Agents of structural change

Frank Neffke, Matté Hartog, Ron Boschma, Martin Henning

October 2013*

Abstract

We study who introduces most structural change in a region. Are these entrepreneurs or existing firms? Do they come from inside or from outside the region? We approach these questions quantitatively, using employer-employee matched data on all workers in the economy to assess how related new activities are to the existing local economy. We find that incumbents reinforce a region’s implicit capability structure, partly by aligning their production with the local economy, whereas most impulses for structural change come from outside the region, with external firms playing a somewhat bigger role than entrepreneurs that enter the region from elsewhere.

Key words: Structural change, entrepreneurship, diversification, relatedness, regions, RBV.
Agents of structural change

“Our remote ancestors did not expand their economies much by simply doing more of what they had already been doing: piling up more wild seeds and nuts, slaughtering more wild cattle and geese, making more spearheads, necklaces, burins and fires. They expanded their economies by adding new kinds of work. So do we.” (Jane Jacobs, 1969, p. 49)

1. Introduction

In her *Theory of the growth of the firm*, Penrose (1959) famously argues that firms can only sustain growth if they not expand just the scale of their production, but also its scope. What is true for firms holds as well at the aggregate for the economies of cities (Jacobs, 1969): unless cities diversify into new activities, they will be unable to prosper in a changing competitive landscape. Detroit, for instance, went through a particularly devastating episode of this kind when the decline during the Great Recession of its automotive industry hit the city so hard that it eventually defaulted on part of its debt. Unlike a firm, however, a city does not act itself but relies on firms and entrepreneurs to introduce new activities. The aim of this study is to explore which economic agents are responsible for the most salient changes in regional economies. Are these entrepreneurs or existing firms? And, where do the entrepreneurs and firms that introduce most novelty in a location come from? Are they local or do they come from outside the region? Answers to these questions have so far relied on case studies of the rise of industrial clusters, with Silicon Valley as the archetypal example (Saxenian, 1994). By contrast, the present study’s aim is to determine who the main agents of structural change in a region are using an explicitly quantitative approach.

In order to do so, we propose a theoretical framework that highlights commonalities in the development paths of firms and regions. This framework conceives regions as endowed with bundles of resources, such as physical infrastructure, specialized labor markets and research organizations. From these resources, capabilities emerge that are often industry specific – they benefit some industries but not others. Together, these capabilities constitute a region’s capability base. We maintain that, even though regions do not act autonomously, regional capability bases outline the feasible development paths of local economies in the same way that the diversification opportunities for firms are conditioned by their resource bases.
Regional capability bases are not static but evolve when local firms undertake new economic activities. However, not all new activities transform a region to the same extent. Some require the same capabilities as existing activities, whereas others require capabilities that are new to the region. The former are easier to develop, but only the latter expand a region’s capability base and open up new opportunities for diversification. Therefore, we distinguish industrial change from structural change. Whereas industrial change is merely a change in a region’s industrial activities, structural change represents a more profound shift in production that involves changes in the region’s capability base.

This study contributes to the existing literature in three different ways. First, we construct a capability-based theoretical framework for regional diversification. We use this framework to derive hypotheses about which agents induce most structural change in a region from agents’ differential reliance on, and capacity to extract rents from, regional capability bases. Second, we propose quantitative instruments that infer the amount of structural change a region undergoes from the extent to which employment shifts to industries that are unrelated to the existing industries in the region. Third, we test these instruments on a comprehensive employer-employee linked dataset covering every worker in the Swedish economy between 1994 and 2010. These data allow us to assess (1) how much structural change is implied in observed shifts in regions’ industry mixes and (2) which economic agents transform a region’s capability base the most. The different types of economic agents we distinguish are the owners of existing establishments on the one hand, and the founders of new establishments on the other. Among the founders of new establishments, we further differentiate new establishments that belong to existing firms from those that belong to entrepreneurs. Finally, we subdivide these firms and entrepreneurs into local founders and founders from outside the region.

Whereas we find substantial shifts in regions’ industry mixes, the implied changes in underlying capability bases are minimal. This combination of volatility at the surface and inertia at the level of capability bases occurs because various agents shift regional capability bases in opposite directions. Whereas the growth, decline and industrial reorientation of existing establishments tend to reinforce a region’s existing capability base, new establishments truly change it and hence induce structural change. However, there are marked differences among establishment-founder types. Entrepreneur-owned establishments induce most structural change in the short run, but in the long run, new establishments of existing firms increasingly acquire this role. Moreover, we find that founders from outside the region and, especially, expanding firms from outside the region induce much more structural change than their local counterparts. These results prove to be robust, regardless of whether relatedness is measured in
terms of human capital requirements, input-output linkages or industries’ positions in the industry classification system.

The text is structured as follows. In section 2, we highlight the similarities in diversification in regions and firms and derive hypotheses on which agents will induce most structural change. In section 3, we introduce the data. In section 4, we describe our approach to measuring industrial and structural change. In section 5, we present the empirical findings. Section 6 summarizes and concludes.

2. Theory

The resource-based view (RBV) of the firm (Wernerfelt, 1984; Barney, 1991) conceives firms as bundles of resources. This perspective has been highly influential, shedding light on at least two important areas in management research. First, the RBV has inspired theoretical and empirical work on the relationship between resource characteristics and economic rents. Second, the RBV has shed light on how a firm’s resources affect its diversification path. Following Lawson (1999), we argue that the notion of a resource or capability base is not only applicable to firms, but also to regions.¹ This statement builds on four observations: (1) firms do not only leverage firm-internal but also locally available, firm-external resources and capabilities, many of which display characteristics associated with sustained competitive advantage; (2) these local resources are often fungible and (3) they sometimes grow as they are used more intensively; and (4) because resources become obsolete when technologies or demand changes, regions decline if their resource bases are not updated regularly. We will argue that observations (1) to (3) suggest that regional diversification is a path-dependent process, whereas observation (4) suggests that, analogously to the importance of dynamic capabilities in firms, regional capability bases must adapt to new economic realities.

However, two important differences between the capability bases of regions and those of firms must be noted. First, because firms control their internal capability bases, they can extract rents from them. In contrast, given that external resources are shared within a region, it is not immediately clear who will appropriate the rents of regional capability bases. Second, regional capability bases do not develop of the volition of a central actor. Instead, a region depends on firms and entrepreneurs for introducing new

¹ We prefer the term “regional capabilities” to Lawson’s (1999) “regional competences”, because being capable, more so than being competent, evokes the question “capable of what?” This outcome-specific orientation of capabilities will turn out to be pivotal in our measurement of structural change.
capabilities and retiring old ones. Indeed, the main question of the present study is how regional capability bases change, or to be more precise, who changes them.

A regional analogue to the resource-based view of the firm

Resources confer sustained competitive advantage to their owners if these resources are valuable, rare and hard to imitate and substitute (Barney, 1991). To extract rents, firms have to be in full command of a resource’s usage. That is why the RBV traditionally highlights firm-internal resources. However, firms also use locally available external resources. For instance, firms draw on a region’s available physical infrastructure, specialized supplier networks and skilled workforce. Although authors in different fields use different vocabularies to describe them, such regional, firm-external resources are identified in a variety of literatures. For instance, economic geographers argue that firms benefit from agglomeration externalities that derive from intra-regional labor market pooling, input-output linkages and knowledge spillovers (Glaeser et al., 1992; Henderson et al., 1995; Almeida and Kogut, 1999; McCann and Simonen, 2005; Faggian and McCann, 2006). In cluster research, elements of Porter’s (1990, 2000) diamond, such as the availability of production factors and the non-traded goods and services of supporting industries, constitute regional resources. Finally, the learning region and regional innovation systems frameworks (Cooke and Morgan, 1998) highlight the importance of regions’ untraded interdependencies (Storper, 1995), i.e., inter-firm and inter-organizational networks of knowledge exchange, which have also been described as localized capabilities (Maskell and Malmberg, 1999).

Regardless of what they are called, regional resources can help local firms compete in global markets if they are valuable, rare, inimitable and non-substitutable (VRIN). Many of the regional resources described above would fit this definition. That regional resources are often valuable and non-ubiquitous is all but beyond dispute in all the strands of regional research referred to above. Furthermore, in analogy to the inimitability requirement, the advantages regional capabilities provide are often highly localized because many of them are non-traded. However, regional capabilities are not necessarily non-substitutable, particularly if establishments can access firm-internal capabilities. We will leave this observation for what it is for now but return to it at the end of this section.

