Related Variety and Regional Development: A Critique

Harald Bathelt

Department of Geography and Planning

University of Toronto

Toronto, Ontario M5S 3G3

Canada

harald.bathelt@utoronto.ca

Michael Storper

Department of Geography and Environment

London School of Economics

London WC2A 2AE

UK

and

Luskin School of Public Affairs

UCLA

Los Angeles, CA 90095-1656

m.storper@lse.ac.uk

Key words:

economic geographies of places

evolutionary economic geography (EEG)

regional development

regional specialization

related variety

JEL codes:

L23

R11

Acknowledgments

This article 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 three reviewers and editor Jim Murphy, for constructive criticism. We are also indebted to Adams Aghimien, Max Buchholz, and He (Shawn) Shuang for superb research assistance.

Abstract

Evolutionary approaches in economic geography have contributed substantially to the growing body of knowledge of regional development processes and their underlying mechanisms. One key concept in the literature on evolutionary economic geography is that of *related variety*. Herein, regional industry structure is represented through the level of related variety of technologies, skills, or outputs. The related variety concept proposes that regional economic development is favored when an economy diversifies into products or technologies that are closely related to the stock of existing activities. In this article, we raise substantive questions regarding the internal logic of the concept of related variety, its spatial expressions, measurement specifics, empirical regularities and biases, and its possible short-

and long-term effects on regional development. Based on this investigation, we make suggestions for improvements to future research.

There is a long tradition of thought about regional economic development that sees it as a process whereby economies evolve gradually along pathways that are dependent on the existing stock of economic activities or knowledge bases, or, conversely, how they may break away from these pathways and reinvent themselves into something radically different (Chinitz 1961; Storper and Walker 1989; Boschma and Lambooy 1999; Bathelt and Glückler 2000). The former is the central concern of a prominent school of evolutionary economic geography (EEG), an approach that can be traced back to the seminal work of Boschma and Frenken (2006) (Martin and Sunley 2006; Essletzbichler and Rigby 2007; Boschma and Martin 2010a). As Boschma and Martin (2010b: 6-7) describe, the goal of this approach is to investigate "the processes by which the economic landscape—the spatial organization 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 case studies of regions, clusters, and innovation systems toward comparative statistical and analytical work that aims to generalize 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 article does not aim to review this entire corpus of EEG research, but rather to scrutinize the development and deployment of one of its central concepts, that of *related variety* in technologies, industries, and activities.[1]

[1] Other work in EEG that focuses on topics, such as regional path development, path creation, or lock-in (Coenen et al. 2017; Hassink, Isaksen, and Trippl 2019; MacKinnon et al. 2019), is often based on qualitative studies and does not employ the related variety concept empirically.

A starting point for this literature is the pathbreaking study by Frenken, van Oort, and Verburg (2007) on the determinants of regional growth in the Netherlands. This study led to a stream of publications that, broadly speaking, suggest that regional economies that are characterized by high levels of related variety experience more growth or positive development with respect to productivity, employment, innovation, or patenting than those regions with low levels of related variety. Studies by Content and Frenken (2016) and Boschma (2017), among others, summarize the main accomplishments of this body of research, identify missing aspects and suggest directions for future work.

EEG has also developed an increasingly influential position on regional development policy, especially in the EU, 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. As such, expanding related variety is quite different from the notion that a region should be widely diversified in what it does as a hedge against downturns. In a recent assessment, Martin and Sunley (2022, 70) demur, warning 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." Studies in EEG often imply that relatedness is a causal factor in tracing the specific contours and qualities of the development pathway of a region, its *roads taken and not taken*, whether these are favorable or not.

Notwithstanding the volume of research on related variety, relatively little research has thus far thoroughly assessed the veracity of the core assumption, that is, that increasing related variety is a key foundation of successful regional economic development and the decisive factor shaping roads taken and not taken (Henning 2019; Martin and Sunley 2022). In this article, we address this gap by examining the fundamental assumptions behind and characteristics of the related variety approach along multiple important dimensions. In doing this, we aim to bring together critical comments from previous work (e.g., Whittle and Kogler 2020) and combine this with novel arguments and empirical evidence into a comprehensive critique of the related variety approach. Though this body of research is currently moving in new and different directions, we focus here on the core literature and its findings on the role of related variety in regional development.

Our article begins by considering the logic and assumptions behind the related variety concept. We then examine the measurement and potential geographic expressions of related variety. Empirical studies rest on a set of assumptions that are 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 then correctly measure its effects on development. This is followed by a presentation of our own results on certain empirical regularities in the US that call into question some of the implications of this research. We conclude by emphasizing the need for a fuller conceptualization and discussion of the pathways of regional economic change and make suggestions for future studies. We pose this as a sympathetic critique whose goal is to suggest to the research community several areas to revisit, rethink and improve upon previous research, and possibly go back to the drawing board in conceptualizing future work. We provide suggestions on what we believe could be important starting points for future studies, but we do not offer complete answers to all the issues we raise. Our intention is to stimulate a renewed discussion of the concept, its underlying premises, and implications in the real world.

[level1] Related Variety and Regional Development: A Conceptual and Empirical Survey

Boschma and Frenken (2011a) extensively argue that technologically related industries are better for regional economic development than a collection of nonrelated industries (Boschma and Frenken 2006; Frenken, van Oort, and Verburg 2007). This speaks to a long-standing discussion in the regional development field about whether specialization or diversification processes are better for long-term regional development (Duranton and Puga 2005; Kemeny and Storper 2015). Frenken, van Oort, and Verburg (2007) attempt 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.[2] The assumption behind this is that related variety generates the benefit of having activities that are closely related, enough that they will naturally create knowledge spillovers among them and trigger technological dynamism through the combination and recombination of such knowledge, but that these activities will be broad enough to harvest the benefits of tapping into many different areas of economic change and hedging risks against downturns. Extending this technological argument, it is proposed that certain forms of geographic proximity and technological proximity are often linked, in that related activities are likely to (1) colocate, which leads to knowledge spillovers; and (2) coevolve from learning that goes along with these spillovers. Interactive learning is said to require that the cognitive, social, and geographic *distance* (a difference) between economic actors not be too great. A first observation in this article 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.

In terms of intellectual lineage, the related variety concept refers to the work of Marshall (1920), Jacobs (1969), and Audretsch and Feldman (1996) among others. It thus places itself in the context of researchers who have explored various dimensions of regional economic specialization, including clusters, agglomerations, and regional innovation systems (Cooke, Uranga, and Extebarria 1997; Malmberg and Maskell 2002; Duranton and Puga 2005). Most of that research is concerned with causal topics, such as why economies specialize; whether bigger metro areas are more diversified than

---

[2] While this research often emphasizes the diversification aspect over specialization; this could also be viewed from a clustering perspective.

smaller ones; and with outcomes, such as whether diversification better protects economies from downturns than specialization. The focus of related variety research is somewhat different, however, as it is more about the advantages conferred by a relatedly diverse economy as opposed to a conventionally specialized or diverse one. The spillover potential of technological proximity is assumed to be enhanced when related (but diverse) subsectors are geographically colocated. Whether or not the presumed knowledge spillovers among related activities are strengthened by colocation is said to depend on a number of factors. These include whether the sectors are new or mature, routinized or nonroutinized, their underlying innovation potential, and their inherent potential for exploiting complementarities.

Table 1 gives an overview of the thirty most-cited empirical journal publications on related variety according to Scopus citations (November 29, 2021). The table characterizes each article in terms of its geographic basis, number of observations, dependent variables, related and unrelated variety indicators, some key findings, the specific mechanisms that are considered, and the corresponding analysis of geographic outcomes and peculiarities in the results. Though this table presents only a snapshot of the existing literature on related variety, it provides a good starting point for analyzing the most influential work in this field of research. Most of the studies follow the above line of argument and the methodology Frenken, van Oort, and Verburg (2007) introduce, deploying 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 between 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 directly investigate the underlying assumptions mentioned above nor provide an in-depth explanation of underlying mechanisms. None of the thirty studies listed in Table 1 presents empirical results on or causal identification of the specific mechanisms that drive successful or less successful regions.[3] 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 in these studies.

