key findings, specific mechanisms that are explained, and the analysis of geographical outcomes and peculiarities in the results. While this table only presents a snapshot of the existing literature on related variety, it provides a good starting point to analyze the most influential work in this field of research. Most studies broadly follow the methodology introduced by Frenken et al. (2007), deploying positive regional development outcomes as dependent variables in relation to a set of independent variables, the core of which are measures of technological relatedness (i.e. related and unrelated variety). These studies often, albeit not always, find a positive relationship to such regional performance variables and related variety, itself assumed to be driven by knowledge spillovers and economic collaboration. But it stands out that these studies do not actually directly investigate these mechanisms nor provide an in-depth explanation for them.[2] None of the 30 studies listed in Table 1 presents empirical information about the specific mechanisms that drive successful or less successful regions. There is also some confusion about what is being captured by the dependent variables of regional performance: in some studies, technological relatedness stimulates regional growth, while in others it is productivity, employment growth or innovation. It is not entirely clear what the channels of causality are in these different cases in relation to the underlying theoretical setup around related variety. Complicating matters, related variety analyses typically do not consider alternative explanations in their models; we did not find systematic exploration of control variables or robustness tests.

[Table 1 about here]

The EEG literature is also largely silent about whether relatedness leads to direct relationships within the regional economy, indirect ones, or something else. For example, firms that operate in technologically related sectors or are vertically connected within a value chain may indeed have potential to cooperate and generate synergies but this requires that these firms actually do collaborate. There is a broad literature that shows that intense collaborations in industrial agglomerations or clusters are rare and cannot be assumed (Bathelt et al. 2004). The mere fact that firms are co-located in the same region and are in technologically related fields (as defined and measured in the EEG literature) does not mean that they operate in the same fields of application, and hence that they benefit

[2] Content et al. (2019) are aware that the mechanisms as to how regional knowledge spillovers are created are not fully clear and that evidence needs to be provided for these effects (Content & Frenken 2016). An exception is the study of Miguelez & Moreno (2018), which tackles underlying mechanisms rather than assuming them.

from close cooperation or are even interested in it. A number of firms within the same region may use opportunities to collaborate and benefit from this (e.g. Breul et al. 2015), or they may distrust each other or see one another as close competitors. Regional firms need to have a reason for collaboration and do not automatically mobilize themselves to engage in joint action.

Alternatively, one could instead argue – as in the cluster debate (Bathelt et al. 2004) – that any positive impact of related variety is likely to spread through indirect knowledge and labor market effects, and that actual collaborations between firms may be secondary. This is consistent with the standard micro-foundations of agglomeration (sharing, matching, learning model) where learning may be an emergent property of sharing and matching (Duranton & Puga 2004). Boschma & Frenken (2011a), in this vein, point to a number of potential mechanisms that can be associated with or lead to technological relatedness. They stress four such mechanisms: producer-user relations, interdependencies in the production system, technological complementarities between industries, and interdependencies in technology development. The point is that the assumed logic of spillovers due to statistically-identified regional related variety may take many and varied forms, and does not flow from the measure of related variety itself. Such different possible spillovers would need to be validated empirically and in real geographical contexts.

We noted that the related variety literature seems to aim to resolve an opposition between the benefits of specialization and those of diversification in regional development, by combining the two under a single umbrella. The long tradition of studying regional economies via the optic of their specialization, which is sometimes further specified as specialization through clustering of some set of densely interrelated activities, seems to use the opposite theoretical labeling from the variety part of the EEG framework. The EEG literature does not give us clear guideposts about where it sits in relation to the body of research on specialization and diversification (Kemeny and Storper, 2015). Is this merely a semantic difference, or is it a difference that matters in real terms? The EEG literature often invokes Jacobs (1969) as an inspiration for the importance of diversification (Frenken et al. 2007). Jacobs, however, never advanced a clear formal model of diversity and regional growth, so this does not resolve the question of whether it is more sensible to label a regional economy diversified, specialized, or relatedly diversified. This ambiguity also shows up when comparing the approach with Porter's (1990) diamond model, where Porter's emphasis is on specialization in the form of related and supporting industries. One may ask: what is different between the two (Ketels 2016)? We will show in the following sections that this issue of whether an economy is relatedly diverse or specialized is not just a question of semantics, by discussing how well relatedness and variety are actually defined and measured in EEG (section 3) and then demonstrate that there is no clear idea in the EEG literature of what a relatedly diverse industrial structure might look like (section 4).

A final point in this initial discussion of assumptions and results is that, even though the related variety concept does not exclude the role of external processes and linkages per se[3], it focuses on intra-regional processes and pays little attention to the extra-regional geography of spillovers and linkages (Content & Frenken 2016). In reality, many firms are linked to other regions and countries worldwide through subsidiaries or partnerships that have developed over time (e.g. Cantwell 1989; Crescenzi & Iammarino 2017; Li & Bathelt 2018; 2021; Bathelt & Buchholz 2019; Yeung 2021). Through these linkages, they are able to share knowledge over distance (Bathelt & Henn 2014) and may be less dependent on and less engaged in home-region knowledge spillovers. It is unclear whether regional related variety would be a strong influence on these firms' (and their respective regions') performances. They are able to participate in localized knowledge spillovers simultaneously in different places (Malecki 2010). Is this because technological closeness is more important than geographical proximity or that they are sometimes substitutes and at other times complements? Certainly, a deep understanding of these dimensions would be essential to establish when related variety is a positive property of a regional economy as opposed to a spatially-extensive value chain or field of endeavor. As it is, this key question in economic geography is not taken on by EEG in a sustained way. In sum, this discussion indicates that it cannot be assumed, a priori, that related variety (even assuming a method for successfully capturing it empirically) would create significantly higher levels of interaction within a region than some other compositions of activities. Since such interaction is one of the channels to superior regional performance posited by EEG researchers, it also cannot be assumed up front that related variety always positively stimulates regional performance or long-term development.[4] Below, we scrutinize some of these links in more detail.

---

[3] See, for instance, the studies by Saviotti & Frenken (2008), Boschma & Iammarino (2009) and Miguelez & Moreno (2018).

[4] An investigation by Spencer et al. (2010) asks similar questions with respect to the impact of industrial clusters on regional performance. While not focused on related variety, they investigate in a study of Canadian city-regions in the early-2000s whether regional performance depends on the degree of clustering. Their study presents descriptive evidence that clustered industries are generally located in better-performing regions with higher incomes and lower unemployment than non-clustered industries. Typically, city-regions with a higher employment share in clusters have a higher average income, employment growth and patenting intensity, although there is also large variation and this relationship cannot be confirmed for all cases. As with the related variety concept, the direction of causality remains opaque. Do clusters cause this development or are they attracted to high-income regions that may have higher skill levels?

3. Measuring Related Variety

A core concept and technique in EEG is to operationally define and then measure technological relatedness as suggested by Frenken et al. (2007). They quantify related and unrelated variety by using entropy measures (Theil 1972; Reardon & Firebaugh 2002) based on the regional industry structure. According to this approach, technological relatedness is measured by simultaneously analyzing industrial classification systems (such as SIC or NAICS) at different aggregation levels. If a region's employment is broadly distributed across many different sectors at a high level of aggregation (e.g. 2- or 3-digit) this is viewed as an indication of unrelated variety as these sectors are assumed not to be technologically related – a reasonable assumption in many cases. But when looking at more finely-grained aggregation levels of industries (e.g. 4-, 5- or 6-digit), a broad representation of sub-sectors within a major sector is viewed as an indication of technological relatedness within a field. Regional related variety as defined by Frenken et al. (2007) adds all main sector scores, weighted by sector size. Notice that the same index that is interpreted as indicating unrelated variety at a high aggregation level is then viewed as an expression of related variety when applied to fine-grained industry levels. The validity of this "within versus between" distinction, at some level of granularity in industry codes, is the core of the matter in the sense that the entire edifice of related variety studies depends on it. Indeed, the EEG project around related variety is based on a statistical artifact: "proximity" within classes of industries. And this artifact is in turn extrapolated from another artifact: the industrial classification system as a whole. In this, it differs from directly observable categories of economic reality, such as employment, wages, incomes, sales and so on.

While methodologically elegant, the use of a double artifact should be subject to the greatest care and prudence. To start with, industrial classification systems have never been designed with sub-sectoral interrelations, technological spillovers or common developmental dynamics in mind. They are based on the classification of outputs following a logic of cognate end-uses (e.g. cars, clothes, leisure). These classifications are mobilized by EEG to extrapolate relatedness, which is then in turn assumed to shape change, evolution and dynamism in regional economies due to spillovers and other features of being closely related. In an attempt to defend the "principle of relatedness", Hidalgo et al. (2018) refer to many different ways how activities could be related through a "shared knowledge base". They assert that it is a general and multi-scalar principle of modern economic development through the way that related industries constitute a "product space" but they then, admittedly, only infer the existence of such a space by the composition of a region's exports, where relatedness itself is defined as an extrapolation from something else. Meanwhile, attempts to directly scrutinize the veracity of relatedness measures, such as that by Delgado et al (2016, p. 1), find that co-location patterns, input-output links and similarity in occupations "[outperform] other methods in capturing a wide range of inter-industry linkages, including the grouping of industries within the same three-digit NAICS".