Apart from fulfilling VRIN conditions, like firm-internal resources, external resources are often specific to the economic activities they feed into. For instance, specialized car parts suppliers are of little use to pharmaceutical firms. Likewise, a local labor market that supplies skilled actuaries is much more valuable to insurance companies than to operators of spas. Because these resources are specific, the regional capability base they constitute supports some economic activities, but not others.
At the same time, external resources are often fungible (Teece, 1982). For example, although the presence of skilled mechanical engineers may not be useful to all economic activities, their services are valued in various manufacturing and business services industries. Moreover, external resources are not just used in, but also expanded by, local economic activities. For instance, workers are attracted to places with many employment opportunities that fit their skill profiles. Similarly, specialized suppliers are attracted to regions that host many potential clients. These processes are self-reinforcing: the firms that use specialized resources are attracted to regions where they are available at the same time that these specialized resources are attracted by the presence of firms willing to pay for them (Duranton and Puga, 2004).

**Regional capability bases and related diversification**

In sum, regional capability bases can provide competitive advantage to local firms, are often specific, yet fungible and grow as they are being used. In the context of firms, resources with these characteristics are typical preconditions for related diversification (Penrose, 1959; Teece 1982) because firms tend to expand into new activities to leverage their existing resources (Montgomery and Wernerfelt, 1989; Peteraf, 1993). However, because firms can also leverage existing regional resources, related diversification will not only characterize firm growth, but also the development of regional economies as a whole.²

Although the notion of related diversification has traditionally formed an integral part of the RBV discourse, economic geographers have emphasized the relatedness of regional diversification only relatively recently. That is not to say that the importance of related industries has not been noted. First of all, pioneering work on the role of inter-industry relatedness in regions is found in cluster research (Porter, 1998, 2003; Maskell, 2005; Delgado et al., 2013). For instance, Delgado et al. (2010) show that the presence of related industries increases entrepreneurial activity. Similarly, Florida et al. (2012) argue that related industries in a region give rise to geographies of scope that stimulate the local economy and Neffke et al. (2012) find that the presence of related industries raises establishments’ survival rates. Second, in urban economics, Ellison et al. (2010) and Dauth (2010) use a variety of relatedness measures to disentangle the effects of different Marshallian externality channels. Yet, the question of how

---

² Because a region’s capability base is not just accessible to existing firms, this related diversification does not have to rely solely on existing local firms. It may also be driven by start-ups or by firms in other regions that decide to invest in new facilities in the region. We will get back to this shortly.
relatedness affects the diversification of regions has not received nearly as much attention as the question of how it affects the diversification of firms.

Recently, however, Boschma and Frenken (2009) and Frenken and Boschma (2011) have argued that regional development is characterized by a branching process in which new, yet related activities are spun out of existing activities. This conjecture finds an increasing amount of empirical support. At the national level, Hildalgo et al. (2007) show that countries diversify their export portfolios according to such a branching logic. Neffke et al. (2011) show that similar processes are at work in the long-term development of Swedish regions, a result that has subsequently been replicated for regions in Spain (Boschma et al., 2012) and the United States (Essletzbichler, 2013; Muneepeerakul et al., 2013). We argue that this can be best understood by building on the conceptual apparatus of the RBV. Accordingly, regions grow through related diversification for similar reasons that firms do: regions have capability bases that are valuable, rare, specific to the existing set of economic activities and hard to access from outside the region.

**Structural change**

Related diversification is often insufficient to guarantee the long-term survival of a firm. In the long run, economic environments are not static. Changes in technologies may render resources obsolete and erode incumbent firms’ competitive advantage (Tushman and Anderson, 1986). Changes in demand or competitive pressures can have similar effects. Therefore, a preeminent question in RBV research is how firms manage to rearrange their resource configurations, which has led to research into how higher-level dynamic capabilities are used to update and acquire new lower-level operational capabilities (Henderson and Cockburn, 1994; Teece et al., 1997; Eisenhardt and Martin, 2000; Helfat and Peteraf, 2003).

However, not only firms face resource obsolescence. There is widespread agreement among cluster researchers and economic geographers that also regions must adapt to changes in the economic environment (Grabher, 1993; Pouder and St. John, 1996; Glaeser, 2005). That is, once the existing regional capabilities do not yield any competitive advantage anymore, the regional capability base must be renewed or lose its attraction. In much the same way as the new resource configurations (Eisenhardt and Martin 2000) that dynamic capabilities generate go beyond changing a firm’s product portfolio, renewal of the regional capability base goes beyond a mere change in the region’s industrial

---

3 An exception is Helfat and Peteraf (2003), who argue that capabilities change endogenously, following a capability life cycle.
employment composition. Indeed, capabilities exist at a deeper level than the products they help produce (Lawson, 1999). To distinguish between these two levels of regional diversification, we refer to a change in the industrial composition of a local economy as *industrial change*, whereas a transformation of the local capability base is denoted by *structural change*.

**Rents of regional capabilities**

Who will appropriate the rents from regional capabilities? In principle, the capability base of a region is available to all firms that locate there. Therefore, although local firms may gain a competitive advantage over firms outside the region, *a priori*, no such advantage should exist over local competitors. For instance, Pouder and St. John (1996) argue that, within a local cluster, firms are at *competitive parity*. Consequently, if firms can freely enter the region, the rents of a superior regional capability base do not necessarily accrue to the firms that use it. Instead they may end up with the owners of relatively inelastic local production factors, such as labor or land.\(^4\) However, regardless of who appropriates these rents, a region’s carrying capacity for a given industry will depend on the regional capability base’s general quality and on the extent to which it fulfills the industry’s particular needs. As a consequence, regional diversification will still be predominantly related diversification even if firms are unable to reap the benefits.

However, it is also likely that competitive parity within a region is too strong an assumption. Accessing regional capabilities often becomes easier as firms grow roots in a region (Grabher, 1993; Pouder and St. John, 1996; Bathelt et al., 2004). For instance, preferred access to local suppliers may require long-standing relationships (Ghemawat, 1986). Similarly, given the importance of social networks in job search, local firms will have less trouble finding the right workers than newcomers (Sorenson and Audia, 2000). In line with this, Dahl and Sorenson (2012) show that “regional tenure”, *i.e.*, the number of years an entrepreneur worked in a region, is almost as potent a predictor of a venture’s success as industry tenure.

Because access to the regional capability base requires time, entry is restricted. This allows local incumbents to appropriate at least some of the rents to external capabilities. It also suggests that not all actors benefit equally from the regional capability base. Consequently, the importance firms attach to

\(^4\) Indeed, urban economists sometimes seek evidence for agglomeration externalities in elevated wages or house prices instead of in the profits of local firms (Rosenthal and Strange, 2004). For instance, Glaeser (2005) argues that much of the competitive advantage that Boston’s highly educated workforce offered in the knowledge economy era has shown up in its house prices and wage levels, instead of in a growing population, as had been the case elsewhere.
regional capabilities, and therewith, the degree to which these capabilities affect corporate strategy, differs by firm, an insight that we use to develop hypotheses on who induces most structural change in regions.

**Agents of structural change**

As we noted earlier, the regional analogue to the dynamic capabilities that allow a firm to reconfigure its resources resides in the ways in which economic agents affect the industrial composition of a region. We distinguish among a number of different types of economic agents. First, there are the region’s existing establishments. Existing establishment affect the regional employment structure, and in its wake, the regional capability base, when they expanding employment, reduce it, change industrial orientation (industry switching) or exit the region. Second, change can be introduced by new establishments that enter the region. These new establishments are set up by existing firms or by entrepreneurs, which themselves can either come from inside (local agents) or from outside the region (non-local agents).

Which of these agents are most likely to introduce new capabilities to a region? Our discussion on rent appropriation suggested that not all agents are equally well-positioned to access regional capabilities. Indeed, establishments will differ in (1) their access to regional capabilities, (2) their access to capabilities in other regions and (3) their overall reliance on regional capabilities.

Starting with the first, we argue that agents that can access regional capabilities more easily will be more likely to build on the existing regional capabilities and, therewith, less likely to introduce new capabilities in the region. Because access to regional capabilities is easier if firms have developed ties with other actors in the region, local incumbents should be less likely to induce structural change than new establishments:

**Hypothesis 1:** Incumbent establishments are less likely to induce structural change in the region than new establishments.

Second, agents differ in the extent to which they can access capabilities that are currently unavailable in the own region but that are available elsewhere. For instance, Grabher (1993) and Pouder and St. John (1996) argue that firms are often strongly conditioned by a local dominant logic. In contrast, agents that enter the region from the outside bring with them the capabilities they developed in other regions, infusing the region with new ideas, skills and relations. This suggests that local agents are less likely to change the region’s capability base than agents that enter the region from elsewhere:
Hypothesis 2: New establishments of local entrepreneurs and firms are less likely to induce structural change in the region than those of non-local entrepreneurs and firms.