[Table 1 about here]

It is also not clear as to whether relatedness leads to relationships between firms, within labor markets, or in innovation systems at the regional scale, and whether these are direct or indirect in nature, traded or untraded. For example, firms that operate in technologically related sectors or are vertically positioned within a value chain may indeed have natural potential to cooperate and generate synergies, but this does not mean that these firms actually collaborate within a region (Whittle and Kogler 2020). There is a broad literature that shows that intense collaborations in industrial agglomerations or clusters are rare and do not always flow from mere colocation (Bathelt, Malmberg, and Maskell 2004), since even firms that are colocated and are in technologically related fields (as defined and measured by the related variety concept) may still operate in different fields of application, hence different markets. As such, they may not benefit from close cooperation or not even be interested in it. Moreover, even when technological proximity exists and benefits from collaboration within the same

[3] Content, Frenken, and Jordaan (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 and Frenken 2016). An exception is the study by Miguelez and Moreno (2018), which directly tackles underlying mechanisms rather than assuming them.

region are possible, firms may not exploit such opportunities (Breul, Broekel, and Brachert 2015; Essletzbichler 2015) since, for example, they may distrust each other or see one another primarily as rivals. In other words, identifying the logical possibility of vertical linkages or technological proximity leaves too many characteristics of firms unobserved to justify assuming that they would automatically collaborate and engage in joint action.

Alternatively, as has been argued in cluster research (Bathelt, Malmberg, and Maskell 2004), positive effects of related industries may spread through indirect knowledge and labor market effects. This is consistent with the standard microfoundations of agglomeration (sharing, matching, learning model) where learning may be an emergent property of sharing and matching rather than direct collaboration (Duranton and Puga 2004). Boschma and Frenken (2011a), in this vein, point to four potential mechanisms that can be associated with or lead to technological relatedness: producer-user relations, interdependencies in the production system, technological complementarities between industries, and interdependencies in technology development. Thus, even though statistically identified regional related variety may be associated with knowledge spillovers or direct collaborations, the direction of causality is far from clear. To establish such links would require the identification of specific mechanisms and empirical validation in real geographic contexts.

As noted above, the related variety literature treats specialization and diversification in regional development as two sides of a single coin. There is a long tradition of studying regional economic specialization in general, and more recently in the form of the clustering or colocation of a set of densely interrelated activities, whether vertically, horizontally, or indirectly. The clustering of such inter-related firms is usually considered a mechanism that underpins specialization. But the related variety literature relabels it as a form of variety, which does not give us clear guideposts about where the related variety concept is situated 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 broader EEG literature often invokes Jacobs (1969) as an inspiration for the importance of diversification (Frenken, van Oort, and Verburg 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 varied. 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 the 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 defined and measured in the corresponding literature section (see "Measuring Related Variety"), and then demonstrate that there is no clear idea of what a real relatedly diverse industrial structure might look like (see "What Does a Regional Economy with High Related Variety Look Like?).

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,[4] it focuses on intraregional processes and pays little attention to the extraregional geography of spillovers and linkages (Content and Frenken 2016). As Boschma (2017, 357) notes, "the literature on regional diversification has primarily focused on the role of local capabilities … [and] neglected the role of extra-regional linkages and actors that might affect regional diversification." 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 and Iammarino 2017; Bathelt and Buchholz 2019; Li and Bathelt 2021; Yeung 2021). Through these linkages, they can generate knowledge over distance (Bathelt and 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'

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

(and their respective regions') performances. Because of their broad geographic reach, they can participate in localized knowledge spillovers simultaneously in different places (Malecki 2010). Is this because technological closeness is more important than geographic 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.

In sum, this discussion indicates that it cannot be expected, a priori, that a regional economy with high related variety (even if perfectly measured) would create have significantly higher levels of interaction within the region than some other mix of activities. Since such interaction is claimed by related variety research to be one of the channels to superior regional performance, it also cannot be assumed up front that related variety always positively stimulates regional performance or long-term development. Below, we scrutinize some of these links in more detail.

[level1] Measuring Related Variety

A core part of many related variety studies is to operationally define and then measure technological relatedness in the way Frenken, van Oort, and Verburg (2007) suggest. They quantify related and unrelated variety by using entropy measures (Theil 1972; Reardon and Firebaugh 2002) based on the regional industry structure: technological relatedness is measured by simultaneously analyzing industrial classification systems (such as SIC or NAICS) at different levels of granularity. For example, if a region's employment is broadly distributed across many different sectors at a high level of aggregation (e.g., two or three digit), this is considered to be a high level of unrelated variety, since sectors at this level are assumed to be technologically different or unrelated to one another, a reasonable assumption at a first glance in many cases. By contrast, when looking at more finely grained classes of industries (e.g., four, five, or six digit), a diverse set of subsectors within a two- or three-digit class is considered to indicate a high level of technological relatedness. This approach is applied to a region by Frenken, van Oort, and Verburg (2007), who add all main sector scores, weighted by size. Thus, the same index that indicates unrelated variety at a high aggregation level is then viewed as an expression of related variety when applied to more finely 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. In this sense, related variety is based on a statistical artifact: *proximity* within classes of industries but diversity or distance across them. And this artifact is in turn extrapolated from another artifact: the industrial classification system as a whole. In this respect, the related variety measure differs from directly observable categories of economic reality such as employment, wages, incomes, sales, and so on.

While methodologically elegant, a double artifact should be subject to the greatest care and prudence in its use. To start with, industrial classification systems have never been designed with subsectoral 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 the related variety approach to extrapolate relatedness, which is then assumed to shape change, evolution, and dynamism in regional economies as a result of spillovers and other features of being closely related. In an attempt to defend the *principle of relatedness*, Hidalgo et al. (2018) refer to the many different ways activities can be related through a *shared knowledge base*. They assert that it is a general and multiscalar principle of modern economic development through the way that related industries constitute a *product space*. However, 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. As Whittle and Kogler (2020, 101) powerfully state, it "is a rather weak justification to assume that just because two industries share a common two/three-digit classification that this automatically implies [they are] related"

(Essletzbichler 2015; Fitjar and Timmermans 2017) or that "the hierarchical structure of industry classifications reflects the prevalence of scope economies among industries" (Neffke and Henning 2013, 301). Attempts to directly scrutinize the veracity of relatedness measures, such as that by Delgado, Porter, and Stern (2016, 1), find that colocation 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 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 the related variety concept 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 US, 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 can this change be considered evolution into a related variety of photography, or a technological rupture based on the application of new general-purpose digital technologies to transform photographic imaging into something fundamentally different? The fact that related variety is essentially a static concept has also generated criticism from within EEG (Juhász, Broekel, and Boschma 2021; Kuusk and Martynovich 2021).

The logic of technological relatedness does make a certain sense if indeed all subsectors within a main industrial sector are part of the same vertical value chain, that is, making inputs that link to one another, and using real interrelations to do so, as well as indirectly linking by drawing from a common labor pool. To be operationalizable, however, it also would seem to require that all main sectors (e.g., at the two-digit NAICS level) be considered *unrelated* to each other. But this is obviously not the case.

Manufacturing sectors are often vertically linked to producer services or machinery industries, and some sectors have strong input–output relations with others such as computer hardware and software, automobiles and metal fabricating, or chemicals and pharmaceuticals and biotechnology. Conversely, if we use two-digit NAICS codes as a starting point of our analysis, some have more related subsectors than others—and in some main sectors, subgroups are hardly linked to each other.

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 would be. Indeed, while industries at the three-digit level in the classification system are more homogeneous than at the two-digit level, there is still substantial diversity within three-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 pointed out above, the classifications are not designed to capture input–output relations, knowledge commonalities, or overlapping labor demands. 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 multifaceted statistical testing of relationships prior to constructing the statistical artifact of relatedness.[5]

Some of these issues have begun to be addressed. 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

---

[5] We suspect that the related variety index captures a wide array of different forms of relatedness. A particularly potent example of this comes from the findings of Storper et al. (2015) in analyzing high-tech industries in California. At the six-digit NAICS level, the study reveals substantial differences in wages (up to 50 percent) for occupations between two comparable metropolitan regions. This implies that, at least for some industries, even six-digit industry codes may not be very homogeneous in what they are doing *on the ground* across places; at worst, they may be little more than chaotic descriptive aggregations.

downstream linkages benefit from coagglomeration (Duranton and Puga 2005; Mudambi 2008; Timmer, Miroudot, and de Vries 2019; Delgado and Mills 2020). While the majority of related variety studies focus on industrial relatedness, an increasing number of investigations now deploy different indicators. These include *skill relatedness* (Boschma, Minondo, and Navarro 2011; Boschma, Eriksson, and Lindgren 2014; Fitjar and Timmermans 2017), *knowledge variety* (Tavassoli and Carbonara 2014), or *occupational and educational related variety* (Wixe and Andersson 2017). Most promising are studies that measure relatedness through observable interactions and relations between firms such as labor mobility or copatenting (Neffke, Henning, and Boschma 2011; Boschma, Eriksson, and Lindgren 2014; Fiorgo and Mayerhofer 2018). Although a review of this work is beyond the scope of this analysis, such *revealed relatedness* approaches are clearly a step forward in terms of measurement, since the identification of related economic activities is based on real linkages. However, the use of indicators that focus on revealed relatedness also creates new challenges because it is less clear how then to define unrelated variety. Thus far, these studies do not investigate how revealed relatedness would have a stronger positive impact on learning, innovation, and growth capabilities, compared to a less related industry structure. There is little conceptual debate about which of these versions of the related variety concept is most appropriate. 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.