Another potentially important limitation is that industrial classification systems look backward and react slowly to changes in industry structure or the emergence of new industries, and yet the latter is a key dynamic process that EEG asks us to consider. Think about the case of photo-optics, which was once an industry involving films, cameras and lenses, while photography is now basically just a function of the digital production and storage industries. In the United States, the three major firms (Xerox, Eastman Kodak, Bausch and Lomb) generated great regional prosperity in their hometown of Rochester, New York. The problem for these firms and Rochester itself is that "photography" and "imaging" are broad output or product areas whose technology of production has migrated from optical to digital. Where should they be placed in the NAICS system? And is this change a related evolution or a rupture based on the application of new general-purpose technologies across many different output fields? The fact that related variety is essentially a static concept has also generated criticism from within EEG (Juhász et al. 2021; Kuusk and Martynovich 2021).

The logic of technological relatedness does make a certain sense if indeed all sub-sectors within a main industrial sector are part of the same vertical value chain, i.e. making inputs that go into one another and using real interrelations to do so. To be operationalized, however, it also would seem to require that all main sectors (e.g. at the 2-digit NAICS level) be considered "unrelated" to each other. But this is not self-evident from a theoretical or on-the-ground viewpoint. Manufacturing sectors are for instance closely linked to producer services or machinery industries and some industrial sectors are linked to each other such as computer hardware and software, automobiles and metal fabricating, chemicals and pharmaceuticals and biotechnology – or many other examples. Conversely, if we use 2-digit NAICS codes as a starting point of our analysis, some have more related sub-sectors than others – and in some main sectors sub-groups are hardly linked to each other. The latter can be illustrated in the case of manufacturing: NAICS 31 includes sub-sectors such as food, beverages, textile, apparel, leather and footwear and NAICS 32 sub-sectors such as paper, printing, petroleum/coal, chemicals, plastics, pesticides, pharmaceuticals and cosmetics. NAICS 33 seems a bit more internally coherent around metal-working industries but it is also rather diverse with sub-sectors such as steel, metal fabricating, machinery, engines/turbines, computers and semiconductors. This becomes even more diverse when looking at services sectors that often appear as rest categories in industrial classification systems. Examples are NAICS 56, which consists of a range of administrative, support, waste management and remediation services, or NAICS 55, which includes sub-sectors such as legal services, accounting, computer, scientific research and advertising services.

One could argue that these concerns are based on using a very high level of aggregation, and that they disappear as we get more granular. But there is no clear definition what the right level of aggregation is. Indeed, while industries at the 3-digit level in the classification system are more homogenous than at the 2-digit level, there is still substantial diversity within 3-digit industry groups and related outside categories still exist, especially in services sectors. It seems unlikely that there is an ideal level that eliminates unrelated industries and draws the line around related ones because as we point out above, the classifications are not designed to draw this distinction in any rigorous way. That does not automatically mean that they are not useful, but – to restate our point – the classifications that define related industries and separate them from unrelated ones do not flow from the aggregation level itself. This makes it essential to do some combination of statistical testing of relationships and ground-truthing them prior to constructing the statistical artifact of relatedness. We suspect that the index as deployed in much EEG research is capturing a wide array of different forms of relatedness. A particularly potent example of this comes from the findings in Storper et al. (2015) in analyzing high-tech industries in California. Their study revealed substantial differences in wages (up to 50%) within 6-digit NAICS occupations between two California metropolitan regions. This implies that, at least for some industries, even 6-digit industry codes may not be very homogeneous in what they are actually capturing "on the ground", across places; at worst, they may be little more than chaotic descriptive aggregations.

It should be noted that recent studies on industrial linkages and agglomerations argue that the nature of linkages that constitute relatedness is shifting from a sectoral logic that dominated the manufacturing era to one where related occupations, functions, activities or downstream linkages particularly benefit from co-agglomeration (Duranton & Puga 2005; Mudambi 2008; Timmer et al. 2019; Delgado & Mills 2020). While the majority of related variety studies focus on industrial relatedness, there is now an increasing number of investigations that use different indicators. These include "skill relatedness" (Boschma et al. 2011; Boschma et al. 2014), "knowledge variety" (Tavassoli & Carbonara 2014) or "occupational and educational related variety" (Wixe & Andersson 2017). Most promising are studies that measure relatedness through observable interactions and relations between firms, such as labor mobility or co-patenting (Neffke et al. 2011; Boschma et al. 2014; Fiorgo & Mayerhofer 2018). Still these studies do not investigate how such supposedly revealed relatedness[5] would have a stronger positive impact on learning, innovation and growth capabilities, compared to a less related industry structure. There is little conceptual discussion of which related variety concept is most fitting and should be used. The studies listed in Table 1 are focused on the empirical application of the related variety concept, not on questioning or adding to its theoretical premises.

Considering our arguments up to this point, it can be asked what the concept of relatedness as specifically operationalized in EEG exactly means, both from a static (input-output interdependence) and dynamic (developmental-innovative-co-growth) perspective. What we have not been able to find are systematic discussions or investigations of this in the EEG literature, whether in the form of a statistical or ground truthing exercise.

4. What Does a Regional Economy with High Related Variety Look Like?

Despite some similarities, the EEG notion of related variety goes beyond Porter's conception, in the sense that it attempts to capture the entire regional industry structure in some way. Note, however, that none of the studies reviewed in Table 1 provides a closer analysis of regions with high or low related variety (or unrelated variety) or a precise characterization of their overall industry structure. In fact, it seems quite unclear what specific form of regional industry structure would create a high degree of related variety, and what form would not. In order to explore this point, Figure 1 attempts to represent logically the industry structure of a regional economy with different hypothetical degrees of related variety. It builds on the observation that related variety is associated with regional specialization and clustering and presents the related and unrelated variety scores for 10 different scenarios of regional clustering. For reasons of simplicity, it is assumed that the economy portrayed in the figure includes 10 sectors which each consists of 10 sub-sectors. It is also assumed that the 10 sectors are completely technologically unrelated, which is unlikely in reality but confirms with the assumptions of the related variety concept. It is further assumed that these sectors form a perfect related cluster if each of the corresponding sub-sectors is equally well-developed (for instance in terms of employment). Assuming a region with an overall employment of 100,000 people, we constructed different scenarios of clustering: Scenario 1 characterizes a region with 1 very large cluster, in which all employees are distributed across the sub-sectors of a single cluster sector, and scenario 10 refers to

[5] Although measures of revealed relatedness are increasingly being applied they cannot yet be considered to be a new standard. While they identify related economic activities based on actual linkages, the use of such indicators makes it challenging to define unrelated variety.

a situation with 10 small clusters of equal size, with employment in each cluster again being equally distributed across all relevant sub-sectors. The difference between the scenarios in Figure 1 is the number of clusters and their size.[6] The plot reveals that related variety is highest when then regional economy is made up of 4 (fully developed) clusters. The related variety score decreases when the number of clusters is reduced (1 cluster corresponding with maximum specialization) or when it is increased (10 clusters corresponding with maximum diversity). Unrelated variety, in contrast, increases monotonically with the number of clusters. These trends are not surprising for those who work with the concept. But it is not clear conceptually why precisely 4 clusters should be better than 2 or 3 clusters or why 6 clusters would basically be as good as 2 clusters.

[Figure 1 about here]

From a policy perspective, this appears quite problematic, but we are unaware of any careful discussions of these issues in the EEG literature itself. The effects of technological relatedness would also undoubtedly depend on and vary with the specific industries involved, and not just the number related complexes and how internally or externally related they are as an artifact of the use of the classification system. Any useful policy implications therefore would need to consider the specific economic context of a region and the reality of industries in the contemporary economy. While some papers pull together and identify empirical regularities in studies on related variety (Boschma & Frenken 2011b; Content & Frenken 2016), it seems unclear what these can tell us if we lack a clear understanding how related variety translates into real geographies.

5. Related Variety and Regional Economic Structure and Performance

To get an initial sense of how related variety varies across a country's urban system in relation to structural features of regional economies, we treated related variety as a dependent variable and verified through scatterplots and simple regressions how this indicator is associated with variables such unrelated variety, ln(population), median income and education levels, as well as with employment shares in 2-digit NAICS industries. The results using data from the Quarterly Census of Employment

---

[6] The same fundamental results in terms of related and unrelated variety scores can be found irrespective of overall employment of the regional economy since employment shares rather than absolute numbers are decisive for the computation. However, as revealed further down, the industry structure in large and small cities is in reality rather different producing some bias of high related variety values toward large cities.

and Wages (U.S. Bureau of Labor Statistics 2020) for metropolitan statistical areas (MSAs) in the United States for the year 2017 are shown in Figures 2 and 3. The figures reveal some surprising regularities and potential biases in the concept itself.