Third, agents differ in the extent to which they depend on local capabilities. In particular, new establishments of existing firms often have access to their parents’ firm-internal capabilities to substitute for regional capabilities. Therefore, these establishments can develop activities that rely on capabilities that had previously been unavailable in the region. In as far as these new capabilities leak to others in the region, the regional capability base expands. In contrast, entrepreneur-owned establishments do not have access to parent-firm capabilities. This suggests that entrepreneurs will be more reliant on regional capabilities and induce less structural change than expanding firms.

At the same time, since the writings of Schumpeter (1942), entrepreneurship has been associated with new combinations, innovation, and structural change. For instance, entrepreneurs are typically more risk-taking (Cramer et al., 2002) and creative (Zhao and Seibert, 2006) than the average population, qualities that make them likely sources of structural change. Given these contradictory considerations, we put both hypotheses to the test:

**Hypothesis 3a:** New establishments of expanding firms induce more structural change than those of entrepreneurs.

**Hypothesis 3b:** New establishments of entrepreneurs induce more structural change than those of expanding firms.

3. **Data**

The data we use to test these hypotheses are derived from the administrative records of Sweden. These records contain yearly information on individuals’ workplaces and incomes for the country’s entire workforce. The income information distinguishes among income derived from wages, from a private business or from both, which allows identifying entrepreneurs as workers with incomes from private businesses. Each establishment is linked to a firm identifier that is shared by all establishments that are owned by the same parent firm. Furthermore, information is provided on the location and industry affiliation of the establishments.

---

5 The “leaking” of capabilities can occur through all the self-reinforcing mechanisms of the spatial concentration of resources that were described before, such as knowledge spillovers, the formation of specialized labor pools and a growing availability of specialist suppliers.
We collapse the individual level data at the region-industry level to analyze the economic dynamics in 110 labor market regions in Sweden between 1994 and 2010. Labor market areas are ideally suited for measuring the capabilities embedded in the local labor force – the component of the regional capability base we focus on – because they are defined such that most commuting takes place within their borders. We use industries defined at the 4-digit level of the European NACE classification, which distinguishes among over 700 different industries. Finally, we focus on the years 1994 to 2010, because the prevailing industry classification systems (NACE 1.0 and NACE 1.1) can be harmonized without much difficulty, whereas classificatory changes beyond these years are hard to accommodate.

An important assumption in this paper is that locally available capabilities influence an establishment’s location choice. However, in some industries, location choice is severely restricted by the availability of natural resources or by the need to be close to population centers. Therefore, when defining a region’s industry mix, we focus on 259 traded, non-natural-resource-based industries in the private sector, excluding non-traded services (e.g., retail stores, restaurants), government activities and natural-resource-based activities (e.g., mining and agriculture).

4. Measurement
To test the hypotheses formulated in section 2, we have to quantify the diversification (or increased focus) of the regional capability base that each agent type brings about. However, it is useful to first clarify some terminology. We already distinguished diversification of a region’s industry mix (industrial change) from diversification of a region’s capability mix (structural change). However, the term “diversification” is also ambivalent because it can be used in a static and in a dynamic sense. The static use of the word refers to the existing variety within a portfolio of economic activities, whereas the dynamic use of the word refers to a change of the portfolio and involves a comparison of the portfolio at two different points in time.

We will call the static, level-of-variety counterpart to industrial change the diversity of a region. Commonly used measures include the count of different activity classes in a portfolio, or the Herfindahl or entropy index for the size distribution of these activity classes. The dynamic, change-of-portfolio notion of diversification, i.e., industrial change, can be measured by turn-over rates of activity classes or

---

6 The Stata code for this can be downloaded from the authors’ websites.
7 See Appendix A. However, because all industries contribute to the local capability base, we still take the omitted industries into account when measuring the match between local industries and their regional capability base.
by, for instance, the cosine distance of a portfolio’s activity size distributions over time. Moving to the context of structural change, the static notion of diversification refers to the coherence of economic activities in a portfolio in terms of overlap in capability requirements.\footnote{Although the word “coherence” generally evokes positive associations, caution is warranted. Coherent regions are not necessarily better off than incoherent regions. On the one hand, coherent regions have a compact capability base, which is easier to maintain. On the other hand, this compactness also limits diversification options. Indeed, it is likely that, particularly in the long run, some intermediate level of coherence is optimal. Again, this argumentation mirrors an existing discussion in the RBV literature on the optimal level of diversification of firms (Palich et al., 2000). The issue of optimality is left for future research.} The dynamic notion, structural change, refers to a change in the portfolio’s implied capability base. Table 1 summarizes these distinctions along the static-dynamic and the activity-capability axes.

TABLE 1 ABOUT HERE

Although we cannot directly observe regions’ capability bases, the existence of capabilities manifests itself in a region’s industry mix.\footnote{This is not just a practical problem of gathering the required data. Because capabilities exist at a deeper level than the activities that use them, they are to some extent inherently latent. The easiest, if ultimately not the only, way to assess whether a capability is present, is to determine whether the object of the capability is achieved. For instance, a firm’s strong bio-technology capabilities are expressed in its display of processes and products that use these bio-technology capabilities.} That is, we can infer that a region that produces cars must host car-making capabilities. Obviously, to deduce for each product X in a region that the region must have X-making capabilities achieves little. However, we can make progress by using information on the relatedness among economic activities. Industries are commonly thought of as related if they require similar resources or capabilities (Farjoun, 1994; Teece et al., 1994; Bryce and Winter, 2009). Therefore, a region expands its capability base, whenever it diversifies into an industry that is unrelated to its current portfolio of industries.

This is best illustrated by an example. Take a region with a traditional focus on making shoes. For this region, diversification into sandal manufacturing represents a much smaller change than diversification into the motion picture business. After all, making movies requires a vast array of new capabilities, such as those in special effects, casting agencies and movie studios. At the same time, there is little use for shoemakers’ leather stitching skills. Therefore, a successful launch of a local motion picture industry will significantly change the regional capability structure, paving the way for a whole set of new industries, related to film-making, to emerge.

Based on the above considerations, we quantify regional coherence and structural change by first determining how related industries are to one another in terms of their capability requirements. Given
such estimates, we calculate how related an industry is to the basket of industries that constitute a region’s industry mix. We will call this the *regional capability match*, or simply *match*, of an industry to a region. Next, regional coherence is quantified as the average regional match of all industries in the region. Finally, structural change is measured by assessing how strongly the current industry mix conforms to the region’s past capability base. This procedure is summarized in Table 2 and explained in greater detail in the following sections.

**TABLE 2 ABOUT HERE**

**Inter-industry relatedness: skill relatedness**

The first building block in measuring coherence and structural change is relatedness. Inter-industry relatedness can be measured in several ways (for an overview, see Neffke and Henning, 2013). In this paper, we focus on relatedness in terms of similarities in professional skills or *skill relatedness*. We do this for three reasons. First, the skills embedded in a firm’s human capital are among its most valuable resources (Grant 1996; Grant and Spender 1996) and often condition the firm’s diversification path (Porter, 1987; Neffke and Henning, 2013). Second, human capital can and is shared among firms in a region. It therewith acts as an important channel of local knowledge exchange and local externalities (Almeida and Kogut, 1999). Third, similarities in human capital requirements are quantifiable as shown by Neffke and Henning (2013). The logic behind Neffke and Henning’s skill-relatedness measure is that workers are in general reluctant to switch to jobs where their current skills are not valued. Therefore, industries that demand similar skills typically display large labor flows among them. Inter-industry skill relatedness should thus reveal itself in excessive inter-industry labor flows.

Using the simplified version proposed in Neffke et al. (2013), we measure the skill relatedness between a pair of industries, \(i\) and \(j\), as the ratio of observed to expected worker flows between the industries, where expectations are based on overall mobility rates in both industries.

\[
SR_{ij} = \frac{F_{ij}}{(F_{i}F_{j})/F_{..}}
\]  

10 Our findings (reported in section 5) do not depend on the use of this particular relatedness indicator. Indeed, as documented in Appendix D, all main results also hold when measuring relatedness using value chain linkages or the industry classification system.
\( F_{ij} \) represents the observed labor flow from industry \( i \) to industry \( j \). Where the index \( i \) or \( j \) is replaced by a dot, the flows are summed over this omitted category, such that \( F_i = \sum_j F_{ij} \), \( F_j = \sum_i F_{ij} \) and \( F_\cdot = \sum_{i,j} F_{ij} \). \(^{11}\)

The term \( \left( F_i, F_j \right) / F_\cdot = F_i F_j / F_\cdot \) in equation (1) captures the expected flows from \( i \) to \( j \), assuming that \( j \) receives a share of total worker flows from \( i \) that is proportional to the share of inflows \( j \) receives from any industry in the economy. We take \( SR_{ij} \) values greater than one to signal that industries are skill related, whereas values between zero and one indicate the industries are unrelated. The resulting skill-relatedness index is highly predictive of corporate diversification (Neffke and Henning, 2013), stable over time and similar for workers in different wage categories and occupations (Neffke et al., 2013).