In light of this discussion, it seems reasonable to ask what the concept of relatedness means in terms of real-world linkages and relations (Gong and Hassink 2020), both from a static (input–output interdependence) and dynamic (developmental–innovative–cogrowth) perspective. What we have not been able to identify are systematic discussions or investigations of this in the related variety or other EEG literature, whether in the form of a statistical or *ground truthing* exercise.

[level1] What Does a Regional Economy with High Related Variety Look Like?

Despite some similarities, the notion of related variety aims to go beyond Porter's conception of related and supporting industries in the sense that it attempts to capture the entire regional industry structure in some way and not just a single cluster sector. 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. What would the typical economic geography of a region with high related variety look like?[6] In order to explore this point, Figure 1 attempts to represent the hypothetical industry structure of a regional economy and its degree of related and unrelated variety through different scenarios. It builds on the observation that related variety is associated with regional specialization and clustering and presents the related and unrelated variety scores for ten different scenarios of regional clustering. For reasons of heuristic clarity, assume that the economy portrayed in the figure includes ten sectors that each consist of ten subsectors, and that the ten sectors are completely technologically unrelated (unlikely in reality but consistent with the assumptions of the related variety literature). Assume, in addition, that these sectors establish a perfectly related cluster if each of the corresponding subsectors is equally well developed (for instance in terms of employment). In this exercise, we assume a region with an overall employment of 100,000 people who work within and across a number of clusters: scenario one characterizes a region with one very large cluster, in which all employees are equally distributed across the subsectors of this clustered sector, and scenario ten refers to a situation with ten smaller clusters of equal size, with employment in each cluster sector again being equally distributed

[6] To make things even more complicated, Uhlbach, Balland, and Scherngell (2022) find in a study about the impact of EU research funding on new regional specializations that positive effects seem largest when the level of relatedness is neither too high nor too low. Yet, as in many other studies, it essentially remains unclear what that precisely means.

across all relevant subsectors. The difference between the scenarios in Figure 1 is the number of clusters and their size.[7] Although based on simple assumptions that can be refined in future work, the graph is quite instructive. It reveals that unrelated variety scores are low when there are few clusters and that they increase monotonically with the number of clusters in a regional economy. Related variety scores, in contrast, remain constant regardless of the number of (full-fledged) clusters. This would lead us to conclude, following the related variety argument, that a regional economy with one large cluster is just as good as an economy with two, three or even ten smaller clusters (each fully developed). Should we therefore indeed associate high related variety with a specific sort of cluster structure, no matter how many clusters there are (Bathelt and Zhao 2016)? Or how else can we envision concrete economic geographies of related variety? The trends in Figure 1 may not be surprising when considering the mathematical construct behind related and unrelated variety – but it is not clear conceptually why one large cluster, which represents a highly-specialized economy, would be just as good as ten small clusters, corresponding with maximum diversity.[8] Overall, it seems that related variety scores are high, as long as the cluster sub-sectors have no gaps and are equally represented in an economy.

[Figure 1 about here]

From a policy perspective, this appears problematic, as it is unclear exactly what industry structure the approach defines as optimal and that could become a goal of regional policy. We are unaware of any

[7] The same 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.

[8] In fact, in the scenario with ten clusters, related and unrelated variety scores are the same.

detailed discussions of these issues in the related variety or broader EEG literature itself. The effects of technological relatedness would undoubtedly depend on and vary with the specific industries involved, and not just the number of related complexes and how internally or externally related they are as an artifact of the use of the industrial classification system. Any useful policy implications would therefore need to consider the specific economic context of a region and the reality of industries in the contemporary economy. While some articles pull together and identify empirical regularities in studies on related variety (Boschma and Frenken 2011b; Content and Frenken 2016), it remains unclear what these can tell us if we lack a clear understanding how related variety translates into real geographies and especially those that generate prosperity.

Even if we assume that related variety positively stimulates regional development, its effect will thus likely vary according to the type of region, or interact with other determinants of growth that vary in different combinations with related variety such as whether a regional economy is traditional and trust based or not, the types and levels of long-standing collaboration practices, policy-generated alliances, specific localized skill sets or historic resource-related advantages. Different contexts and institutional set-ups can impact regional development in specific ways and generate different causal relationships between variables that can trigger growth, decline, or stagnation (Storper 1997, 2009; Storper et al. 2015; Glückler and Bathelt 2017; Buchholz 2019; Gong and Hassink 2020). In regression models, such as those commonly used in related variety studies, there are always specific regional cases where regularities between dependent and independent variables cannot be verified or where large positive or negative residuals can be found (Buchholz and Bathelt 2021). Regularities may help understand average trends but they do not do justice to diversity of economies found on the ground. In contrast, a systematic analysis of outliers and regional cases without expected relationships would help identify and isolate additional influences and alternative explanations to the related variety argument.

We are not aware, however, of broad systematic analyses of this kind in the literature.[9] Such investigations could also strengthen the basis for relatedness policies in relation to other influences (Iammarino, Rodríguez-Pose, and Storper 2019). This effort would require fine-grained and case-based geographic analyses of how the results of related variety studies play out in real space.

Related Variety, City Scale, and Regional Economic Structure

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 as unrelated variety, ln(population) and population density, as well as with employment shares in two-digit NAICS industries. The results for metropolitan statistical areas (MSAs) in the US, using data from the Quarterly Census of Employment and Wages (US Bureau of Labor Statistics 2020) and the five-year American Community Survey (downloaded from the National Historical Geographic Information System—Manson et al. 2021), for the year 2017 are shown in Figure 2 and in the online appendix. The scatterplots suggest some surprising regularities and potential problems 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.[10] From the definition of the two indicators, this close relationship is not obvious or may even

[9] As a starting point, however, see Fitjar and Timmermans (2017).

[10] Castaldi, Frenken, and Los (2015) also find a positive correlation between unrelated and related variety in their study at the US state level but are neither concerned about this relationship nor investigate it further.

be counterintuitive, but a possible explanation can be found considering scale. Both related and unrelated variety are strongly associated with city size as measured by ln(population). More specifically, we find that with growing city size, related and unrelated variety (adopting Frenken, van Oort, and Verburg's (2007) operationalization of these notions for the purpose of this test) are systematically increasing among US MSAs. The concept unintentionally associates large cities with high related variety and thus implies that the secret to favorable regional development is to be big, since this generates higher related variety. 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, there are also more well-developed subsectors within each main sector (Macheras and Stanley 2017). According to the entropy measure by Frenken, van Oort, and Verburg (2007), this will result in a high degree of technological relatedness in a large city. Small cities are less likely to have such an industry structure. Fully fledged subsectors 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 subsectoral 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 growing city size. As such, this creates bias toward larger cities and makes it difficult to derive lessons about regional development dynamics, other than saying that large cities have stronger development potential than small cities.

In a complex urban system, such as that of the US with its many cities of different sizes, related variety thus appears to operate as a proxy of city scale.[11] While this is potentially worrisome, it is interesting to note that the studies in Table 1, except for four,[12] 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 proxy urbanization economies, in their model formulations. No study in the table questions the use of this variable. Figure 2 shows that the association of population density with related variety is much weaker than that with ln(population). With an R-squared value of 0.40 between population density and ln(population), population density is clearly not a strong control for the scale issues raised above. 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 scale of the cluster itself, and how this would affect the ability of each related activity (or firm, for that matter) within the cluster to exploit scale efficiencies (Kemeny and Storper 2015).

It is notoriously hard to identify general principles that link regional economic structure and regional economic performance. This is also illustrated in the online appendix, which shows, for all US MSAs, the relationships between employment shares in two-digit NAICS industries and related variety scores. The figure implies that some aggregate sectors have a negative association with related variety when they dominate a large part of overall employment, which is more likely to be the case in smaller regions. When a main sector has a very high employment share beyond 20 or 30 percent, related

---

[11] This may also explain why recent studies use indicators, such as related variety density (consisting of a quotient of different related/unrelated variety indicators—e.g., Balland et al. 2019), that reduce the impact of scale on the overall relatedness measure computed. While this seems an improvement, there are also concerns with indicators, such as related variety density (Uhlbach, Balland, and Scherngell 2022) or a combination of relatedness and complexity variables (Deegan, Broekel, and Fitjar 2021), since these are even more difficult to make sense of in real geography or policy terms than conventional related variety.