[Figure 2 about here]

The first strong positive relationship that emerges in Figure 2 is that between related and unrelated variety.[7] From the definition of the two indicators, this close relationship is not obvious or may even be counter-intuitive (see also Figure 1) but an explanation can be found considering scale. Both, related and unrelated variety have a strong positive association with city size, as measured by ln(population). More specifically, we find that with increasing city size, related and unrelated variety (adopting the EEG's operationalization of these notions, for the purpose of this test) are systematically increasing among U.S. MSAs. The concept thus unintentionally associates large cities with high related variety and essentially favors larger cities over smaller cities in terms of prospects for regional development. A plausible explanation can be found when reconsidering the definition of related and unrelated variety and how industries develop differently in cities of different size. On the one hand, large cities often have large and diversified economies that therefore host a wide range of major sectors. At a highly aggregated level, it can thus be expected that unrelated variety increases with city size. But with increasing city size, we are also more likely to find many well-developed sub-sectors within each main sector. According to the entropy measure by Frenken et al. (2007), this will result in a high degree of technological relatedness in a large city. On the other hand, small cities are less likely to have such an industry structure. Fully-fledged sub-sectors across multiple main sectors are quite uncommon in small cities. In contrast, we are more likely to find a concentration of employment in selected sectors and can expect the economic structure to be characterized by many gaps. While such gaps or low representations at the main sector level lead to lower unrelated variety scores, the existence of gaps in the sub-sectoral structure in small cities conversely also results in lower related variety scores. Overall, this generates a situation in which both related and unrelated variety systematically increase with city size. As a result, this creates bias toward larger cities and makes it extremely difficult to interpret findings in terms of regional development dynamics, other than saying large cities have stronger

[7] Castaldi et al. (2015) also found a positive correlation between unrelated and related variety in their study at the U.S. state level but are neither concerned about this relationship nor investigate it further.

development potential than smaller cities. Essentially, in a complex urban system such as that of the United States with many cities of different sizes, related variety operates as a proxy of city size.[8] While this is potentially worrisome, it is interesting to note that most studies in Table 1, except for 4[9], do not even integrate an indicator of regional scale, such as population size, in their models. Instead, they use population density which is meant to consider urbanization economies in the model formulations. No study in the table questions the use of this variable, even though it is far from clear that such density measures are an adequate representation of urbanization economies. It should also be noted that this discussion echoes debates in the economics of agglomeration literature as to whether localization economies are captured principally by some relative measure of concentration (high share) or whether a main source is the scale of the cluster itself, and how this would affect the ability of each related activity within the cluster to exploit scale efficiencies (Kemeny and Storper 2015).

But it is notoriously hard to find solid generalizations about regional economic structure and its relationship to performance. Figure 3 implies that some aggregate sectors have a negative association with related variety when they dominate a large part of overall employment, and this is more likely to be the case in smaller regions (see also Figure 1). When a main sector has a very high employment share beyond 20 or 30% the related variety score decreases – the implication being that the overall impact on regional development is negative. Indeed, for agriculture (NAICS 11), mining (NAICS 21), traditional manufacturing (NAICS 31) and retail (NAICS 44), we find that very high regional employment shares are associated with decreasing related variety. The related variety approach does not see positive development outcomes in regions with a single dominant industry cluster but rather emphasizes the danger of potential lock-in and a lack of alternatives in such a situation. However, this is not always the case. Storper et al. (2015) note that the San Francisco Bay Area showed a 4-fold increase in the direct share of IT-based employment from 1970 to 2010, which does not seem to be associated with lock-in, but with extraordinary technological learning and dynamism. The related variety literature would undoubtedly conclude that that is because of the concentration of a wide variety of related sub-sectors; but then it is saying little about the overall regional industrial structure

[8] This may also explain why recent studies use indicators, such as related variety density (consisting of a quotient of different related variety indicators – e.g. Balland et al. 2019), to reduce the impact of scale on the overall related variety measure computed. While this is certainly a step forward, much previous research has used the approach by Frenken et al. (2007). There are also other concerns since related variety density or a combination of relatedness and complexity variables (Deegan et al. 2021) are even more difficult to make sense of in policy terms than conventional related variety.

[9] See Boschma et al. (2009), Cainelli & Iacobucci (2012), Ebersberger et al. (2014) and Tavassoli & Jienwatcharamongkhol (2016).

and a lot about the dynamism of a specialized regional cluster. All of this echoes the point made in a classical paper by Chinitz (1961) that what counts is not just whether a region is specialized or clustered, but the nature of the specialization itself. Using the EEG semantic, we could say that it is not just whether a region has a lot of related variety, but in what activities its related variety is expressed and whether it is a big region with other unrelated or related activities or not.

Because of different regional contexts and industry configurations, it would be important to study regions with low related variety associated with a dominant single cluster yet positive development outcomes. And vice versa, it would be interesting to investigate the reverse situation. Only through such comparative analysis is it possible to identify different mechanisms at work that produce varying regional outcomes. None of the studies reviewed in Table 1 asks such questions systematically.

[Figure 3 about here]

Figure 3 also shows that those sectors that do not become as dominant regionally as the above-mentioned sectors have a positive association with related variety. Among those are information/media industries (NAICS 51), real estate (NAICS 53), professional, scientific and technical services (NAICS 54), administrative, support, waste management and remediation services (NAICS 56) and arts, entertainment and recreation services (NAICS 71). While some of these industries have been emphasized in the creative class literature (Florida 2002; 2017), it is hard to construct a direct relationship between these industry groups and a conception of technological relatedness or to explain superior regional outcomes through associated knowledge spillovers and development triggers. It should be noted, however, that these sectors are typically more developed in larger as opposed smaller cities, which is consistent with our prior observation regarding scale.

6. Short- and Long-Run Effects of Related Variety

In EEG, related and unrelated variety are often used as independent variables to explain regional development outcomes, such as economic growth, changes in employment or innovation activities. This is typically done in large-N regression models with multiple independent variables. As in some regional economics work (e.g. Glaeser et al. 1992), the strategy of such modeling approaches is typically to explain regional development over a time interval of a few years, using related variety and other variables as independents at the beginning of this time period. This is also the approach of

Frenken et al. (2007) in their study of regional income and employment growth in Dutch regions. While their analysis only draws from a small number of regions overall, many other EEG studies investigate regional development using a similar approach based on much larger sample sizes, which is a step forward. When looking more closely, however, it appears that many studies have a are cross-sectional design. Among the most-cited related variety papers, about half use a conventional cross-sectional approach to predict regional development over time based on related variety at the beginning of that time period. The other half of studies in Table 1 neither conduct a true panel analysis nor incorporate some dynamic approach.

Following the methodology by Frenken et al. (2007), Buchholz & Bathelt (2021) use related and unrelated variety measures in 2010 to explain changes in income and employment levels in U.S. MSAs between 2010 and 2017. Interestingly, this study is able to largely reproduce the findings by Frenken et al. (2007) for the U.S. urban system. The study finds that unrelated variety is positively and significantly associated with income changes, while related variety is positively and significantly linked to employment increases (Table 2). This, in and of itself, suggests that there are different channels of causality at work, as the U.S. urban system is sharply divided between urban regions with high income growth versus those with high population growth.

In addition, the question arises whether it is really the best approach to explain change in the dependent variable by using some state of the independent variables from the past. If we are interested in causal explanations, it would be more intuitive to conduct a panel analysis to investigate what kind of *changes* in related variety go along with *changes* in the dependent variable. Such an analysis is much more plausible when it is the goal to explain whether changes in related variety indeed explain regional development and to derive policy suggestions from this. Making such a shift and conducting a panel analysis for U.S. MSAs with the same data (Buchholz & Bathelt 2021), the outcome is different from before and differently from what would have been expected (Table 2). The corresponding results of the panel analysis show that related variety is no longer significant and unrelated variety seems to be negatively associated with income changes albeit at a low significance level.

[Table 2 about here]

While other studies have come to different conclusions using panel analyses, we should not be overly surprised that the results of cross-sectional studies cannot automatically be transferred to a panel. In fact, in a short time frame of few years, changes in the related and unrelated variety structure of U.S. regions are small and scores only change incrementally (Buchholz & Bathelt 2021). In many cases, shifts in the localized industry structure are not large enough to explain changes in regional development. To be fair, findings of related variety studies are not consistent in this respect and some report different results. For instance, Castaldi et al. (2015) find in a study of U.S. states over a longer time span from 1977 to 1999 that related variety has a positive and significant impact on patenting activity. Overall, however, about as many studies in Table 1 suggest a significantly positive impact of related variety on regional growth as studies that find insignificant or negative impacts.

The above arguments suggest that we need to ask serious questions regarding causality. Is it acceptable to view regional development as the consequence of a favorable industry structure if this industry structure does not change or stays stable in the short time observed? More generally, can related variety be considered the cause of positive regional performance or is it rather a reflection of it (Martin & Sunley 2022)? Related variety in Figure 2 is positively correlated with the share of workers with a college degree and with median income and negatively correlated with the share of workers without high school degree. It would be difficult to identify a mechanism through which related variety directly causes better skill levels or higher incomes unless it indicates positive specialization in innovation activities in a technological field; even so, it would certainly be plausible to speculate about reverse causality.