**Industry-region capability match**

Skill relatedness is a quantity that characterizes industry-industry pairs. To assess how related a given industry is to a region’s economy, a region-industry relationship, we calculate how much related employment the industry finds in the region. The more related employment there is, the stronger the industry’s match with the region’s capability base is supposed to be. Let \( E_{i\cdot r t}^\text{rel} \) be all employment in industries related to industry \( i \) in region \( r \) in year \( t \).\(^{12}\)

\[
E_{i\cdot r t}^\text{rel} = \sum_{j \neq i} E_{j\cdot r t} \cdot I(SR_{ij} > 1)
\]  

(2)

where \( E_{j\cdot r t} \) represents the employment of industry \( j \) in region \( r \) in year \( t \) and \( I(SR_{ij} > 1) \) an indicator function that evaluates to one if its argument is true and to zero otherwise. The match of industry \( i \) to region \( r \) in year \( t \) is defined as the degree to which the region is overspecialized in industries related to industry \( i \), or as the location quotient of related employment:

\[
LQ_{i\cdot r t}^\text{rel} = \frac{E_{i\cdot r t}^\text{rel} / E_{\cdot r t}}{E_{i\cdot t}^\text{rel} / E_{\cdot t}}
\]

(3)

where \( E_{\cdot r t} \) is the total employment in the region in year \( t \), \( E_{i\cdot t}^\text{rel} \) the total employment in related industries in the country, and \( E_{\cdot t} \) the overall employment in the country. If \( LQ_{i\cdot r t}^\text{rel} \) is greater than one,

\(^{11}\) Detailed industry-industry labor flows can be derived from our data because we can follow all Swedish workers throughout their careers over the period 1994-2010. This results in yearly skill relatedness estimates that we average over the entire period to reduce measurement errors as suggested by Neffke et al. (2013). The exact procedure is described in Appendix B.

\(^{12}\) Because they add to the regional capability base, related employment includes all employment in non-traded, public sector and natural-resource-based industries.
the employment share of related industries in the region exceeds their share in the national economy. If it is smaller than one, the region has a smaller share of related industries than the national economy does.

By construction, \( LQ_{irt}^{rel} \) has a strongly asymmetric distribution: whereas an overrepresentation of related industries ranges from 1 to infinity, the underrepresentation of related industries lies between zero and one.\(^{13} \) This will affect the average match values we construct later on. We therefore transform \( LQ_{irt}^{rel} \) as follows:

\[
\overline{LQ}_{irt}^{rel} = \frac{LQ_{irt}^{rel} - 1}{LQ_{irt}^{rel} + 1}
\]

\( \overline{LQ}_{irt}^{rel} \) now ranges from -1 (no related employment) to +1 (a complete concentration of all related employment in region \( r \)). Because \( \frac{LQ_{irt}^{rel} - 1}{LQ_{irt}^{rel} + 1} = -\frac{(1/LQ_{irt}^{rel}) - 1}{(1/LQ_{irt}^{rel}) + 1} \), a given level of overrepresentation of related employment has the same magnitude but opposite sign as the same level of underrepresentation. For instance, if \( LQ_{irt}^{rel} = 2 \), \( \overline{LQ}_{irt}^{rel} = \frac{1}{3} \), whereas \( LQ_{irt}^{rel} = \frac{1}{2} \) implies \( \overline{LQ}_{irt}^{rel} = -\frac{1}{3} \).

**Regional coherence and structural change**

Whereas capability match is a characteristic of a local industry, i.e., of a region-industry pair, coherence is a region characteristic. We define coherence as the employment-weighted average capability match of a region’s industries:

\[
C_{rt} = \sum_{i} \frac{E_{i.r.t}}{E_{..t}} \overline{LQ}_{irt}^{rel}
\]

(5)

The coherence tells us how related a region’s industry mix is to the local economy as a whole. The higher the coherence, the more related the industries in the region are to one another. As a baseline, we also calculate how strongly the national industry mix matches the capability base of a region \( r \):

\[
B_{rt} = \sum_{i} \frac{E_{i.t}}{E_{..t}} \overline{LQ}_{irt}^{rel}
\]

(6)

where \( \frac{E_{i.t}}{E_{..t}} \) is industry \( i \)’s share in total national employment.

\(^{13} \) For instance, an industry for which related industries are twice as large in the region as in the national economy, \( M_{irt} \) equals 2. However, in the reverse situation (related industries’ share of the national economy is twice as large as the one of the regional economy), \( M_{irt} \) equals 0.5.
The dynamic counterpart to coherence, structural change, can be measured in much the same way. However, instead of asking how related a region’s industry mix is to the current local economy, we ask how related the industry mix is to the local economy of year $T$:

$$S_{rt,T} = \sum_i \frac{E_{irt}}{E_{rt}} \frac{Q_{irT}}{LQ_{irT}}, T < t$$  \hspace{1cm} (7)

**Structural change by agent type**

The regional industry mix changes when economic agents create or destroy employment in local industries. When agents create employment in industries that strongly match the existing capability base, agents reinforce the focus of that capability base. When the agents destroy employment in such industries, the capability base’s focus shifts. When the capability match is weak, employment creation shifts the capability base and employment destruction deepens it. Therefore, we subdivide all establishments into types that create and those that destroy employment. Incumbent establishments that retain their present industry affiliation are divided into three groups: growing, declining and exiting incumbents. Incumbents that switch industries create employment in the industry they enter and, at the same time, destroy employment in the industry they leave. Therefore, we split industry switchers into two artificial types: “out-switching” incumbents and “in-switching” incumbents. For new establishments, we distinguish among those of local expanding firms, of non-local expanding firms, of local entrepreneurs and of non-local entrepreneurs. Table 3 provides an overview of all agent types.

**TABLE 3 ABOUT HERE**

A detailed description of how we determine establishment ownership and geographic origins is provided in Appendix C. In short, we first identify expanding firms through the establishments’ firm identifiers. If this firm identifier is shared with any other establishment, we know that the establishment is part of a larger firm. Furthermore, a new establishment is said to belong to a local firm, if in the year prior to its founding, the parent firm employed most of its employees outside the new establishment’s labor market area. By contrast, if the founding of a new establishment also leads to the creation of a new firm, we regard the establishment as entrepreneur-owned. We identify the entrepreneurs in such establishments as the workers who earn income from a private business. The origin of an entrepreneur-owned establishment is determined by its entrepreneur prior workplace. Again, entrepreneurs coming from within the new establishment’s labor market area are regarded to be local entrepreneurs, whereas

---

14 In principle, firms may also move to another region. However, this is such a rare event that we do not further explore it.
all others are seen as non-local entrepreneurs. This approach allows us to identify the origins of all new subsidiaries of existing firms and of some 35,000 out of about 60,000 entrepreneur-owned establishments. The 25,000 establishments whose origins we cannot determine are dropped from the analyses.

The structural change associated with each agent type depends on the average capability match to the local economy in the base year, $T$, of the industries in which agents create or destroy employment up until the current year, $t$:

$$A_{rt,T}^a = \sum_{e \in e(a)} \left[ \frac{\Delta_{et,T}^a}{\Delta_{e(r)t,T}^a} \right] \hat{L}q_{rel}^{re} \cdot (e)_{r(e)T},$$

where $\frac{\Delta_{et}^a}{\Delta_{e(r)t}^a}$ is the total employment created (or destroyed) by an establishment of type $a$ in industry $i(e)$ in region $r(e)$ between the base year $T$ and the current year $t$ ($\Delta_{et,T}^a$) as a share of the total employment created (destroyed) by all of agent $a$’s establishments in the region ($\Delta_{e(r)t,T}^a$). In other words, $A_{rt,T}^a$ shows how strongly an agent type’s new (or destroyed) employment conforms to the local economy of year $T$. To facilitate interpretation, we normalize $A_{rt,T}^a$ by subtracting the average match of the existing employment mix in the base year, that is, by subtracting the region’s coherence in year $T$:

$$\tilde{A}_{rt,T}^a = A_{rt,T}^a - C_{rT}$$

Agents that add employment to industries with above average capability match ($\tilde{A}_{rt,T}^a > 0$) deepen the focus of the current capability base, whereas agents that add employment to below-averagely matched industries ($\tilde{A}_{rt,T}^a < 0$) shift it. The opposite happens if agents destroy employment.