[12] See Boschma, Eriksson, and Lindgren (2009), Cainelli and Iacobucci (2012), Ebersberger, Herstad, and Koller (2014), and Tavassoli and Jienwatcharamongkhol (2016).

variety scores generally decrease—implying 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 in US MSAs are associated with decreasing related variety. The related variety logic 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 may not always be the case.

Storper et al. (2015) note that the San Francisco Bay Area showed a fourfold increase in the direct share of information technology–based employment from 1970 to 2010, which has not been associated with a negative lock-in effect, but rather with extraordinary technological learning and dynamic spin-offs, development of new products, and other types of diversification. The related variety literature would likely conclude that that is because of the general principle of the concentration of a wide variety of related subsectors; but it could equally well be argued that it is because the Bay Area is concentrated in a specific and technologically dynamic cluster of the economy. Contrast this to the fate of Rochester, formerly concentrated in physical photo-optics, as we discussed earlier. It echoes the point made in a classic article by Chinitz (1961) that what counts is not just whether a region is specialized or clustered, but rather the nature of the specialization itself. Using a similar 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. These examples illustrate that different regional contexts and industry configurations may have a completely different impact. To study this, it could be interesting to compare different regions with low related variety associated with a dominant single cluster and, at the same time, positive development outcomes, as, for example, the Bay Area's dominant cluster in information technology and Detroit's dominant cluster in mechanical engineering. It would be equally interesting to investigate the reverse situation. Only through such comparative analyses 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 and searches for answers.

The online appendix also shows that those sectors, which 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), there is no clear relationship between these industry groups via the concept of technological relatedness and knowledge spillovers and development triggers. These sectors are, however, typically more developed in larger cities, which is consistent with our prior observation regarding scale.

Short- and Long-Run Effects of Related Variety

In empirical studies 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. This is typically done in large-*N* regression models with multiple independent variables. As in some regional economic studies (e.g., Glaeser et al. 1992), such modeling uses a regional development indicator as the dependent variable, with related variety and other variables as independents. In an essentially cross-sectoral setting, the dependent variable is typically measured over a time interval of a few years, whereas the independents are measured in a single year at the beginning of the interval. This is also the approach of Frenken, van Oort, and Verburg (2007) in their study of regional income and employment growth in Dutch regions.[13]

[13] While their analysis draws from a small number of regions, other studies use a similar approach with much larger sample sizes, which improves reliability.

Upon closer examination, about half of the studies in Table 1 use a similar cross-sectional setting to predict regional development over time based on related variety at the beginning of that time period. The other half conduct a panel analysis or incorporate some other dynamic approach. Following the methodology by Frenken, van Oort, and Verburg (2007), Buchholz and Bathelt (2021) use related and unrelated variety measures in 2010 to explain changes in income and employment levels in US MSAs between 2010 and 2017. They largely reproduce the findings of Frenken, van Oort, and Verburg (2007) for the US urban system and find 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, since growth and change in the US urban system is mostly divided between those urban regions with high income growth versus those with high population growth (Kemeny and Storper 2012).

In addition, there are some inherent limitations when attempting to explain *change* in a dependent variable through independent variables that measure some *state* in the past. It would be more intuitive to conduct a panel analysis and investigate instead what kind of *changes* in related variety go along with *changes* in the dependent variable. To test this, Buchholz and Bathelt (2021) conduct a panel regression analysis for US MSAs with the same data as above. According to the panel results, related variety no longer has a significant regression coefficient, and unrelated variety seems to be negatively associated with income changes albeit at a low significance level (Table 2).

[Table 2 about here]

While other panel analyses have come to different conclusions regarding the effects of related/unrelated variety, we should not be overly surprised that the results of cross-sectional studies cannot automatically be transferred to a panel setting. In a short time frame of few years, changes in the related and unrelated variety structure of US regions are small, and scores only change incrementally (Buchholz and Bathelt 2021). In many cases, shifts in the local industry structure are not large enough to explain changes in regional development. To be fair, findings of related variety studies are not all mutually consistent in this respect, and some report different results. For instance, Castaldi, Frenken, and Los (2015) find in a study of US states over a longer time span from 1977 to 1999 that related variety has a positive and significant impact on patenting activity. Overall, about as many studies in Table 1 report a significantly positive impact of related variety on regional growth as studies that find insignificant or negative impacts. This suggests the need for deeper probing of causality (Rutten 2020). Can regional development be considered the consequence of a favorable industry structure if this industry structure does not change or remains stable in the short time span observed? In this regard, should related variety be considered the cause of positive regional performance or is it rather a reflection of it (Martin and Sunley 2022)?[14]

Concerns also arise when considering the effects of related variety in the long run. The literature gives little consideration 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: (1) some radical technologies replacing previous user technologies (e.g., digital photography replacing photo optics, film, and

[14] An investigation by Spencer et al. (2010) asks similar questions but with respect to the impact of industrial clusters on regional performance. In a study of Canadian city-regions in the early 2000s, they show that city-regions with a higher employment share in clusters have a higher average income, employment growth, and patenting intensity, although they find large variations. As with the related variety concept, however, the direction of causality remains opaque. Do clusters cause such development or are they attracted to high-income regions that have higher skill levels?

mechanical cameras); (2) some radical technologies having new types of user complementarities (fields of users not previously linked such as when machines become digitally guided as with the emerging self-driving cars); and (3) some radical technologies making possible completely new types of activity (the telephone making possible remote hearing of voices) (Petralia 2020). Related variety is not likely to explain these cases and—at least in some of them—crucially positive effects on regions hosting the breakthrough process could not be detected by the methodological 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, Dallas, 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. Though there were communications equipment industries in the Bay Area, the evolution into semiconductor dominance was initially far from evident from the Bay Area's overall industrial structure at the time (Saxenian 1985, 1994; Scott and Storper 1987), which was dominated by natural resource processing industries, branch plant manufacturing, refining, and an important port complex. Moreover, several places with heavy concentrations in chip design and manufacture in the 1950s and 1960s, such as Phoenix, Dallas, Los Angeles, and New Jersey, subsequently fell behind the Bay Area (Storper et al. 2015).

The above discussion shows that while the related variety approach aims to explain differences in regional economic development patterns, the actual effects of technological relatedness are far from clear—neither in the short nor the long term.

Conclusion: Related Variety, Place, and Geography

As we have noted throughout this article, most of the literature on related variety seeks to identify macroregularities or mechanisms that are said to shape development of larger populations of cities or regions. This is certainly an important step toward scientific results in economic geography, but more remains to be done. 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 patterns. The studies in Table 1 either do not discuss geographic variation at all, briefly present variable distributions over space, or use regional dummy variables. Only one of the thirty studies links to a specific regional case study. Neffke, Henning, and Boschma (2011) mention the case of Linköping in their analysis of the impact of reveled relatedness on industrial transformation. However, this case is declared as arbitrary and remains descriptive. None of the articles 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. Economic 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 natural geography, migration and workforce change, regional land use and housing dynamics, education provision, the influence of racism, segregation and class relations, as well as many other dimensions. It is thus awkward that many related variety 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 from deeper engagement with the wider social science of economic development in which the rate and direction of technological change is seen from this multiplicity of angles (e.g., Mokyr 1990).

To be clear, the purpose of this article is neither to give a complete overview of the 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 (1) the geographic expressions of this approach and to better understand its regularities, (2) grasping how related variety indicators vary in real economies and what mechanisms are at play, (3) identifying limitations and biases of the relatedness measure, and (4) investigating 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 city, region, and country cases. While we focus in our analysis on the most-cited related variety studies, many of our remarks are equally relevant for recently published work in this field.