The EEG literature also has no real engagement with the question of whether disruptive technological change can undermine the advantages of relatedness or even generate positive advantages to previously unrelated activities. In the economics of technology literature, there is considerable effort to distinguish minor, within-paradigm technological changes from major, disruptive change (Perez 2010; Petralia 2020). The terminology is wide and varied: disruptive, general-purpose, radical, paradigmatic shifts, and so on; but reasons explored for why some changes are more important than others include: (a) some radical technologies replace previous user technologies (e.g. digital photography replaced photo optics and film and mechanical cameras); (b) some radical technologies have new types of user complementarities (fields of users that were not previously linked, such as when machines become digitally guided as with the emerging self-driving cars); and (c) some radical technologies make possible completely new types of activity (the telephone made possible the remote hearing of voices) (Petralia, 2020). Related variety as used in the EEG literature is not likely to explain these cases and – at least in some of them – crucially positive effects on regions hosting the breakthrough process would not be detected by the EEG set up.

To take the most glaring example of this, consider that the silicon-based semiconductor was invented in New Jersey in 1954, and that the major centers of semiconductor production from the 1950s through the early 1970s included New Jersey, Los Angeles County, Arizona and New York, but that by the mid-1970s, the San Francisco Bay Area (also known as Silicon Valley) had taken the lead (Scott and Storper, 1987). Though there were communications industries in the Bay Area, the evolution into semiconductor dominance was initially far from evident from the Bay Area's overall industrial structure (Saxenian 1985; 1994), which was dominated by natural resource processing industries, branch plant manufacturing, refining and an important port complex. And reverse processes were also not predictable, i.e. the places with a heavy concentration in chip design and manufacture in the 1950s and 1960s that lost out subsequently (Storper et al. 2015).

One possible reason for the relative lack of attention to technological and geographical disruption may be the European origins of EEG work. Europe lost out on radical technological change leadership in the 3rd industrial revolution as there are virtually no household names of technology leaders, which are overwhelmingly American and located in the San Francisco Bay Area and Seattle (Soskice 2021). Thus, an implicit bias may be associated with the economic reality in much of Europe, where more incremental technological change dominates, whereas radical innovation and discontinuous geographies may be more dominant in the United States. Thus, pathways of development, and within them the role of relatedness, may be shaped not only by technology, but also by institutional varieties of capitalism (Hall & Soskice 2001). The problem is not, then, necessarily the empirical conclusions from many European EEG studies, but the elevation of those findings to a general theory about technology, geography, and regional development.

7. Conclusion: Related Variety, Place and Geography

As we have noted throughout this paper, most of the literature on related variety studies as it currently stands is not about place and explaining geographic variation in development. Its thrust is to identify macro regularities or mechanisms in a larger population of cities or regions. This is certainly an important step forward, but EEG should not stop there. As our review of the core of related variety studies in Table 1 shows, publications typically end by presenting and interpreting the significance levels and direction of relationships in multiple regressions but do not investigate regional variations and deviations from the identified relationships. The studies in Table 1 either do not discuss geographical variation at all, briefly present variable distributions over space or use regional dummy variables. Only 1 of the 30 studies links to a specific case study. Neffke et al. (2011) mention the case of Linköping in their analysis of the impact of reveled relatedness on industrial transformation in the region. However, this case is declared as arbitrary and remains descriptive. None of the papers in Table

1 identifies regions where the empirical models fit well to explain regional development or discusses other regions where the models do not provide an adequate reasoning.

More generally, this observation brings us to a broader theory debate. Development is one of the noisiest problems to solve in social science, and there is no consensus about the sources of economic dynamism, development and decline and how they interact. At a minimum, regional development involves institutional influences, cultural factors, connectivity and geography, migration and workforce change, regional land use and housing, education, racism and segregation and class relations, as well as many other dimensions. It is thus awkward that many EEG studies concentrate their attention on a theoretical framework that can be characterized as “technology drives technology which drives development”. This has a trace of technological determinism, in our view, and could benefit by drawing on the wider social science of economic development in which technological change is seen from a multiplicity of angles (e.g. Mokyr 1990).

To be clear, the purpose of this paper is neither to give a complete overview of the entire related variety debate, nor to question the usefulness of the core concept within EEG, but rather to identify ways in which it can reach its promise. In order to do so, we believe that it should build out from where it has begun, investigating such issues as: (i) the geographical expressions of this approach and to better understand its regularities, (ii) to grasp how related variety indicators vary in real economies and what mechanisms are at play, (iii) to identify limitations and biases of the relatedness measure, and (iv) to investigate the outcomes of statistical analyses systematically in relation to concrete regional development contexts. The latter should include substantive engagement with wider explanations for regional performance, alternative hypotheses, robustness checks and many other issues that are discussed in regional economics, international business, development theory and economic geography as a whole. A starting point could involve better description and interpretation of data and results for real-world cases. While we focus in our analysis on the most-cited related variety studies, many of our remarks are equally relevant for recent work in this field.

Most importantly, EEG research should investigate in which regions its regularities do or do not apply and which mechanisms can explain this. Residuals and distributions are helpful to the improvement of concepts. There are always regional cases where identified regularities work and other regions where this is not the case and where large positive or negative residuals can be identified (Buchholz & Bathelt 2021). Regularities may help understand important average trends but they do not do justice to diversity of place. Different contexts and institutional set-ups can impact regional development in specific ways as they can generate different causal relationships between variables that trigger different mechanisms of growth, decline or stagnation (Storper 1997; 2009; Storper et al. 2015;

Glückler & Bathelt 2017; Buchholz 2019; Gong & Hassink 2020). These investigations would also strengthen the basis for relatedness policies in relation to other forces (Iammarino et al. 2019). We hope that this sympathetic critique contributes to a discussion of how we can collectively improve our understanding of regional development and the possible role that relatedness plays in it.

References (Top-30 Most-Cited Related Variety Articles According to SCOPUS Citations, 29 November 2021)

Aarstad, J., Kvitastein, O.A. & Jakobsen, S.-E. (2016) Related and unrelated variety as regional drivers of enterprise productivity and innovation: A multilevel study. *Research Policy*, 45(4): 844-856.

Antonietti, R. & Cainelli, G. (2009) The role of spatial agglomeration in a structural model of innovation, productivity and export: A firm-level analysis. *Annals of Regional Science*, 46(3): 577-600.

Basile, R., Pittiglio, R. & Reganati, F. (2017) Do agglomeration externalities affect firm survival? *Regional Studies*, 51(4): 548-562.

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

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

Boschma, R., Minondo, A. & Navarro, M. (2011) Related variety and regional growth in Spain. *Papers in Regional Science*, 91(2): 241-256.

Cainelli, G. & Iacobucci, D. (2012) Agglomeration, related variety, and vertical integration. *Economic Geography*, 88(3): 255-277.

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

Castaldi, C., Frenken, K., & Los, B. (2015) Related variety, unrelated variety and technological breakthroughs: An analysis of US state-level patenting. *Regional Studies*, 49(5): 767-781.

Content, J., Frenken, K. & Jordaan, J.A. (2019) Does related variety foster regional entrepreneurship? Evidence from European regions. *Regional Studies*, 53(11): 1531-1543.

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

Ebersberger, B., Herstad, S.J. & Koller, C. (2014) Does the composition of regional knowledge bases influence extra-regional collaboration for innovation? *Applied Economics Letters*, 21(3): 201-204.

Eriksson, R.H. (2011) Localized spillovers and knowledge flows: How does proximity influence the performance of plants? *Economic Geography*, 87(2): 127-152.

Firgo, M. & Mayerhofer, P. (2018) (Un)related variety and employment growth at the sub-regional level. *Papers in Regional Science*, 97(3), 519-547.

Frenken, K., van Oort, F. & Verburg, T. (2007) Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41(5): 685-697.

Fritsch, M. & Kublina, S. (2018) Related variety, unrelated variety and regional growth: The role of absorptive capacity and entrepreneurship. *Regional Studies*, 52(10): 1360-1371.

Guo, Q., He, C. & Li, D. (2016) Entrepreneurship in China: The role of localisation and urbanisation economies. *Urban Studies*, 53(12): 2584-2606.

Hartog, M., Boschma, R. & 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.

Howell, A., He, C., Yang, R. & Cindy, C. (2018) Agglomeration, (un)-related variety and new firm survival in China: Do local subsidies matter? *Papers in Regional Science*, 97(3): 485-501.

Lazzeretti, L., Innocenti, N. & Capone, F. (2017) The impact of related variety on the creative employment growth. *Annals of Regional Science*, 58(3): 491-512.

Liang, J. & Goetz, S.J. (2018) Technology intensity and agglomeration economies. *Research Policy*, 47(10): 1990-1995.

Miguelez, E. & Moreno, R. (2018) Relatedness, external linkages and regional innovation in Europe. *Regional Studies*, 52(5): 688-701.

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.

Saviotti, P.P. & Frenken, K. (2008) Export variety and the economic performance of countries. *Journal of Evolutionary Economics*, 18(2): 201-218.

Sedita, S.R., De Noni, I., & Pilotti, L. (2017) Out of the crisis: An empirical investigation of place-specific determinants of economic resilience. *European Planning Studies*, 25(2): 155-180.

Tavassoli, S. & Carbonara, N. (2014) The role of knowledge variety and intensity for regional innovation. *Small Business Economics*, 43(2): 493-509.