5. Results

Diversity and industrial change in Swedish regions

Figure 1 shows the evolution of the diversity of Swedish regions. For each year, it depicts the employment entropy of regions’ industry mixes averaged over all regions. On average, regions show no tendency of becoming more or less specialized: average diversity stays constant throughout the entire time period. However, as shown in Figures 2 and 3, this apparent stability masks significant industrial...
change. Figure 2 shows that 23% of all local industries\textsuperscript{15} in 2010, appeared in or after 1994 and that 27% of the local industries in 1994 had disappeared by 2010. Moreover, not only is there significant turnover in local industries, regions’ industrial employment compositions become increasingly dissimilar to their 1994 compositions (Figure 3).

Coherence and structural change

Figure 4 shows the coherence of regions and how it evolves over time. The average coherence exceeds its proportional employment baseline in every single year by several standard deviations, which shows that regions are much more coherent than they would have been by chance. We interpret this finding as an indication that the industry composition of a regional economy draws on a narrower set of regional capabilities, in much the same way that a firm’s product portfolio is based on the firm’s core competences. Given the observed industrial change, one would expect the capability base of regions to change as well, which would be reflected in the coherence and structural change indicators. However, the average coherence between 1994 and 2010 fluctuates marginally between 0.02 and 0.05, without any statistically significant shifts. Regional capability bases could still have changed if they developed in a way that gradually shifted their capability bases\textit{ en bloc} while keeping coherence constant. Figure 5 shows this is not the case. Regions are not drifting away from their 1994, base-year capability bases. On the contrary, the 2010 regional industry mixes match their corresponding 1994 regional economies on average as strongly as they match the 2010 regional economies. Apparently, Sweden’s regions have undergone little to no structural change since the mid-1990s.

\textsuperscript{15} A local industry is defined as a region-industry combination, such as for instance shipbuilding-in-Gothenburg.
Agents of structural change

Short-term structural change

A possible reason for why we find on average no structural change at the regional level is that different agents change the capability base in opposite ways, thereby cancelling each other’s actions out in the aggregate. We investigate this by first looking at the structural change each agent induces within a period of one year (Figure 6).\footnote{Incumbents are defined as all establishments that exist in the base year, 1994. However, to increase the sample of agents that set up new establishments, we take all new establishments between 1994 and 1999. Next, we record the structural change induced one year after they are founded. That is, for new establishments, we pool $\bar{A}_{T}^{a}$, $\bar{A}_{T+1}^{a}$, $\bar{A}_{T+2}^{a}$, $\bar{A}_{T+3}^{a}$, $\bar{A}_{T+4}^{a}$, and $\bar{A}_{T+5}^{a}$.} Whether agents trigger structural change depends on whether they generate or destroy employment and in which industries they do so. The vertical axis lists all different agent types. On the horizontal axis, we plot the corresponding $\bar{A}_{t,T}^{a}$ values, together with their 95% confidence interval. Agent types that generate employment are denoted by a green plus sign, those that destroy employment by a red minus sign. Positive values of $\bar{A}$ indicate that an agent type is generally found in industries that match the region more strongly than the (employment-weighted) average local industry. If the agents generate employment in strongly matched industries, the current focus of the capability base increases. If they destroy employment in such industries, the focus weakens. Negative values of $\bar{A}$ correspond to industries with a below-average match to their regions. In these industries, the reverse holds: employment creation diversifies the capability base, whereas employment destruction narrows it.

The agent-specific structural change values are overlaid on a diagram (in blue) that shows where agents create or destroy employment in the distribution of capability match values of existing local industries. This graph therewith illustrates the economic significance of an agent’s $\bar{A}$ value. For instance, a green plus at $\bar{A} = -0.05$ means that the corresponding agent type on average creates employment in local industries in the 33rd match-percentile, placing it in the bottom third of all existing employment in Swedish local industries in terms of the industries’ capability match with their regions.

The absence of structural change at the aggregate level indeed conceals that different agents draw the regional capability base in opposite directions. Incumbent establishments tend to increase the
specialization of the local capability base. If they grow, they do so predominantly in above-averagely matched local industries. If they shrink or close down, they mostly reduce employment in below-averagely matched local industries. However, what affects the capability base most is incumbent establishments’ industry switching, shifting employment from, on average, the 36th to the 53rd match-percentile.

New establishments, in contrast, tend to diversify a region’s capability base. Indeed, in support of Hypothesis 1 (incumbents induce less structural change than new establishments), almost all types of new establishments have negative $\tilde{A}$ values and therefore induce structural change. The only exception is new establishments of expanding local firms, which reinforce the existing capability base. Since these new establishments represent incumbent growth that needs to be accommodated in new facilities however, it is not surprising that they behave much like growing incumbents.

Within the group of new establishments, those that belong to existing firms are on average found in more strongly matched industries (40th match-percentile), than those of entrepreneurs (26th match-percentile). This supports Hypothesis 3b over Hypothesis 3a: entrepreneurs induce more structural change than expanding firms.

Regarding the agents’ geographical origins, we find that non-local agents (from outside the region) induce much more structural change than those from within the region. On average, local expanding firms create employment in the 59th match-percentile, against the 30th for non-local firms. The difference between local (31st match-percentile) and non-local entrepreneurs (22nd match percentile) is smaller, but still economically and statistically highly significant. We therefore conclude that there is strong support for Hypothesis 2.

The fact that non-local agents seem to widen the capability base of a region, suggests that they are important in the diffusion of capabilities. We explore this notion by looking at non-local agents’ home regions. If these agents indeed diffuse capabilities, there should be a high capability match of the agents’ industries with their home region’s capability base. Table 4 shows that this is indeed the case: new establishments’ capability match to their owners’ home regions is much higher than to the establishments’ host regions.

TABLE 4 ABOUT HERE
**Long-term structural change**

So far, we have investigated employment dynamics on a one-year horizon. However, structural change is typically associated with a much longer time horizon. To induce long-lasting structural change, establishments in industries with a low capability match need to survive and grow. To investigate this, Figure 7 summarizes the structural change each agent type induces over a time period of 10 years, *i.e.*, it displays $\tilde{A}_{T+T}$ for $t = T + 10$. Again, the cumulative distribution of match values for the base year 1994 is provided as a reference.

FIGURE 7 ABOUT HERE

The patterns of long-term structural change are very similar to those of short-term structural change. On a ten-year horizon, incumbents still reinforce the focus of the base year’s capability base. However, this must now be attributed almost fully to growing incumbents, while the effect of exits and product switching is much more muted. It is important to note that the incumbents and new plants now consist of establishments that existed or where founded since 1994 and were able to stay in business for at least 10 years. Any differences between 1- and 10-year structural change figures arise, therefore, from differences in the short- and long-term survival and growth rates at different points of the match distribution. In particular, the fact that exits and industry switching shift to the right when comparing Figure 6 to Figure 7 (*i.e.*, occur at lower match values) suggests that unrelated activities were selected out early on through establishment closures and adjustments of establishments’ industry orientation. The group of growing incumbents has also shifted to the right. This suggests, in line with the finding that firms benefit from nearby related economic activity (Delgado et al., 2010; Neffke et al., 2012), that incumbents in highly-matched industries outperformed incumbents in low-match industries.

Turning to new establishments, we see that entrepreneur-owned establishments’ $\tilde{A}$-values also shifted to the right. Growth and/or survival of these establishments has thus been concentrated in the higher-matched part of the distribution. In contrast, the new establishments of existing firms remained at the same match-value (local firms) or even moved leftward (non-local firms). This means that expanding firms in low-match industries grew more and/or survived longer than those in high-match industries. As a consequence, new establishments of (non-local) expanding firms all but surpass entrepreneurs as the main agents of structural change.

To summarize, in the long run, we still find strong support for Hypothesis 1 and 2. However, the support for Hypothesis 3a over Hypothesis 3b has disappeared: expanding firms and entrepreneurs induce equal
amounts of structural change. However, if we limit the analysis to non-local agents, expanding firms induce more structural change than entrepreneurs.