From our analysis, a number of suggestions emerge that address the points raised here, for work on related variety in particular and regional economic evolution in general. First, we have argued for scrutinizing results in a variety of ways. For example, it would be helpful to confront general results on related variety with ground-truthed results on real regional economies over time, probing examples deeply. This could come in the form of well-chosen case studies, and notably those that examine what might appear to be confounding cases, such as economic success with radical change, or at least sampling from economies across the distribution of results rather than merely reporting on means. In further probing broad statistical results, we advocate using systematic control variables and testing alternative explanations from the literature on economic growth and development. Whether in case studies or further empirical work, institutional and contextual influences on outcomes should be considered and rigorously operationalized as part of the set-up for additional or alternative explanations. Regional economies should be viewed and compared over time, in real panel data or—with case studies—considering cases over time in a rigorous and comparative way. A second major area of work, where already experiments are taking place, could scrutinize how well operationalizations of related variety using industrial censuses correspond with real relatedness, considering alternative indexes of relatedness (such as knowledge bases), but insuring that in all cases, the work is not purely based on alternative assumptions but on ground-truthed notions of relatedness. Clear ideas about evolution without related variety or, at least, what would constitute growth based on new pathways or branching into areas that are not strongly related to the past, should be operationalized and considered as alternative explanations. A third area of work would more systematically consider the role of system-wide or extraregional connections and forces (such as linkaging and sorting) on the evolution of relatedness but also on the outcome variable of development.

Taking up these challenges could propel research on related variety back into the mainstream of important debates about regional growth and development and overcome its tendency to separate from such work. This would enrich efforts in cognate areas in regional economics, innovation studies, and development studies and vice-versa. It would also go a long way toward overcoming the perception that the related variety approach is technologically deterministic or better defend it in the wider world of ideas.

References

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

Aarstad, J., Kvitastein, O. A., and Jakobsen, S.-E. 2016. Related and unrelated variety as regional drivers of enterprise productivity and innovation: A multilevel study. *Research Policy* 45 (4): 844–56.

Antonietti, R., and 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., and Reganati, F. 2017. Do agglomeration externalities affect firm survival? *Regional Studies* 51 (4): 548–62.

Boschma, R., Eriksson, R., and Lindgren, U. 2009. How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity. *Journal of Economic Geography* 9 (2): 169–90.

Boschma, R., Eriksson, R. H., and Lindgren, U. 2014. Labour market externalities and regional growth in Sweden: The importance of labour mobility between skill-related industries. *Regional Studies* 48 (10): 1669–90.

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

Boschma, R., Minondo, A., and Navarro, M. 2011. Related variety and regional growth in Spain. *Papers in Regional Science* 91 (2): 241–56.

Cainelli, G., and Iacobucci, D. 2012. Agglomeration, related variety, and vertical integration. *Economic Geography* 88 (3): 255–77.

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

Castaldi, C., Frenken, K., and Los, B. 2015. Related variety, unrelated variety and technological breakthroughs: An analysis of US state-level patenting. *Regional Studies* 49 (5): 767–81.

Content, J., Frenken, K., and Jordaan, J. A. 2019. Does related variety foster regional entrepreneurship? Evidence from European regions. *Regional Studies* 53 (11): 1531–43.

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

Ebersberger, B., Herstad, S. J., and Koller, C. 2014. Does the composition of regional knowledge bases influence extra-regional collaboration for innovation? *Applied Economics Letters* 21 (3): 201–4.

Eriksson, R. H. 2011. Localized spillovers and knowledge flows: How does proximity influence the performance of plants? *Economic Geography* 87 (2): 127–52.

Firgo, M., and Mayerhofer, P. 2018. (Un)related variety and employment growth at the sub-regional level. *Papers in Regional Science* 97 (3): 519–47.

Frenken, K., van Oort, F., and Verburg, T. 2007. Related variety, unrelated variety and regional economic growth. *Regional Studies* 41 (5): 685–97.

Fritsch, M., and Kublina, S. 2018. Related variety, unrelated variety and regional growth: The role of absorptive capacity and entrepreneurship. *Regional Studies* 52 (10): 1360–71.

Guo, Q., He, C., and Li, D. 2016. Entrepreneurship in China: The role of localisation and urbanisation economies. *Urban Studies* 53 (12): 2584–606.

Hartog, M., Boschma, R., and Sotarauta, M. 2012. The impact of related variety on regional employment growth in Finland 1993–2006: High-tech versus medium/low-tech. *Industry and Innovation* 19 (6): 459–76.

Howell, A., He, C., Yang, R., and 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., and Capone, F. 2017. The impact of related variety on the creative employment growth. *Annals of Regional Science* 58 (3): 491–512.

Liang, J., and Goetz, S. J. 2018. Technology intensity and agglomeration economies. *Research Policy* 47 (10): 1990–95.

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

Neffke, F., Henning, M., and Boschma, R. 2011. How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. *Economic Geography* 87 (3): 237–65.

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

Sedita, S. R., De Noni, I., and Pilotti, L. 2017. Out of the crisis: An empirical investigation of place-specific determinants of economic resilience. *European Planning Studies* 25 (2): 155–80.

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

Tavassoli, S., and Jienwatcharamongkhol, V. 2016. Survival of entrepreneurial firms: The role of agglomeration externalities. *Entrepreneurship and Regional Development* 28 (9 10): 746–67.

van Oort, F., de Geus, S., and Dogaru, T. 2015. Related variety and regional economic growth in a cross-section of European urban regions. *European Planning Studies* 23 (6): 1110–27.

Wixe, S., and Andersson, M. 2017. Which types of relatedness matter in regional growth? Industry, occupation and education. *Regional Studies* 51 (4): 523–36.

(ii) Other References Cited

Audretsch, D. B., and Feldman, M. P. 1996. R&D spillovers and the geography of innovation and production. *American Economic Review* 86 (3): 630–40.

Balland, P. A., Boschma, R., Crespo, J., and Rigby, D. L. 2019. Smart specialization policy in the European Union: Relatedness, knowledge complexity and regional diversification. *Regional Studies* 53 (9): 1252–68.

Bathelt, H., and Buchholz, M. 2019. Outward foreign-direct investments as a catalyst of urban-regional income development? Evidence from the United States. *Economic Geography* 95 (5): 442–66.

Bathelt, H., and Glückler, J. 2000. Netzwerke, Lernen und evolutionäre Regionalentwicklung [Networks, learning and evolutionary regional development]. *ZFW [German Journal of Economic Geography]* 44 (3–4): 167–82.

Bathelt, H., and Henn, S. 2014. The geographies of knowledge transfers over distance: Toward a typology. *Environment and Planning A* 46 (6): 1403–24.

Bathelt, H., and Zhao, J. 2016. Conceptualizing multiple clusters in mega-city regions: The case of the biomedical industry in Beijing. *Geoforum* 75 (October): 186–98.

Bathelt, H., Malmberg, A., and 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): 351–64.

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

______. 2011a. Technological relatedness and regional branching. In *Beyond territory: Dynamic geographies of knowledge generation, diffusion, and innovation*, ed. H. Bathelt, M. P. Feldman, and D. F. Kogler, 64–81. Oxford: Oxford University Press.

______. 2011b. The emerging empirics of evolutionary economic geography. *Journal of Economic Geography* 11 (2): 295–307.

Boschma, R. A., and Lambooy, J. G. 1999. Evolutionary economics and economic geography. *Journal of Evolutionary Economics* 9 (4): 411–29.

Boschma, R. A., and Martin, R., eds. 2010a. *Handbook of evolutionary economic geography*. Cheltenham, UK: Edward Elgar.

______. 2010b. The aims and scope of evolutionary economic geography. In *Handbook of evolutionary economic geography*, ed. R. A. Boschma, and R. Martin, R., 3–39. Cheltenham, UK: Edward Elgar.

Breul, M., Broekel, T., and Brachert, M. 2015. 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): 133–50.

Buchholz, M. 2019. Organizations, institutions and networks in local scenes: The growth of San Francisco Bay Area punk rock. *Geoforum* 103:158–66.

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

Cantwell, J. A. 1989. *Technological innovation and multinational corporations*. Oxford: Basil Blackwell.

Coenen, L., Asheim, B., Bugge, M. M., and Herstad, S. J. 2017. Advancing regional innovation systems: What does evolutionary economic geography bring to the policy table? *Environment and Planning C: Politics and Space* 35 (4): 600–20.

Chinitz, B. 1961. Contrasts in agglomerations: New York and Pittsburgh. *American Economic Review* 51 (2): 279–89.

Content, J., and Frenken, K. 2016. Related variety and economic development: A literature review. *European Planning Studies* 24 (12): 2097–112.

Cooke, P., Uranga, M. G., and Extebarria, G. 1997. Regional innovation systems: Institutional and organizational dimensions. *Research Policy* 26 (4–5): 475–91.

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

Deegan, J., Broekel, T., and 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., and Mills, K. G. 2020. The supply chain economy: A new industry categorization for understanding innovation in services. *Research Policy* 49:104039.