Tavassoli, S. & Jienwatcharamongkhol, V. (2016) Survival of entrepreneurial firms: The role of agglomeration externalities. *Entrepreneurship and Regional Development*, 28(9-10): 746-767.

van Oort, F., de Geus, S. & Dogaru, T. (2015) Related variety and regional economic growth in a cross-section of European urban regions. *European Planning Studies*, 23(6): 1110-1127.

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

Other References Cited

Audretsch, D.B. & Feldman, M.P. (1996) R&D spillovers and the geography of innovation and production. *American Economic Review*, 86(3): 630-640.

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.

Bathelt, H. & Buchholz, M. (2019) Outward foreign-direct investments as a catalyst of urban-regional income development? Evidence from the United States. *Economic Geography*, 95(5): 442-466.

Bathelt, H. & Glückler, J. (2000) Netzwerke, Lernen und evolutionäre Regionalentwicklung (Networks, learning and evolutionary regional development). *ZFW (German Journal of Economic Geography)*, 44(3-4): 167-182.

Bathelt, H. & Henn, S. (2014) The geographies of knowledge transfers over distance: Toward a typology. *Environment and Planning A*, 46(6): 1403-1424.

Bathelt, H., Malmberg, A. & Maskell, P. (2004) Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 28(1): 31-56.

Boschma, R. (2017) Relatedness as a driver of regional diversification: A research agenda. *Regional Studies*, 51(30): 351-364.

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. & Frenken, K. (2011a) Technological relatedness and regional branching. In Bathelt, H., Feldman, M.P. & Kogler, D.F. (eds) *Beyond Territory: Dynamics Geographies of Knowledge Generation, Diffusion, and Innovation* (pp. 64-81). Oxford: Oxford University Press.

Boschma, R. & Frenken, K. (2011b) The emerging empirics of evolutionary economic geography. *Journal of Economic Geography*, 11(2): 295-307.

Boschma, R.A. & Lambooy, J.G. (1999) Evolutionary economics and economic geography. *Journal of Evolutionary Economics*, 9(4): 411-429.

Boschma, R.A. & Martin, R. (eds) (2010a): *Handbook of Evolutionary Economic Geography*. Cheltenham, Northampton, MA: Edward Elgar.

Boschma, R.A. & Martin, R. (2010b) The aims and scope of evolutionary economic geography. In Boschma, R.A. & Martin, R. (eds) *Handbook of Evolutionary Economic Geography* (pp. 3-39). Cheltenham, Northampton, MA: Edward Elgar.

Breul, M., Broekel, T. & Brachert, M. (2021) Die Treiber der räumlichen Emergenz und Konzentration der Photovoltaik-Industrie in Deutschland (The drivers of the spatial emergence and clustering of the photovoltaic industry in Germany). *ZFW (German Journal of Economic Geography)*, 59(3): 133-150.

Buchholz, M. (2019) Organizations, institutions and networks in local scenes: The growth of San Francisco Bay Area punk rock. *Geoforum*, 103: 158-166.

Buchholz, M. & Bathelt, H. (2021) Models of regional economic development: Illustrations using U.S. data. *ZFW (The German Journal of Economic Geography)*, 65(1): 28-42.

Cantwell, J.A. (1989) *Technological Innovation and Multinational Corporations*. Oxford: Basil Blackwell.

Chinitz, B. (1961) Contrasts in agglomerations: New York and Pittsburgh. *American Economic Review*, 51(2): 279-289.

Content, J. & Frenken, K. (2016) Related variety and economic development: A literature review. *European Planning Studies*, 24(12): 2097-2112.

Cooke, P., Uranga, M.G., Extebarria, G. (1997) Regional innovation systems: Institutional and organizational dimensions. *Research Policy*, 26(4-5): 475-491.

Crescenzi, R. & Iammarino, S. (2017) Global investments and regional development trajectories: The missing links. *Regional Studies*, 51(1): 97-115.

Deegan, J., Broekel, T. & Fitjar, R.D. (2021) Searching through the haystack: The relatedness and complexity of priorities in smart specialization strategies. *Economic Geography*, 97(5): 497-520.

Delgado, M., Porter, M. & Stern, S. (2016). Defining clusters of related industries. *Journal of Economic Geography*, 16(1): 1-38.

Delgado, M. & Mills, K.G. (2020) The supply chain economy: A new industry categorization for understanding innovation in services. *Research Policy*, 49: #104039

Duranton, G. & Puga, D. (2004) Microfoundations of urban agglomeration economies. In Henderson, J.V. & Thisse, J.-F. (eds) *Handbook of Regional and Urban Economics, Vol. 4* (pp. 2063-2117). Amsterdam: Elsevier.

Duranton, G. & Puga, D. (2005) From sectoral to functional urban specialization. *Journal of Urban Economics*, 57(2): 343-370.

Essletzbichler, J. & Rigby, D. (2007) Exploring evolutionary economic geographies. *Journal of Economic Geography*, 7(5): 549-571.

Florida, R.L. (2002) *The Rise of the Creative Class: And How It's Transforming Work, Leisure, Community and Everyday Life*. New York, NY: Basic Books.

Florida, R.L. (2017) *The New Urban Crisis: How Our Cities Are Increasing Inequality, Deepening Segregation, and Failing the Middle Class – And What We Can Do About It*. New York, NY: Basic Books.

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

Glückler, J. & Bathelt, H. (2017) Institutional context and innovation. In Bathelt, H., Cohendet, P., Henn, S. & Simon, L. (eds) *The Elgar Companion to Innovation and Knowledge Creation* (pp. 121-137). Cheltenham, Northampton, MA: Edward Elgar.

Gong, H. & Hassink, R. (2020) Context sensitivity and economic-geographic (re)theorising. *Cambridge Journal of Regions, Economy and Society*, 13(3): 475-490.

Hall, P.A. & Soskice, D.W. (eds) (2001) *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford, New York, NY: Oxford University Press.

Henning, M. (2019) Time should tell (more): Evolutionary economic geography and the challenge of history. *Regional Studies*, 53(4), 602-613.

Hidalgo, C.A., Balland, P.A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Glaeser, E., He, C., Kogler, D.F., Morrison, A., Neffke, F., Rigby, D., Stern, S., Zheng, S. & Zhu, S. (2018) The principle of relatedness. In Morales, A.J., Gershenson, C., Brah, D., Minai, A.A. & Bar-Yam, Y. (eds) *Unifying Themes in Complex Systems IX: Proceedings of the Ninth International Conference on Complex Systems* (pp. 451-457). Cham: Springer VS.

Iammarino, S., Rodriguez-Pose, A. & Storper, M. (2019) Regional inequality in Europe: Evidence, theory and policy implications. *Journal of Economic Geography*, 19(2): 273-298.

Jacobs, J. (1969) *The Economy of Cities*. New York, NY: Random House.

Juhász, S., Broekel, T. & Boschma, R. (2021) Explaining the dynamics of relatedness: The role of co-location and complexity. *Papers in Regional Science*, 100(1): 3-21.

Kemeny, T. & Storper, M. (2015) Is specialization good for regional economic development? *Regional Studies*, 49(6): 1003-1018.

Ketels, C. (2016) Smart specialisation: Opportunities and Challenges for Regional Innovation Policy. *Competitiveness Review*, 26(4): 476-478.

Kuusk, K. & Martynovich, M. (2021) Dynamic nature of relatedness, or what kind of related variety for long-term regional growth. *Tijdschrift voor Economische en Sociale Geografie*, 112(1): 81-96.

Li, P. & Bathelt, H. (2018) Location strategy in cluster networks. *Journal of International Business Studies*, 49(8): 967-989.

Li, P. & Bathelt, H. (2021) Spatial knowledge strategies: An analysis of international investments using fuzzy set Qualitative Comparative Analysis (fsQCA). *Economic Geography*, 97(4): 366-389.

Malecki, E.J. (2010) Global knowledge and creativity: New challenges for firms and regions. *Regional Studies*, 44(8): 1033-1052.

Malmberg, A. & Maskell, P. (2002) The elusive concept of localization economies: Towards a knowledge-based theory of spatial clustering. *Environment and Planning A*, 34(3): 429-449.

Marshall, A. (1920) *Principles of Economics*. 8th edn., London: Macmillan and Co.

Martin, R. & Sunley, P. (2006) Path dependence and regional economic evolution. *Journal of Economic Geography*, 6(4): 395-437.

Martin, R. & Sunley, P. (2022) Making history matter more in evolutionary economic geography. *ZFW – Advances in Economic Geography*, 66(2): 65-80.

Mokyr, J. (1990) *The Lever of Riches: Technological Creativity and Economic Progress*. New York, NY, London: Oxford University Press.

Mudambi, R. (2008) Location, control and innovation in knowledge-intensive industries. *Journal of Economic Geography*, 8(5): 699-725.

Perez, C. (2010) Technological revolutions and techno-economic paradigms. *Cambridge Journal of Economics*, 34(1): 185-202.

Petralia, S.G. (2020) Mapping general-purpose technologies with patent data. *Research Policy*, 49(7): #104013.

Porter, M.E. (1990) *The Competitive Advantage of Nations*. New York, NY: Free Press.

Reardon, S.F. & Firebaugh, G. (2002) Measures of multigroup segregation. *Sociological Methodology*, 32(1): 33-67.