**Robustness**

So far, we can summarize our findings as follows:

1) Swedish regions are coherent and this coherence does, on average, not change significantly;
2) on average, there is little to no structural change in Swedish regions;
3) existing establishments reinforce the existing capability base of regions whereas new establishments decrease it;
4) entrepreneur-owned establishments induce more structural change in the short than in the long run, whereas the reverse holds for the new establishments of existing firms; and
5) non-local agents induce significantly more structural change than agents from within the region.

These results have been derived using our skill relatedness measure. That is, we measured the capability base in terms of the skills of the local workforce. However, the choice of this particular relatedness index is not critical. Indeed findings 1) to 5) all hold also with other relatedness indicators. We show this by replicating graphs corresponding to Figures 4, 5, 6 and 7 both for input-output relatedness and for the widely used industry-classification-based relatedness index (see Appendix D).

6. **Conclusion**

**Summary and discussion**

In our theoretical framework, we highlighted the many parallels that exist between the RBVs description of firm growth and the way in which regional economies develop. Both firms and regions grow not just by enlarging the scale but also the scope of their economic activities. Firms increase their scope predominantly by entering fields in which they can leverage existing resources. This then leads to related diversification. However, because there are many resources that exist not within firms, but among groups of firms, it is sensible to think of regions as being endowed with capability bases. Regional capability bases are specific to the economic activities that use them, yet at the same time fungible. As a consequence, also the economic development of regions is characterized by related diversification. In line with this, we find that, although regions’ industry mixes change constantly, this change does not seem to affect the underlying capability base much. Apparently, at the aggregate, structural change progresses much more slowly than industrial change.
However, this conceals opposite dynamics at the micro-level of economic agents. We find that the entrepreneurs, who are often put forward as crucial drivers of structural change, indeed play this role, but mostly in the short run. In the long run, entrepreneurs prove hard-pressed to successfully run activities that are unrelated to the existing industries in the region. In contrast, new establishments of existing firms are much more likely to sustain such unrelated activities. Furthermore, actors from outside the region induce much more structural change in a region than local actors. This empirically substantiates the often-held believe that change comes from outside the region (Grabher, 1993; Bathelt et al., 2004). Indeed, we find that the mobility of firms and entrepreneurs is an important channel through which industries and capabilities diffuse to other regions.

Caveats

When interpreting these findings, a number of limitations of our study should be noted. First, we determined the main agents of regional structural change in terms of the intensity, not the amount of structural change they induce. However, some agent types are much more numerous than others. For instance, entrepreneurs set up far more establishments than existing firms do, local entrepreneurs outnumber non-local entrepreneurs 5-to-1 and non-local firms 20-to-1. Therefore, although in intensity terms, local entrepreneurs play a minor role in structural change, as a group, they constitute an important factor in shifting the local capability base.

Furthermore, one can think of alternative implementations of the tools we introduced to measure structural change. For instance, one can insert other inter-industry relatedness measures than the ones used in this study. Moreover, the capability match was quantified as the overrepresentation of related employment in the region. An alternative approach is Hidalgo et al.’s (2007) density measure, which does not calculate how much related employment an industry finds locally, but how related the region’s employment is to this industry. Exploring such alternatives and assessing their relative merits constitutes interesting follow-up research.

Implications for local policy making, corporate strategy and future research

Local policy makers may benefit both from the findings and the tools constructed in this study. For instance, our findings suggest that policies that aim at renewing a local economy will be more successful if they incorporate strategies to attract firms from regions in which the targeted industry is well-established. More fundamentally, however, the study’s explicit reasoning in terms of regional capability bases, together with the indices that have been created, may shift attention away from descriptions of
regions in terms of their industry mix, toward an understanding of the region’s strength and weaknesses at the deeper level of capabilities.

At the level of corporate strategy, managers and entrepreneurs may use capability matches as a tool to identify regions with a suitable capability base for their new activities. Moreover, the distinction between diversity and industrial change on the one hand and coherence and structural change on the other, can be readily extended to the context of firms: also firms can have more or less coherent capability bases, depending on how related the products they produce are.

Finally, the study raises a number of interesting questions for future research. First, our finding that new establishments of existing firms are better able to grow and survive in unrelated environments than stand-alone establishments raises the question of why this is the case. We proposed that corporate establishments draw on their parent firms’ capabilities, but whether this is indeed the case and when, how and across what distance multi-establishment firms can accomplish this are open questions. Another interesting finding is that firms switch affiliations from low-match to high-match industries. This suggests that firm strategies interact with regionally available capabilities in ways that are yet poorly understood. We hope that the analyses in this paper will prove useful in answering these and other questions on how regional economies and their capability bases co-evolve with the firms they host.
REFERENCES


Tables and figures

Table 1: Diversity, industrial change, coherence and structural change

<table>
<thead>
<tr>
<th></th>
<th>Static</th>
<th>Dynamic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industries</td>
<td>Diversity</td>
<td>Industrial change</td>
</tr>
<tr>
<td>Measured by:</td>
<td>Herfindahl, entropy</td>
<td>Measured by:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cosine distance</td>
</tr>
<tr>
<td>Underlying question:</td>
<td>How many different activities are there and how equal is their size distribution?</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Underlying question:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>How fast are new activities introduced and how much does the size distribution of activities change?</td>
</tr>
<tr>
<td>Capabilities</td>
<td>Coherence</td>
<td>Structural change</td>
</tr>
<tr>
<td>Measured by:</td>
<td>see section 3</td>
<td>Measured by:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>see section 3</td>
</tr>
<tr>
<td>Underlying question:</td>
<td>How similar are the capabilities required by the various activities in the portfolio?</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Underlying question:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>To what extent does the capability base change due to changes in the portfolio of activities?</td>
</tr>
</tbody>
</table>
Table 2: Definitions and relationships among quantities

<table>
<thead>
<tr>
<th>quantity</th>
<th>unit of analysis</th>
<th>definition</th>
<th>description</th>
<th>normalization</th>
<th>range</th>
</tr>
</thead>
<tbody>
<tr>
<td>labor low</td>
<td>industry-industry</td>
<td>( F_{ij} )</td>
<td>How many people change jobs from industry ( i ) to ( j )?</td>
<td></td>
<td>([0, \infty))</td>
</tr>
<tr>
<td>skill relatedness</td>
<td>industry-industry</td>
<td>( SR_{ij} = \frac{F_{ij}}{F_i F_j} )</td>
<td>How related are two industries to one another?</td>
<td>( \tilde{SR}<em>{ij} = \frac{SR</em>{ij} - 1}{SR_{ij} + 1} )</td>
<td>([-1, 1))</td>
</tr>
<tr>
<td>employment</td>
<td>industry-region</td>
<td>( E_{irt} )</td>
<td>How many workers does industry ( i ) employ in region ( r ) in year ( t )?</td>
<td></td>
<td>([-1, 1))</td>
</tr>
<tr>
<td>related employment</td>
<td>industry-region</td>
<td>( E_{irt}^{rel} )</td>
<td>How much related employment to industry ( i ) is there in region ( r ) in year ( t )?</td>
<td>( \sum_{r \neq i} E_{irt} I(SR_{ij} &gt; 1) )</td>
<td>([-1, 1))</td>
</tr>
<tr>
<td>capability match</td>
<td>industry-region</td>
<td>( LQ_{irt}^{rel} = \frac{E_{irt}^{rel} E_{rt}}{E_{rt}^{rel} E_{t}} )</td>
<td>How overrepresented are related industries in the region?</td>
<td>( \tilde{LQ}<em>{irt}^{rel} = \frac{LQ</em>{irt}^{rel} - 1}{LQ_{irt}^{rel} + 1} )</td>
<td>([-1, 1))</td>
</tr>
<tr>
<td>coherence</td>
<td>region</td>
<td>( C_{rt} = \sum_{i} \frac{E_{irt}}{E_{rt}} LQ_{irt}^{rel} )</td>
<td>How strongly matched are a region’s activities on average?</td>
<td>( \hat{C}<em>{rt} = C</em>{rt} - B_{rt} )</td>
<td>([-2,2))</td>
</tr>
<tr>
<td>coherence baseline</td>
<td>region</td>
<td>( B_{rt} = \sum_{i} \frac{E_{irt}}{E_{rt}} LQ_{irt}^{rel} )</td>
<td>How strongly matched would a random distribution of activities have been in the region?</td>
<td></td>
<td>([-1,1))</td>
</tr>
<tr>
<td>structural change</td>
<td>region</td>
<td>( S_{rt,T} = \sum_{i} \frac{E_{irt}}{E_{rt}} LQ_{irt}^{rel} )</td>
<td>How strongly matched would a region’s current activities have been in the industry mix of year ( T )?</td>
<td>( \tilde{A}<em>{rt,T}^{a} = A</em>{rt,T}^{a} - C_{rt} )</td>
<td>([-1,1))</td>
</tr>
<tr>
<td>structural change by agent-region</td>
<td>agent-region</td>
<td>( A_{rt,T}^{a} = \sum_{e \in e(a)} \left[ \frac{\Delta e_t / \Delta e_{(e)(r)}^T}{LQ_{(e)(r)}^{rel}} \right] LQ_{(e)(r)}^{rel} )</td>
<td>How much structural change does an agent type induce on average?</td>
<td>( \tilde{A}<em>{rt,T}^{a} = A</em>{rt,T}^{a} - C_{rt} )</td>
<td>([-2,2))</td>
</tr>
</tbody>
</table>