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

Duranton, G., and Puga, D. 2004. Microfoundations of urban agglomeration economies. In *Handbook of regional and urban economics, vol. 4*, ed. J. V. Henderson, and J.-F. Thisse, 2063–117. Amsterdam: Elsevier.

_____. 2005. From sectoral to functional urban specialization. *Journal of Urban Economics* 57 (2): 343–70.

Essletzbichler, J. 2015. Relatedness, industrial branching and technological cohesion in US metropolitan areas. *Regional Studies* 49 (5): 752–66.

Essletzbichler, J., and Rigby, D. 2007. Exploring evolutionary economic geographies. *Journal of Economic Geography* 7 (5): 549–71.

Fitjar, R. D., and Timmermans, B. 2017. Regional skill relatedness: Towards a new measure of regional related diversification. *European Planning Studies* 25 (3): 516–38.

Florida, R. L. 2002. *The rise of the creative class: And how it's transforming work, leisure, community and everyday life*. New York: Basic Books.

_____. 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: Basic Books.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. 1992. Growth in cities. *Journal of Political Economy* 100 (6): 1126–52.

Glückler, J., and Bathelt, H. 2017. Institutional context and innovation. In *The Elgar companion to innovation and knowledge creation,* ed. H. Bathelt, P. Cohendet, S. Henn, and L. Simon, 121–37. Cheltenham, UK: Edward Elgar.

Gong, H., and Hassink, R. 2020. Context sensitivity and economic-geographic (re)theorising. *Cambridge Journal of Regions, Economy and Society* 13 (3): 475–90.

Hassink R., Isaksen A., and Trippl, M. 2019. Towards a comprehensive understanding of new regional industrial path development. *Regional Studies* 53 (11): 1636–45.

Henning, M. 2019. Time should tell (more): Evolutionary economic geography and the challenge of history. *Regional Studies* 53 (4): 602–13.

Hidalgo, C. A., Balland, P. A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Glaeser, E., et al. 2018. The principle of relatedness. In *unifying themes in complex systems IX: Proceedings of the ninth international conference on complex systems,* ed. A. J. Morales, C. Gershenson, D. Brah, A. A. Minai, and Y. Bar-Yam, 451–7. Cham, Switzerland: Springer VS.

Iammarino, S., Rodríguez-Pose, A., and Storper, M. 2019. Regional inequality in Europe: Evidence, theory and policy implications. *Journal of Economic Geography* 19 (2): 273–98.

Jacobs, J. 1969. *The economy of cities*. New York: Random House.

Juhász, S., Broekel, T., and 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., and Storper, M. 2012. The sources of urban development: Wages, housing and amenity gaps across American cities. *Journal of Regional Science* 52 (1): 85–108.

_____. 2015. Is specialization good for regional economic development? *Regional Studies* 49 (6): 1003–18.

Ketels, C. 2016. Smart specialisation: Opportunities and challenges for regional innovation policy. *Competitiveness Review* 26 (4): 476–78.

Kuusk, K., and 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., and Bathelt, H. 2021. Spatial knowledge strategies: An analysis of international investments using fuzzy set Qualitative Comparative Analysis (fsQCA). *Economic Geography* 97 (4): 366–89.

Macheras, A. B., and Stanley, M. 2017. Diversification and specialization across urban areas. *Economic Focus Q4/2017.* Federal Reserve Bank of Richmond. https://www.richmondfed.org/publications/research/econ_focus/2017/q4/district_digest.

MacKinnon, D., Dawley, S., Pike, A., and Cumbers, A. 2019. Rethinking path creation: A geographical political economy approach. *Economic Geography* 95 (2): 113–35.

Malecki, E. J. 2010. Global knowledge and creativity: New challenges for firms and regions. *Regional Studies* 44 (8): 1033–52.

Malmberg, A., and Maskell, P. 2002. The elusive concept of localization economies: Towards a knowledge-based theory of spatial clustering. *Environment and Planning A* 34 (3): 429–49.

Manson, S., Schroeder, J., Van Riper, D., Kugler, T., and Ruggles, S. 2021. *IPUMS National Historical Geographic Information System: Version 16.0* [Data Set]. Minneapolis, MN. https://www.nhgis.org/.

Marshall, A. 1920. *Principles of economics*. 8th ed. London: Macmillan.

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

_____. 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*. London: Oxford University Press.

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

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

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: Free Press.

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

Rutten, R. 2020. Comparing causal logics: A configurational analysis of proximities. *ZFW—Advances in Economic Geography* 64 (3): 134–48.

Saxenian, A. L. 1985. The genesis of Silicon Valley. In *Silicon landscapes* ed. P. Hall and A. Markusen, 20–34. Boston: Allen and Unwin.

_____. 1994. *Regional advantage: Culture and competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.

Scott, A. J., and Storper, M. 1987. High technology industry and regional development: A theoretical critique and reconstruction. *International Social Science Journal* 39 (2): 215–32.

Spencer, G. M., Vinodrai, T., Gertler, M. S., and 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: Guilford Press.

_____. 2009. Regional context and global trade. *Economic Geography* 85 (1): 1–21.

Storper, M., and Walker, R. 1989. *The capitalist imperative: Territory, technology, and industrial growth*. New York: Basil Blackwell.

Storper, M., Kemeny, T., Makarem, N., and Osman, T. 2015. *The rise and fall of urban economies: Lessons from San Francisco and Los Angeles*. Stanford, CA: Stanford University Press.

Theil, H. 1972. *Statistical decomposition analysis*. Amsterdam: North-Holland Publishing.

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

Uhlbach, W.- H., Balland, P. A., and Scherngell, T. 2022. Public R&D funding and new regional specialisations: The contingent role of technological relatedness. *Industry and Innovation* 29 (4): 511–32.

US Bureau of Economic Analysis. 2019. *Regional Economic Accounts* [Data Set]. Washington, DC: US Bureau of Economic Analysis. https://apps.bea.gov/regional/downloadzip.cfm.

US Bureau of Labor Statistics. 2020. *Quarterly Census of Employment and Wages—QCEW. QCEW Data Files, 1993–2017* [Data Set]. Washington, DC: US Bureau of Labor Statistics. https://www.bls.gov/cew/data.htm.

_____. 2021. *Quarterly Census of Employment and Wages—QCEW. QCEW Data Files, 2017* [Data Set]. Washington, DC: US Bureau of Labor Statistics. https://www.bls.gov/cew/data.htm.

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

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.

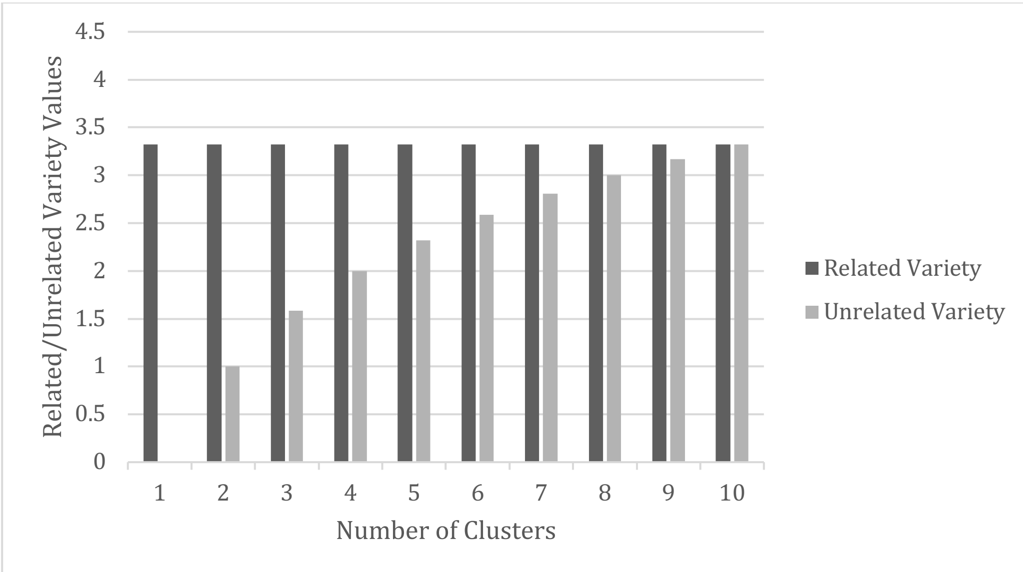

Figure 1. Related and unrelated variety values by regional cluster structure (scenarios).