Saxenian, A.L. (1985) The genesis of Silicon Valley. In Hall, P. & Markusen, A. (eds) *Silicon Landscapes* (pp. 20-34). Boston, MA, London, Sydney: Allen and Unwin.

Saxenian, A.L. (1994) *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA, London: Harvard University Press.

Scott, A.J., Storper, M. (1987). "High Technology Industry and Regional Development: A Theoretical Critique and Reconstruction." International Social Science Journal 112: 215-232, May.

Soskice, D.W. (2021) The US as radical innovation driver: The politics of declining dominance. In Hacker, J., Hertel-Fernandez, A., Pierson, P. & Thelen, K. (eds) *The American Knowledge Economy: Politics, Markets, and Power* (pp. 323-352). Cambridge, New York, NY: Cambridge University Press.

Spencer, G.M., Vinodrai, T., Gertler, M.S. & Wolfe, D.A. (2010) Do clusters make a difference? Defining and assessing their economic performance. *Regional Studies*, 44(6): 697-715.

Storper, M. (1997) *The Regional World: Territorial Development in a Global Economy*. New York, NY: Guilford Press.

Storper, M. (2009) Regional context and global trade. *Economic Geography*, 85(1): 1-21.

Storper, M., Kemeny, T., Makarem, N. & Osman, T. (2015) *The Rise and Fall of Urban Economies: Lessons from San Francisco and Los Angeles*. Stanford, CA: Stanford University Press.

Storper, M. & Walker. R. (1989) *The Capitalist Imperative: Territory, Technology, and Industrial Growth*. New York, NY: Basil Blackwell.

Theil, H. (1972) *Statistical Decomposition Analysis*. Amsterdam: North-Holland Publishing.

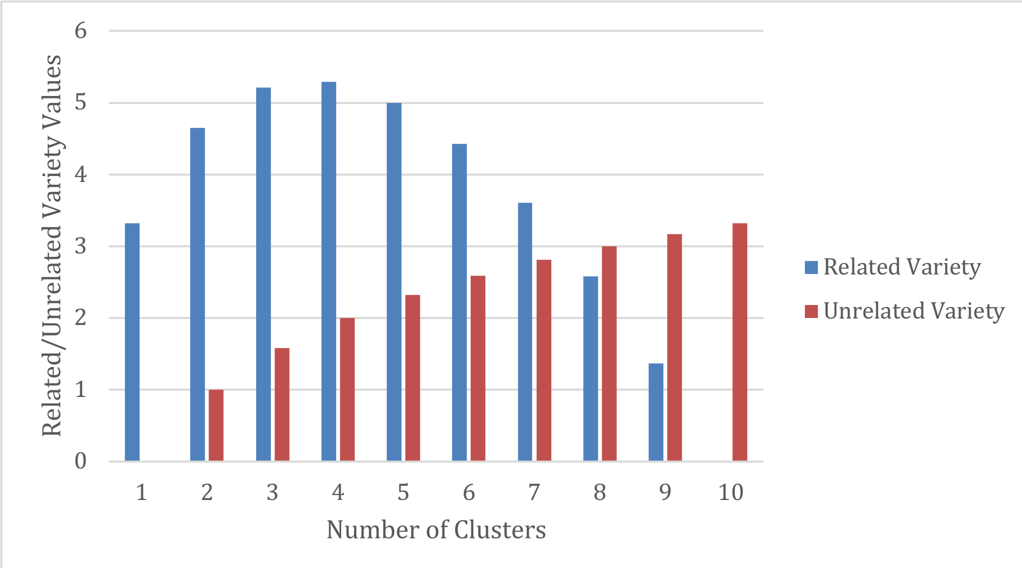
Timmer, M.P., Miroudot, S. & de Vries, G.J. (2019) Functional specialisation in trade. *Journal of Economic Geography*, 19(1): 1-30.

U.S. Bureau of Economic Analysis (2019) *Regional Economic Accounts* [Dataset]. Washington, DC. https://apps.bea.gov/regional/downloadzip.cfm (accessed 3 March 2019).

U.S. Bureau of Labor Statistics (2020) *Quarterly Census of Employment and Wages – QCEW. QCEW Data Files, 1993–2017* [Dataset]. Washington, DC. https://www.bls.gov/cew/datatoc.htm (accessed 21 May 2020).

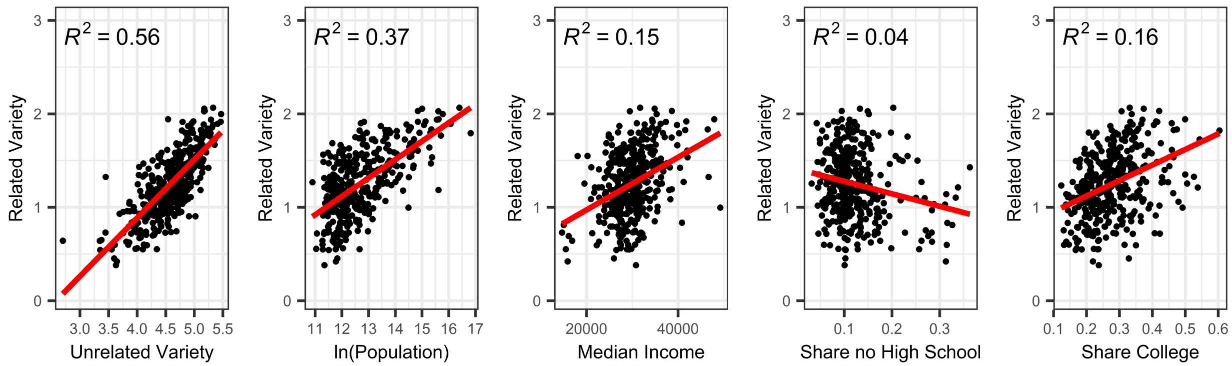
Yeung, H.W.-c. (2021) Regional worlds: From related variety in regional diversification to strategic coupling in global production networks. *Regional Studies*, 55(6): 989-1010.

Figure 1. Related and Unrelated Variety Values by Regional Cluster Structure (Scenarios)



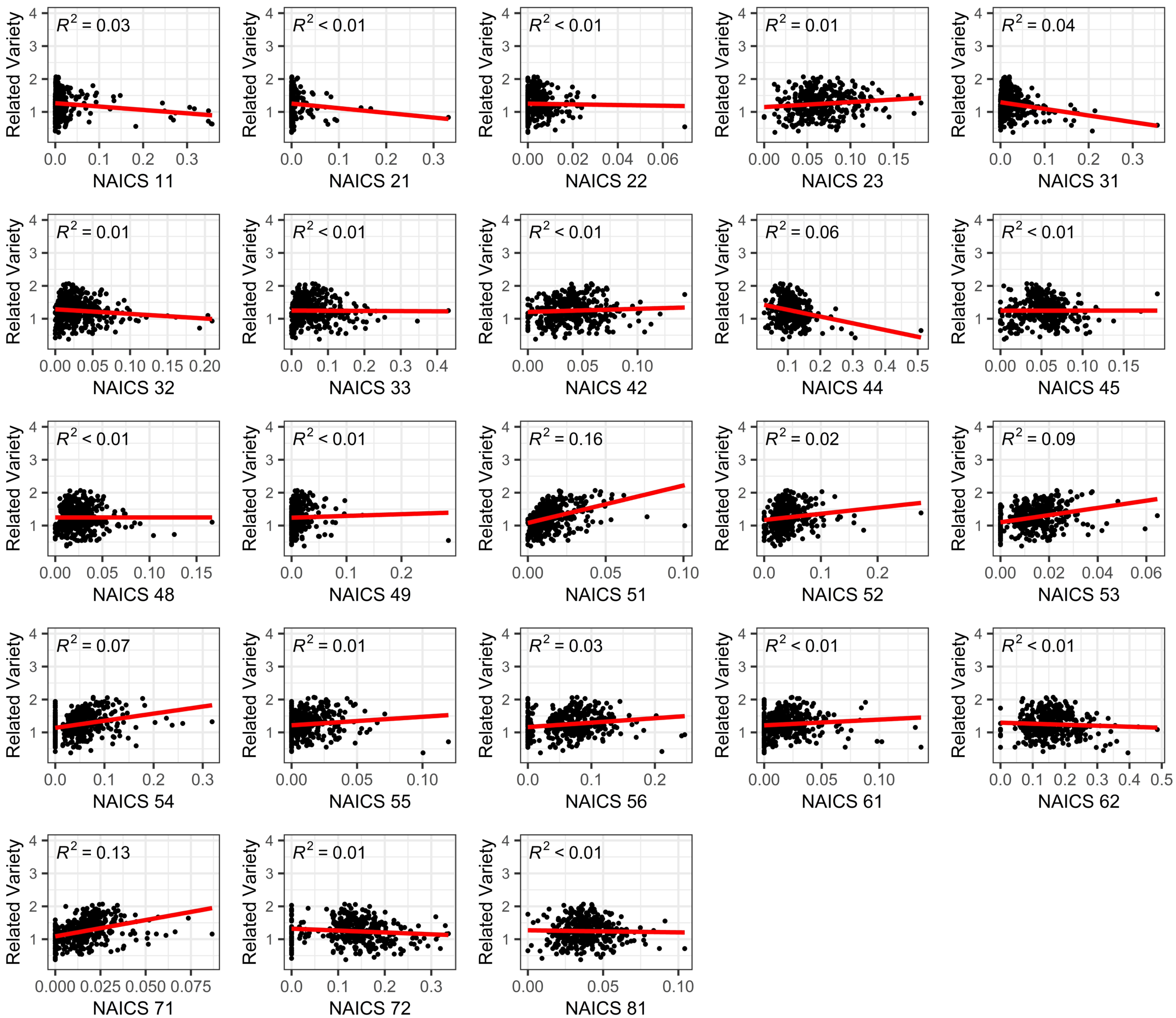
Notes: Related and unrelated variety scores are computed for a hypothetical region with 100,000 workers that are employed in 10 main sectors, each of which has 10 sub-sectors. Scenario 1 assumes that all employees work in 1 main sector, split equally across its 10 sub-sectors; scenario 10 assumes that employment is equally split both across the 10 main sectors and within each across the 10 corresponding sub-sectors.