\[ \Delta e_t = \sum_{e' \in e(e)} \Delta e_{e',t} \]
Table 3: Agent types

<table>
<thead>
<tr>
<th>Agent type</th>
<th>Description</th>
<th>ΔE</th>
<th>Effect on capability base if inds are under-embedded</th>
<th>Effect on capability base if inds are over-embedded</th>
</tr>
</thead>
<tbody>
<tr>
<td>incumbent establishments</td>
<td>existing establishments that ...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>growing</td>
<td>expand their workforce</td>
<td>+</td>
<td>diversify</td>
<td>specialize</td>
</tr>
<tr>
<td>shrinking</td>
<td>reduce their workforce</td>
<td>-</td>
<td>specialize</td>
<td>diversify</td>
</tr>
<tr>
<td>closing</td>
<td>close down</td>
<td>-</td>
<td>specialize</td>
<td>diversify</td>
</tr>
<tr>
<td>industry switchers</td>
<td>existing establishments that ...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>into the industry</td>
<td>switch into the industry</td>
<td>+</td>
<td>diversify</td>
<td>specialize</td>
</tr>
<tr>
<td>out of the industry</td>
<td>switch out of the industry</td>
<td>-</td>
<td>specialize</td>
<td>diversify</td>
</tr>
<tr>
<td>New establishments</td>
<td>new establishments set up by ...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>local expanding firms</td>
<td>pre-existing firm with main employment concentration inside the region</td>
<td>+</td>
<td>diversify</td>
<td>specialize</td>
</tr>
<tr>
<td>non-local expanding firms</td>
<td>pre-existing firm with main employment concentration in another region</td>
<td>+</td>
<td>diversify</td>
<td>specialize</td>
</tr>
<tr>
<td>local entrepreneurs</td>
<td>new firm created by entrepreneur(s) from inside the region</td>
<td>+</td>
<td>diversify</td>
<td>specialize</td>
</tr>
<tr>
<td>non-local entrepreneurs</td>
<td>new firm created by entrepreneur(s) from outside the region</td>
<td>+</td>
<td>diversify</td>
<td>specialize</td>
</tr>
</tbody>
</table>

Column ΔE indicates whether the employment change associated with a given agent type is positive or negative. The final two columns indicate which effect this employment change has on the regional capability base if the change takes place in industries that are less (column 4) or more (column 5) strongly matched to the region than the average existing local industry in the industry (i.e., if the match is below or over the region’s coherence).
Table 4: Knowledge diffusion by outside agents

<table>
<thead>
<tr>
<th>Agent type</th>
<th>Capability match to:</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>home region</td>
<td>host region</td>
</tr>
<tr>
<td>Non-local expanding firms</td>
<td>0.061</td>
<td>-0.019</td>
</tr>
<tr>
<td>Non-local entrepreneurs</td>
<td>0.058</td>
<td>-0.017</td>
</tr>
</tbody>
</table>

The home region is defined as the region in which the new establishment’s parent firm employed most of its workers (non-local firms) or as the region in which the new establishment’s entrepreneur (non-local entrepreneurs) was employed in the year prior to opening up the new establishment.
The figure graphs the development of the average employment entropy of Swedish regions over time. Employment entropy is a measure of how diversified a local economy is and is calculated as $\text{entropy}_{rt} = -\sum_{i=1}^{N} \frac{E_{irt}}{E_{rt}} \ln \frac{E_{irt}}{E_{rt}}$, where $E_{irt}$ denotes the employment in industry $i$, region $r$ and year $t$, and $E_{rt} = \sum_{i} E_{irt}$. It varies from zero when all employment is concentrated in a single industry, to $\ln N$ when all industries have an equal employment share in the region. The error bars depict a 95% confidence interval calculated as $\pm 1.96$ times the entropy’s standard deviation across regions.
The solid blue line depicts the share of local industries (region-industry combinations) in Sweden that existed in 1994, i.e., that had non-zero employment in 1994, that survived at least to year $1994 + t$. The dotted red line depicts the share of local industries existing in 2010 that had existed already in year $2010 - t$. 
Figure 3: Average cosine similarity to the base year 1994 of labor market regions’ employment profiles

The graph depicts the development of the average cosine similarity between a region’s current industrial employment mix and the industrial employment mix of the base year 1994. The cosine similarity measures the similarity of two vectors, in this case, the region’s employment profile at two different points in time: \( \cos \text{ sim}_{rT} = \frac{e_r \cdot e_T}{||e_r|| \cdot ||e_T||} \), where \( e_r = (E_1r, ..., E_rT)' \) a vector whose elements where \( E_{ir} \) correspond to region \( r \)’s employment in industry \( i \) in year \( t \). The cosine distance ranges from -1 (opposite profiles) through 0 (unrelated profiles) to +1 (same profile). The error bars depict a 95% confidence interval calculated as ±1.96 times the cosine similarity’s standard deviation across regions.
The upper line depicts the development of the average coherence of a region’s capability base. It is measured by its local industries’ employment-weighted average capability match to the regional economy as a whole ($C_{rt} = \sum \frac{E_{rt}}{E_r} \bar{Q}_{ir}^{rel}$). As a baseline, the lower line depicts the development of the average capability match of Sweden’s aggregate, national industries to the region’s capability base ($B_{rt} = \sum \frac{E_{rt}}{E_r} \bar{Q}_{ir}^{rel}$). The error bars depict a 95% confidence interval calculated as ±1.96 times the coherence and baseline’s standard deviations across regions.
Figure 5: Structural change in Sweden’s labor market regions

The graph depicts the development of the average capability match of a region's local industries to the local economy of the base year 1994 ($S_{rt,1994} = \sum_{i} \frac{E_{rt}}{F_{rt}} L_{Q,rel}^{i}$), including 95% error bars. In the presence of structural change, $S_{rt,1994}$ should fall (diversification) or rise (increased existing focus) over time. The error bars depict a 95% confidence interval calculated as ±1.96 times the standard deviations in structural change across regions.
Figure 6: Structural change by agent type over a 1-year horizon

The markers shows for each agent type the employment weighted average capability match minus the region’s coherence values of the local industries in which the agent type creates or destroys employment within the first year, averaged across all establishments of agents $a$. That is, the graph shows by how much the agent’s capability match to its region exceeds the regions coherence. Therefore, employment created (destroyed) at $\hat{A} > 0$ correspond to a diversification (further focusing) of the regional capability base. Employment-creating agents are denoted with a green plus sign, employment-destroying agents with a red minus sign. The error bars depict a 95% confidence interval for this average capability match, based on the standard deviation of the capability match across all establishments that belong to the corresponding agent type. To facilitate interpretation of $\hat{A}$-values, the area shaded in blue depicts the employment-weighted cumulative density for the distribution of capability match minus coherence values of all local industries as they existed in 1994. That is, for each $\hat{A}$-value, it provides the percentile in the distribution of match values of the existing economy.
The markers show for each agent type the employment weighted average capability match minus the region’s coherence values of the local industries in which the agent type creates or destroys employment within the first ten years, averaged across all establishments of agents $a$. That is, the graph shows by how much the agent’s capability match to its region exceeds the region’s coherence. Therefore, employment created (destroyed) at $\tilde{A} > 0$ correspond to a diversification (further focusing) of the regional capability base. Employment-creating agents are denoted with a green plus sign, employment-destroying agents with a red minus sign. The error bars depict a 95% confidence interval for this average capability match, based on the standard deviation of the capability match across all establishments that belong to the corresponding agent type. To facilitate interpretation of $\tilde{A}$-values, the area shaded in blue depicts the employment-weighted cumulative density for the distribution of capability match minus coherence values of all local industries as they existed in 1994. That is, for each $\tilde{A}$-value, it provides the percentile in the distribution of match values of the existing economy.
## Appendix A: Classification of industries