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

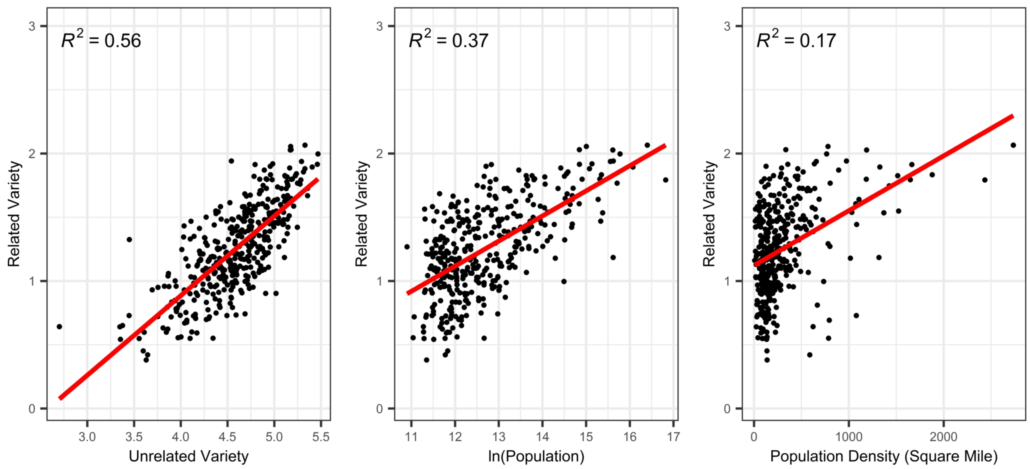

Figure 2. Scatterplots of related variety over unrelated variety, ln(population) and population density for US MSAs, 2017.

Note: Unrelated variety was computed based on three-digit, related variety based on five-digit NAICS codes.

Source: Manson et al. (2021); US Bureau of Labor Statistics (2021).

READTable 1

Characteristics of the Top Thirty Most-cited Articles on Related Variety According to Scopus Citations (November 29, 2021)

Scopus Citations	Authors (Year)	Number of Observations	Dependent Variable(s)	Regional Scale/Size Included?	Panel or Cross-section	Main Findings	Specific Geographic Analysis?	Analysis of Underlying, Mechanisms, Institutions, Alternative Explanations?
1309	Frenken, van Oort, and Verburg (2007)	40 Dutch NUTS 3 regions	(1) Employment growth (1996–2002); (2) productivity growth (1996–2001)	No, but population density(*)	Cross-section	(1) Related variety has positive, highly significant impact on employment growth; unrelated variety negative and insignificant; (2) related variety has negative, significant impact on productivity growth; unrelated variety negative and insignificant	No	No
585	Neffke, Henning, and Boschma (2011)	Plant-level data of 70 Swedish regions; 72,100 membership observations (industry-region combinations)	(1) Membership probability (industry stays in region); (2) entry probability (entry in 5 years); (3) exit probability (exit in 5 years) (1969, 1974, 1979, 1984, 1989, 1994)	No	Equivalent to panel	(1) Regional closeness (based on revealed relatedness) has positive, highly significant impact on membership; extraregional closeness negative and highly significant; (2) regional closeness has positive, highly significant impact on entry; extraregional closeness negative and highly significant; (3) regional closeness has negative, highly significant impact on exit; extraregional closeness negative and highly significant	Largely descriptive, *arbitrary case study* of Linköpings revealed relatedness in industrial transition	No
525	Boschma, and Iammarino (2009)	103 Italian NUTS 3 regions	(1) Employment growth (1995–2003); (2) value-added growth (1995–2003); (3) labor productivity growth (1995–2003)	No, but population density	Cross-section	(1) Related and unrelated variety of exports have positive but insignificant impact on employment growth; (2) related variety has positive, highly significant impact on value-added growth; unrelated variety positive and insignificant; (3) related variety has positive, significant impact on labor-productivity growth; unrelated variety positive and insignificant	Macroregional control variables included, but no specific discussion	No
209	Boschma, Eriksson, and Lindgren (2009)	17,098 job moves to plants in Swedish regions	Plant-level labor productivity growth (2001–3)	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 geographic study	No, but firm-level controls
195	Boschma, Minondo, and 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, and Los (2015)	51 US states over 22 years (877 observations)	(1) Number of patents; (2) number of superstar patents (1977–99)	No	Panel	(1) Related variety has positive, significant effect on patents; unrelated variety negative and insignificant; (2) 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 geographic analysis	No
125	Saviotti, and Frenken (2008)	20 OECD countries over eight 5-year periods (156 observations)	(1) gross domestic product (GDP) per capita growth; (2) labor productivity growth (in eight 5-year periods 1964-2003)	No	Equivalent to panel	(1) Related variety of exports has positive, highly significant impact on GDP per capita growth; unrelated variety negative and highly significant; (2) 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 no specific analysis	No
87	Eriksson (2011)	8,313 plants in Swedish economy (located in differently sized regions)	Plant labor productivity (2001–3)	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 geographic analysis	No; but some controls
84	Hartog, Boschma, and 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	No	No

						different industries, related variety among high-tech sectors has positive, significant impact		
75	Aarstad, Kvitastein, and Jakobsen (2016)	6,584 enterprises in 89 economic-geographic regions in Norway	(1) Enterprise productivity (2010); (2) innovation occurrence (2008–10)	No, but population density	Cross-section	(1) Related variety has positive but insignificant impact on enterprise productivity; unrelated variety negative and highly significant; (2) related variety has positive, significant impact on enterprise innovation; unrelated variety negative and insignificant	No, essentially not a geographic study	No, but firm-level controls
68	van Oort, de Geus, and Dogaru (2015)	205 NUTS 2 regions in 15 EU countries	(1) Employment growth; (2) labor productivity per employee; (3) unemployment growth (2000–10)	No, but population density	Cross-section	(1) Related variety has positive, significant impact on employment growth; unrelated variety also positive and significant; (2) related variety has negative, insignificant impact on productivity growth; unrelated variety positive and insignificant; (3) 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 geographic analysis	No, but numerous controls
68	Boschma, Eriksson, and Lindgren (2014)	72 Swedish functional regions over 5 years (360 observations)	(1) Productivity growth; (2) employment growth; (3) unemployment growth (1998-2002)	No, but population density	Panel	(1) Related and unrelated variety of labor market flows have positive but insignificant impacts on productivity growth; (2) related variety has positive, significant impact on employment growth; unrelated variety negative and significant; (3) related and unrelated variety have negative but insignificant impacts on unemployment growth	No	No, but some controls
67	Tavassoli and Carbonara (2014)	81 Swedish functional regions over 6 years (486 observations)	Number of patent applications per year (2002–7)	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 geographic analysis	No, but numerous controls
61	Cainelli, and 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 macroregions, but no specific discussion	No, but many industry dummies and other controls
60	Antonietti, and Cainelli (2009)	715 Italian manufacturing firms in 103 Italian provinces	(1) research and development (R&D) investment per employee (2003); (2) firm propensity to innovate (2001-3); (3) total factor productivity (2003); (4) firm propensity to export (2001–3)	No, but population density	Cross-section	(1) Related variety has positive, highly significant impact on R&D investment per employee; unrelated variety negative and highly significant; (2) related variety has negative but insignificant impact on firm propensity to innovate; unrelated variety positive and insignificant; (3) related variety has negative but insignificant impact on total factor productivity; unrelated variety positive and insignificant; (4) 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, and 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, and 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 (2001-7)	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 and Moreno (2018)	255 European NUTS 2 regions over 9 years (2,219 observations)	(1) Patents per capita; (2) patent quality: patents weighted by citations (1999-2007)	No	Panel	(1) Related variety has positive, highly significant impact on patenting; unrelated variety negative but insignificant; (2) 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, and Pilotti (2017)	686 Italian local labor systems	Growth in employment rate (2009-13); 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	Macroregional control variables and industrial district dummy included;	No