Figure 2. Scatterplots of Related Variety Over Unrelated Variety, ln(Population), Median Income and Education Levels for U.S. MSAs, 2017 (Source: U.S. Bureau of Economic Analysis 2019; U.S. Bureau of Labor Statistics 2020)



Note: Unrelated variety was computed based on 3-digit, related variety based on 5-digit NAICS codes.

Figure 3. Scatterplots of Related Variety Over Employment Shares in 2-Digit NAICS Codes for U.S. MSAs, 2017 (Source: U.S. Bureau of Economic Analysis 2019; U.S. Bureau of Labor Statistics 2020)



Note: Unrelated variety was computed based on 3-digit, related variety based on 5-digit NAICS codes.

Table 1. Characteristics of the Top-30 Most-Cited Papers on Related Variety According to SCOPUS Citations (29 November 2021)

Scopus citations	Authors (year)	Number of observations	Dependent variable(s)	Regional scale/size included?	Panel or cross-section	Main findings	Specific geographical analysis?	Analysis of underlying, mechanisms, institutions, alternative explanations?
1309	Frenken, van Oort & Verburg (2007)	40 Dutch NUTS 3 regions	a) Employment growth (1996-2002); b) productivity growth (1996-2001)	No, but population density*)	Cross-section	a) Related variety has positive, highly significant impact on employment growth; unrelated variety negative and insignificant; b) related variety has negative, significant impact on productivity growth; unrelated variety negative and insignificant	No	No
585	Neffke, Henning & Boschma (2011)	Plant-level data of 70 Swedish regions; 72,100 membership observations (industry-region combinations)	a) Membership probability (industry stays in region); b) entry probability (entry in 5 years); c) exit probability (exit in 5 years) (1969, 1974, 1979, 1984, 1989, 1994)	No	Equivalent to panel	a) Regional closeness (based on revealed relatedness) has positive, highly significant impact on membership; extra-regional closeness negative and highly significant; b) regional closeness has positive, highly significant impact on entry; extra-regional closeness negative and highly significant; c) regional closeness has negative, highly significant impact on exit; extra-regional closeness negative and highly significant	Largely descriptive, "arbitrary case study" of Linköpings revealed relatedness in industrial transition	No
525	Boschma & Iammarino (2009)	103 Italian NUTS 3 regions	a) Employment growth (1995-2003); b) value-added growth (1995-2003); c) labor productivity growth (1995-2003)	No, but population density	Cross-section	a) Related and unrelated variety of exports have positive, but insignificant impact on employment growth; b) related variety has positive, highly significant impact on value-added growth; unrelated variety positive and insignificant; c) related variety has positive, significant impact on labor-productivity growth; unrelated variety positive and insignificant	Macro-regional control variables included, but no specific discussion	No
209	Boschma, Eriksson & Lindgren (2009)	17,098 job moves to plants in Swedish regions	Plant-level labor productivity growth (2001-2003)	Yes, for firms and regions	Cross-section	Related skill variety of inflowing labor has a positive, highly significant impact on labor productivity growth; unrelated skill variety negative and insignificant	No, essentially not a geographical study	No, but firm-level controls
195	Boschma, Minondo & Navarro (2011)	50 Spanish provinces over three 4-year intervals (150 observations)	Value added growth across (three 4-year time periods 1995-2007)	No, but population density	Equivalent to short panel	Related variety has a positive, highly-significant impact on value-added growth; unrelated variety negative and insignificant	No	No
183	Castaldi, Frenken & Los (2015)	51 U.S. states over 22 years (877 observations)	a) Number of patents; b) number of superstar patents (1977-1999)	No	Panel	a) Related variety has positive, significant effect on patents; unrelated variety negative and insignificant; b) related variety has positive, insignificant effect on superstar patents; unrelated variety positive and highly significant	Some models include a spatial variable; some description of spatial variations of variables; no specific geographical analysis	No
125	Saviotti & Frenken (2008)	20 OECD countries over eight 5-year periods (156 observations)	a) GDP per capita growth; b) labor productivity growth (in eight 5-year periods 1964-2003)	No	Equivalent to panel	a) Related variety of exports has positive, highly significant impact on GDP per capita growth; unrelated variety negative and highly significant; b) related variety of exports has positive, highly significant impact on labor productivity growth; unrelated variety negative and significant	County controls in some models; some description of country export trends, but not specific analysis	No
87	Eriksson (2011)	8,313 plants in Swedish economy (located in differently sized regions)	Plant labor productivity (2001-2003)	No, but population density	Cross-section	0.5 km regions: related and unrelated variety have both negative, but insignificant impact on productivity growth; 5 km and 50 km regions: related variety has positive, highly significant impact on productivity growth; unrelated variety remains negative and insignificant	Separate models for different region sizes; no specific geographical analysis	No; but some controls
84	Hartog, Boschma & Sotarauta (2012)	67 Finnish NUTS 4 regions over 14 years (875 observations)	Employment growth (1993-2006)	No, but population density	Panel	Related variety has positive, but insignificant impact on employment growth; unrelated variety negative and insignificant; however when computed separately for different industries, related variety among high-tech sectors has positive, significant impact	No	No

75	Aarstad, Kvitastein & Jakobsen (2016)	6,584 enterprises in 89 economic-geographical regions in Norway	a) Enterprise productivity (2010); b) innovation occurrence (2008-2010)	No, but population density	Cross-section	a) Related variety has positive, but insignificant impact on enterprise productivity; unrelated variety negative and highly significant; b) related variety has positive, significant impact on enterprise innovation; unrelated variety negative and insignificant	No, essentially not a geographical study	No, but firm-level controls
68	van Oort, de Geus & Dogaru (2015)	205 NUTS 2 regions in 15 EU countries	a) Employment growth; b) labor productivity per employee; c) unemployment growth (2000-2010)	No, but population density	Cross-section	a) Related variety has positive, significant impact on employment growth; unrelated variety also positive and significant; b) related variety has negative, insignificant impact on productivity growth; unrelated variety positive and insignificant; c) related variety has positive, insignificant impact on unemployment growth; unrelated variety also positive and insignificant	Use of spatial lags and differentiation by region size; some description of spatial variations of variables; no specific geographical analysis	No, but numerous controls
68	Boschma, Eriksson & Lindgren (2014)	72 Swedish functional regions over 5 years (360 observations)	a) Productivity growth; b) employment growth; c) unemployment growth (1998-2002)	No, but population density	Panel	a) Related and unrelated variety of labor market flows have positive, but insignificant impacts on productivity growth; b) related variety has positive, significant impact on employment growth; unrelated variety negative and significant; c) related and unrelated variety have negative, but insignificant impacts on unemployment growth	No	No, but some controls
67	Tavassoli & Carbonara (2014)	81 Swedish functional regions over 6 years (486 observations)	Number of patent applications per year (2002-2007)	No	Panel	Related variety has positive, highly significant impact on patent applications; unrelated variety positive, but insignificant	Some description of spatial variations of variables; no specific geographical analysis	No, but numerous controls
61	Cainelli & Iacobucci (2012)	87,688 firms in 103 Italian provinces	Firm-level vertical integration index (2001)	Yes	Cross-section	Vertical related variety has negative, highly significant impact on firm-level vertical integration; unrelated variety positive and highly significant	Analysis separately for macro-regions, but no specific discussion	No, but many industry dummies and other controls
60	Antonietti & Cainelli (2009)	715 Italian manufacturing firms in 103 Italian provinces	a) R&D investment per employee (2003); b) firm propensity to innovate (2001-2003); c) total factor productivity (2003); d) firm propensity to export (2001-2003)	No, but population density	Cross-section	a) Related variety has positive, highly significant impact on R&D investment per employee; unrelated variety negative and highly significant; b) related variety has negative, but insignificant impact on firm propensity to innovate; unrelated variety positive and insignificant; c) related variety has negative, but insignificant impact on total factor productivity; unrelated variety positive and insignificant; d) related variety has positive, but insignificant impact on firm propensity to export; unrelated variety negative and insignificant	No	No, but some industry dummies and other controls
48	Caragliu, de Dominicis & de Groot (2016)	3,614 European firms in 259 European NUTS 2 regions	Percentage employment change (1990-2007)	No	Cross-section	Related variety has negative, but insignificant impact on employment growth; unrelated variety positive and moderately significant	No, but separate models for regions with different density	Industry dummies and other controls; separate sector models
44	Guo, He & Li (2016)	162 sectors in 286 Chinese prefecture-level city-regions over 7 years (35,000-40,000 observations annually)	Newly started privately-owned firms (2001-2007)	No	Cross-section (in each of 7 years)	Related variety has positive, highly significant impact on firm formation in each of 7 consecutive years; unrelated variety positive and highly significant in 2001, but negative and highly significant in 2007, and insignificant in between	No	No; some controls
43	Miguelez & Moreno (2018)	255 European NUTS 2 regions over 9 years (2,219 observations)	a) Patents per capita; b) patent quality: patents weighted by citations (1999-2007)	No	Panel	a) Related variety has positive, highly significant impact on patenting; unrelated variety negative, but insignificant; b) related variety has positive, highly significant impact on patent quality; unrelated variety also positive and highly significant	No	No; some controls
40	Sedita, De Noni & Pilotti, L. (2017)	686 Italian local labor systems	Growth in employment rate (2009-2013); viewed as regional resilience	No, but population density	Cross-section	Related variety has positive, highly-significant impact on growth of employment rate; unrelated variety negative, but insignificant; most interactions between related and unrelated variety and knowledge base variable (share of corresponding industries) insignificant	Macro-regional control variables and industrial district dummy included; some description of spatial variations of variables; no specific geographical analysis	No