### Table A1: Industries included in the analyses

<table>
<thead>
<tr>
<th>Industry codes</th>
<th>Description</th>
<th>Definition industry</th>
<th>Included</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000 - 1499</td>
<td>Agriculture, hunting and forestry + Fishing + Mining and quarrying</td>
<td>Traded, resource-based</td>
<td>no</td>
</tr>
<tr>
<td>1500 - 3999</td>
<td>Manufacturing</td>
<td>Traded, not resource-based</td>
<td>yes</td>
</tr>
<tr>
<td>4000 - 4999</td>
<td>Electricity, gas and water supply + Construction</td>
<td>Non-traded</td>
<td>no</td>
</tr>
<tr>
<td>5000 - 5199</td>
<td>Wholesale and retail trade, repair of motor vehicles, motorcycles and personal and household goods</td>
<td>Non-traded</td>
<td>no</td>
</tr>
<tr>
<td>5200 - 5299</td>
<td>Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods</td>
<td>Non-traded</td>
<td>no</td>
</tr>
<tr>
<td>5500 - 5599</td>
<td>Hotels and restaurants</td>
<td>Non-traded</td>
<td>no</td>
</tr>
<tr>
<td>6000 - 6420</td>
<td>Transport, storage and communication</td>
<td>Non-traded</td>
<td>no</td>
</tr>
<tr>
<td>6500 - 6999</td>
<td>Financial intermediation, except insurance and pension funding</td>
<td>Traded, not resource-based</td>
<td>yes</td>
</tr>
<tr>
<td>7000 - 7199</td>
<td>Real estate + Renting activities</td>
<td>Non-traded</td>
<td>no</td>
</tr>
<tr>
<td>7200 - 7399</td>
<td>Computer and related activities + Research and development</td>
<td>Traded, not resource-based</td>
<td>yes</td>
</tr>
<tr>
<td>7400 - 7499</td>
<td>Other business activities</td>
<td>Traded, not resource-based</td>
<td>yes</td>
</tr>
<tr>
<td>7500 - 7599</td>
<td>Public administration and defense, compulsory social security</td>
<td>public sector</td>
<td>no</td>
</tr>
<tr>
<td>7600 - 8599</td>
<td>Education, Health and social work</td>
<td>public sector</td>
<td>no</td>
</tr>
<tr>
<td>8600 - 9999</td>
<td>Other community, social and personal service activities + Activities of households + Extra-territorial organizations and bodies</td>
<td>public sector</td>
<td>no</td>
</tr>
</tbody>
</table>
Appendix B: Measuring skill-relatedness between industries

We measure skill-relatedness between industries by assessing the labor flows between industry pairs between 1994 and 2010. During these years, about 4.5 million workers in Sweden switched to a new job in another 4-digit industry. First, we use equation (1) to calculate skill relatedness for each year between 1994 and 2010. When years are indexed by \( t \) and summation over omitted categories is indicated by '.' this yields:

\[
SR_{ijt} = \frac{F_{ijt}}{(F_{jt}F_{lt})/F_{t}}. \tag{B1}
\]

Because this measure is highly asymmetric, we use the same transformation as in equation (4) to map it \( SR_{ijt} \) onto the interval \([-1, 1)\):

\[
\tilde{SR}_{ijt} = \frac{SR_{ijt}-1}{SR_{ijt}+1}. \tag{B2}
\]

Hence, industry \( i \) is skill-related to industry \( j \) if \( \tilde{SR}_{ijt} > 0 \). Then, for every industry pair, we average \( \tilde{SR}_{ijt} \) over all yearly flows between 1994 and 2010:

\[
M\tilde{SR}_{ij} = \frac{1}{16} \sum_{t=1994}^{2009} \tilde{SR}_{ijt}. \tag{B3}
\]

Fourth and finally, we symmetrize the measure so that \( SS\tilde{R}_{ij} = SS\tilde{R}_{ji} \): \n
\[
SS\tilde{R}_{ij} = \frac{M\tilde{SR}_{ij}+M\tilde{SR}_{ji}}{2}. \tag{B4}
\]

The actual condition for two industries to be skill related that we evaluate in the indicator function in equation (2) is therefore \( SS\tilde{R}_{ij} > 0 \).
Appendix C: Determining the founders and geographical origins of new establishments

To identify the origins of each establishment, we first determine whether a new establishment is an entrepreneurial entry or an entry by an existing firm. Every establishment has a specific establishment identifier and firm identifier assigned to it by Statistics Sweden (see Andersson and Arvidsson, 2006), which allows following establishments over time regardless of changes in ownership or legal status. Entrepreneurial entries are new establishments that create new firms (both the establishment identifier, variable and firm identifier are new in year $t$). New establishments of existing firms arise when the establishment identifier is new in year $t$ but the establishment’s firm identifier already existed in year $t - 1$.

To determine the geographical origin of new establishments, we proceed as follows. For every new establishment of pre-existing firms, the geographical origin is determined as the region where the parent firm employed most of its workers in the year prior to the new establishment’s creation. To identify the previous location of the founder or founders of entrepreneurial entries, we take a number of steps. First, Statistics Sweden supplies information on employment status to identify workers who derive income from a private venture. If only one person in the new establishment is classified as an entrepreneur according to this variable, we take that person as the establishment’s entrepreneur. The region where she was employed in the previous year is now used as the geographical origin of the new establishment. If a new establishment employs multiple entrepreneurs, and if all these entrepreneurs worked in the same region in the previous year, we take this region as the geographical origin. If no entrepreneur is found and of the new establishment has only 1 employee, we assume that this person is the founder and we take the region in which that person worked in the previous year as the geographical origin. If no entrepreneur is found and if the new establishment has multiple employees, and if all these entrepreneurs worked in the same region the year before, we take this region as the geographical origin of the new establishment.

Using this procedure, we were able to trace the origins of 35,000 new establishments that did not belong to pre-existing firms. All other new establishments were dropped from the analyses.
Appendix D: Alternative relatedness measures

We repeated all analyses reported in the main text with two alternative relatedness indicators to skill relatedness. The first is based on the industry classification system (*NACE-relatedness*). The second relatedness index is based on input-output relations among industries. Below we describe how each relatedness measure is constructed and then replicate Figures 4-7 based on the described index.

Industry-classification-based relatedness (NACE)

To measure NACE-relatedness, we classify the 4-digit industries in the European NACE classification as related when they belong to the same 2-digit sector. For instance, ‘Manufacture of cast iron tubes’ (industry code 2721) and ‘Manufacture of steel tubes’ (industry code 2722) are related because they belong to the same 2-digit sector 27 ‘Manufacture of basic metals’. The corresponding graphs are shown below.

![Figure D1: Coherence of labor market regions’ capability bases (NACE-relatedness)](image-url)

See Figure 4.
Figure D2: Structural change in Sweden’s labor market regions (NACE-relatedness)

See Figure 5.
Figure D3: Structural change by agent type over a 1-year horizon (NACE-relatedness)

See Figure 6.
Figure D4: Structural change by agent type over a 10-year horizon (NACE-relatedness)

See Figure 7.
Input-output relatedness

Input-output linkages are derived from the Swedish input-output table of 1995, which is available from Statistics Sweden. For each pair of industries, \((i, j)\), we calculate the share of industry \(i\)’s inputs that are sourced from industry \(j\) and the share of industry \(i\)’s output that is consumed by industry \(j\). We then average both numbers to arrive at a measure of input-output relatedness between the two industries. If \(CF_{ij}\) represents the value of the commodity flow of industry \(i\) to industry \(j\), then the input-output relatedness between industries \(i\) and \(j\), \(IOR_{ij}\), is given by:

\[
IOR_{ij} = \frac{1}{2} \left( \frac{CF_{ij}}{\sum_k CF_{ik}} + \frac{CF_{ji}}{\sum_l CF_{lj}} \right)
\]

Input-output data are only available at the 2-digit level. Because we use industries at the 4-digit level, we assume that the input-output linkages that exist between two 2-digit sectors are representative of the linkages that exist among the 4-digit industries of which these sectors comprise. We choose the threshold value for when two industries are related in such a way that the same number of industry-pairs are input-output related as skill related. Below, we present the outcomes when \(IOR\) is used as the relatedness measure:

Figure D5: Coherence of labor market regions’ capability bases (input-output relatedness)
Figure D6, Structural change in Sweden's labor market regions (input-output relatedness)

See Figure 5.
Figure D7: Structural change by agent type over a 1-year horizon (input-output relatedness)

See Figure 6.
Figure D8: Structural change by agent type over a 10-year horizon (based on input-output relatedness)

See Figure 7.