						unrelated variety and knowledge base variable (share of corresponding industries) insignificant	some description of spatial variations of variables; no specific geographic analysis	
36	Cortinovis, and van Oort (2015)	260 European NUTS 2 regions over 9 years (2340 observations)	(1) Employment growth; (2) unemployment growth; (3) gross value-added per hour (productivity) growth (2004-12)	No, but population density	Panel	(1) Related variety has negative, highly significant impact on employment growth; unrelated variety negative but insignificant; (2) related variety has positive but insignificant impact on unemployment growth; unrelated variety also positive and insignificant; (3) 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 geographic analysis	No
36	Wixe and Andersson (2017)	290 Swedish municipalities	(1) Employment growth; (2) productivity growth (2002-7)	No, but population density	Cross-section	(1a) Related industry variety has positive, highly significant impact on employment growth; unrelated industry variety negative but insignificant; (1b) related educational variety has negative but insignificant impact on employment growth; unrelated educational variety positive but insignificant; (1c) related and unrelated occupational variety are both insignificant; (2a) related industry variety has negative, highly significant impact on productivity growth; unrelated industry variety negative but insignificant; (2b) related educational variety has positive, highly significant impact on productivity growth; unrelated educational variety positive and significant; (2c) related and unrelated occupational variety are both insignificant	No	No; but separate models for manufacturing and services
30	Basile, Pittiglio, and 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 geographic analysis	No, but many industry and other controls used
27	Fritsch, and 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	Macroregional control variables; some description of spatial variations of variables; no specific geographic analysis	No, some controls
27	Howell, He, Yang, and Cindy (2018)	135,000 Chinese new manufacturing firms in 333 prefecture-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 geographic study	No; firm-level controls and industry dummies used
23	Liang and Goetz (2018)	3,147 U.S. counties	Employment growth (2003-13)	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 X technology intensity positive and highly significant	Some description of spatial variations of variables; no specific geographic analysis	No
23	Tavassoli and 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, and Mayerhofer (2018)	81 Austrian labor market districts in two periods (162 observations)	Employment growth (2000–2006; 2007–13)	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 geographic explanation	No, but numerous controls
19	Ebersberger, Herstad, and Koller (2014)	34,892 region-technology combinations in European NUTS 3 regions over six 5-	(1) Interregional domestic collaboration in patenting; (2) international collaboration (over six	Yes	Equivalent to panel	(1) Related technological variety has negative, highly significant impact on interregional collaboration; (2) it has positive, highly significant impact on international collaboration	No	No

		year periods (209,352 observations)	5-year intervals 1980-2010)					
18	Content, Frenken, and Jordaan (2019)	204 European NUTS 2 regions	(1) Share of regional entrepreneurs; (2) share of necessity-driven entrepreneurs; (3) share of opportunity-driven entrepreneurs (2007-14)	No, but population density	Cross-section	(1) Related variety has positive but insignificant impact on entrepreneurship; unrelated variety negative and highly significant; (2) similar results for necessity-driven entrepreneurship; (3) 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 geographic analysis	No
17	Lazzeretti, Innocenti, and Capone (2017)	103 Italian provinces	Employment difference in cultural/creative industries (1991–2001; 2001–11; 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 2001–11 essentially support these findings	Macroregional control variables; some description of spatial variations of variables; spatial lag and error models; no specific geographic 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 US MSAs, 2010—17, Cross-sectional and Longitudinal Effects

Dependent Variables / Independent Variables	ln(Per-Capita Income) Growth 2010–17 (1)	ln(Employment) Growth 2010–17 (2)	ln(Per-Capita Income) Change 2010–17 (3)	ln(Employment) Change 2010–17 (4)
Intercept	-0.004 (0.005)	-0.016 (0.010)	0.034*** (0.003)	0.099*** (0.009)
Related variety 2010	-0.001 (0.001)	0.004** (0.002)		
Unrelated variety 2010	0.003*** (0.001)	0.000 (0.003)		
ln(Employment density) 2010	0.000 (0.000)	-0.001** (0.001)		
Related variety change 2010–17			0.003 (0.010)	-0.022 (0.036)
Unrelated variety change 2010–17			-0.029* (0.016)	0.108 (0.068)
ln(Employment density) change 2010–17			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 two-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 2010–17 period; columns 3 and 4 refer 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 and Bathelt (2021, 32) based on data from US Bureau of Economic Analysis (2019); US Bureau of Labor Statistics (2020).

Online Appendix

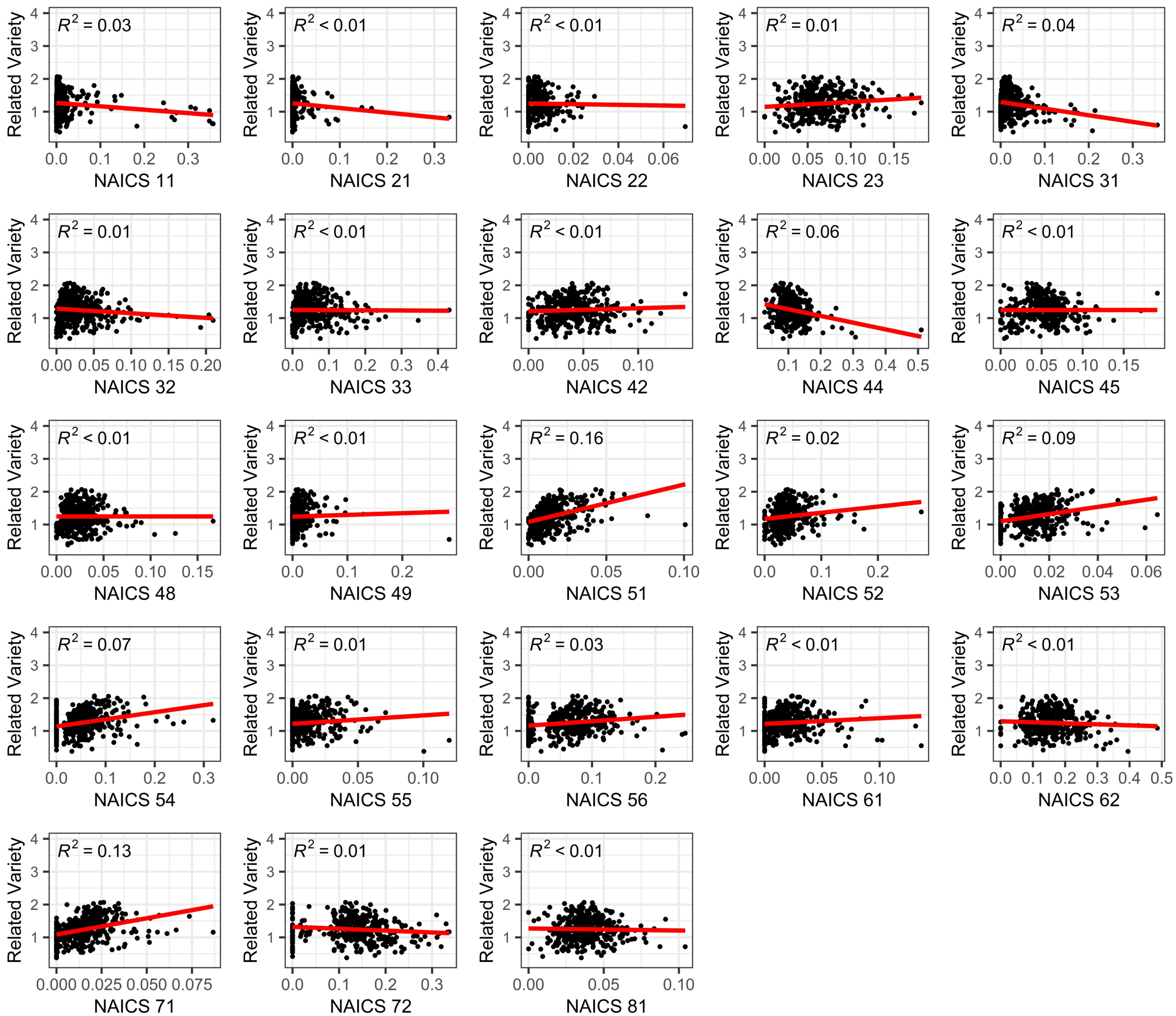

Figure A.1. Scatterplots of related variety over employment shares in two-digit NAICS codes for US MSAs, 2017.
Note: Unrelated variety was computed based on three-digit, related variety based on five-digit NAICS codes.

NAICS 11 (Agriculture, forestry, fishing and hunting); NAICS 21 (Mining, quarrying, and oil and gas extraction); NAICS 22 (Utilities); NAICS 23 (Construction); NAICS 31-33 (Manufacturing); NAICS 42 (Wholesale trade); NAICS 44-45 (Retail trade); NAICS 48-49 (Transportation and warehousing); NAICS 51 (Information); NAICS 52 (Finance and insurance); NAICS 53 (Real estate and rental and leasing); NAICS 54 (Professional, scientific and technical services); NAICS 55 (Management of companies and enterprises); NAICS 56 (Administrative and support, waste management and remediation services); NAICS 61 (Educational services); NAICS 62 (Health care and social assistance); NAICS 71 (Arts, entertainment and recreation); NAICS 72 (Accommodation and food services); NAICS 81 (Other services, except public administration)

Source: US Bureau of Labor Statistics (2021).