36	Cortinovis & van Oort (2015)	260 European NUTS 2 regions over 9 years (2340 observations)	a) Employment growth; b) unemployment growth; c) gross value-added per hour (productivity) growth (2004-2012)	No, but population density	Panel	a) Related variety has negative, highly significant impact on employment growth; unrelated variety negative, but insignificant; b) related variety has positive, but insignificant impact on unemployment growth; unrelated variety also positive and insignificant; c) related variety has negative, but insignificant impact on productivity growth; unrelated variety also negative and insignificant	Separate models for high-tech, medium-tech, low-tech regions; some description of spatial variations of variables; no specific geographical analysis	No
36	Wixe & Andersson (2017)	290 Swedish municipalities	a) Employment growth; b) productivity growth (2002-2007)	No, but population density	Cross-section	a1) Related industry variety has positive, highly significant impact on employment growth; unrelated industry variety negative, but insignificant; a2) related educational variety has negative, but insignificant impact on employment growth; unrelated educational variety positive, but insignificant; a3) related and unrelated occupational variety are both insignificant; b1) related industry variety has negative, highly significant impact on productivity growth; unrelated industry variety negative, but insignificant; b2) related educational variety has positive, highly significant impact on productivity growth; unrelated educational variety positive and significant; b3) related and unrelated occupational variety are both insignificant	No	No; but separate models for manufacturing and services
30	Basile, Pittiglio & Reganati (2017)	164,113 start-up firms in 686 local labor systems (455,000 observations in 3 cohorts)	Likelihood of firm exit (by 2010) for start-up firms (started in 2004, 2005, 2006)	For industries not regions, but population density	Panel	Related variety has positive, but insignificant impact on likelihood of firm exit; unrelated variety negative and highly significant; in manufacturing, related variety reduces the likelihood of firm exits with moderate significance; while unrelated variety is positive, but insignificant in manufacturing	Spatial NUTS 2 dummies included; no specific geographical analysis	No, but many industry and other controls used
27	Fritsch & Kublina (2018)	71 West German planning regions over seven 5-year periods (497 observations)	Regional employment growth (over seven 5-years periods 1999-2008)	No, but population density	Panel	Related variety has a positive, highly-significant impact on employment growth; unrelated variety also positive and highly significant	Macro-regional control variables; some description of spatial variations of variables; no specific geographical analysis	No, some controls
27	Howell, He, Yang & Cindy (2018)	135,000 Chinese new manufacturing firms in 333 pre-fecture-level city-regions (332,500 observations)	Duration of firm survival (1998-2007)	For firms not regions, but labor density	Panel	Related variety has positive, moderately significant impact on firm survival; unrelated variety negative and moderately significant	No; regional dummies used, but essentially not a geographical study	No; firm-level controls and industry dummies used
23	Liang & Goetz (2018)	3,147 U.S. counties	Employment growth (2003-2013)	No, but population density	Cross-section	Related variety has positive, highly significant impact on employment growth; unrelated variety negative and highly significant; interaction effect related variety * technology intensity positive and highly significant	Some description of spatial variations of variables; no specific geographical analysis	No
23	Tavassoli & Jienwatchara-mongkhol (2016)	4682 Swedish knowledge business services firms in 72 functional regions	Hazard of firm exit (1997-2012)	Yes	Cross-section	Related variety has negative, highly significant impact on firm exit hazard; unrelated variety negative and significant	No	No, but numerous individual- and firm-level controls
21	Firgo & Mayerhofer (2018)	81 Austrian labor market districts in two periods (162 observations)	Employment growth (2000-2006; 2007-1013)	No, but population density	Combined cross-sections of two periods	Related variety has positive, significant impact on employment growth; unrelated variety positive and highly significant; in the services sector similar relations are found whereas in manufacturing neither variable is significant	Differentiated models for urban and rural/industrial regions; but no specific geographical explanation	No, but numerous controls
19	Ebersberger, Herstad & Koller (2014)	34,892 region-technology combinations in European NUTS 3 regions over six 5-year periods (209,352 observations)	a) Interregional domestic collaboration in patenting; b) international collaboration (over six 5-year intervals 1980-2010)	Yes	Equivalent to panel	a) Related technological variety has negative, highly significant impact on interregional collaboration; b) it has positive, highly significant impact on international collaboration	No	No

18	Content, Frenken & Jordaan (2019)	204 European NUTS 2 regions	a) Share of regional entrepreneurs; b) share of opportunity-driven entrepreneurs; c) share of necessity-driven entrepreneurs (2007-2014)	No, but population density	Cross-section	a) Related variety has positive, but insignificant impact on entrepreneurship; unrelated variety negative and highly significant; c) similar results for necessity-driven entrepreneurship; b) but for opportunity-driven entrepreneurship impact of related variety positive and significant	Political economy type controls; some description of spatial variations of variables; no specific geographical analysis	No
17	Lazzeretti, Innocenti & Capone (2017)	103 Italian provinces	Employment difference in cultural/creative industries (1991-2001; 2001-2011; 1991-2011)	No, but population density	Cross-sections for different time periods	Related variety has positive, highly significant impact on employment growth in cultural/creative industries; unrelated variety negative, but insignificant (1991-2011); cross-sections 1991-2011 and 2001-2011 essentially support this findings	Macro-regional control variables; some description of spatial variations of variables; spatial lag and error models; no specific geographical analysis	No, but some controls

Notes: In the first step of the literature search, we used the search string TITLE-ABS-KEY (“related variety”) AND (LIMIT-TO (SUBJAREA, “SOCI”) OR LIMIT-TO (SUBJAREA, “ECON”) OR LIMIT-TO (SUBJAREA, “BUSI”)). In the second step, we removed articles from unrelated fields, such as linguistics. Third, we removed all papers that were not analytical in nature.

*) The rationale to include population density is to consider urbanization economies, not scale effects (although it is unclear whether this is an adequate indicator of urbanization economies).

Table 2. Impact of Related Variety Variables on Per-Capita Income and Employment in U.S. MSAs, 2010 – 2017, Cross-Sectional and Longitudinal Effects

Dependent Variables / Independent Variables	ln(Per-Capita Income) Growth 2010-2017 (1)	ln(Employment) Growth 2010-2017 (2)	ln(Per-Capita Income) Change 2010-2017 (3)	ln(Employment) Change 2010-2017 (4)
Intercept	-0.004	-0.016	0.034***	0.099***
	(0.005)	(0.010)	(0.003)	(0.009)
Related Variety 2010	-0.001	0.004**		
	(0.001)	(0.002)		
Unrelated Variety 2010	0.003***	0.000		
	(0.001)	(0.003)		
ln(Employment Density) 2010	0.000	-0.001**		
	(0.000)	(0.001)		
Related Variety Change 2010-2017			0.003	-0.022
			(0.010)	(0.036)
Unrelated Variety Change 2010-2017			-0.029*	0.108
			(0.016)	(0.068)
ln(Employment Density) Change 2010-2017			0.081***	
			(0.027)	
R^2	0.039	0.035	0.060	0.023
Number of MSAs	338	338	338	338

Notes: Linear cross-sectional and panel regression analyses – units of analysis are MSAs. NAICS 2-digit industries were used as a basis when computing related variety measures. Per-capita income is defined as the natural logarithm of total (2017 inflation-adjusted) annual wage and salary income in a MSA divided by total employment. Columns 1 and 2 refer to the growth rates of per-capita income and employment in the 2010-2017 period, columns 3 and 4 to the respective differences. Heteroscedasticity-robust standard errors are reported in parentheses. ***, ** and * correspond to p-values less than 0.01, 0.05 and 0.10, respectively.

Sources: Buchholz & Bathelt (2021, p. 32) based on data from U.S. Bureau of Economic Analysis (2019); U.S. Bureau of Labor Statistics (2020).