

Agents of structural change

The role of firms and entrepreneurs in regional diversification

Frank Neffke, Matté Hartog, Ron Boschma, Martin Henning

April, 2014*

Abstract

Who introduces structural change in regional economies: Entrepreneurs or existing firms? And do local or non-local founders of establishments create most novelty in a region? Using matched employer-employee data for the whole Swedish workforce, we determine how unrelated and therefore how novel the activities of different establishments are to a region's industry mix. Up- and downsizing establishments cause large shifts in the local industry structure, but these shifts only occasionally require an expansion of local capabilities because the new activities are often related to existing local activities. Indeed, these incumbents tend to align their production with the local economy, deepening the region's specialization. In contrast, structural change mostly originates via new establishments, especially those with non-local roots. Moreover, although entrepreneurs start businesses more often in activities unrelated to the existing regional economy, new establishments founded by existing firms survive in such activities more often, inducing longer-lasting changes in the region.

Key words: Structural change, entrepreneurship, diversification, relatedness, regions, resource-based view

Acknowledgments: We are grateful to Sarah Hopkinson for excellent research assistance. We thank Mercedes Delgado, Dario Diodato, Ricardo Hausmann, William Kerr and participants at the CID-LEP seminar and the EMAEE and DRUID conferences for valuable comments. Ron Boschma acknowledges financial support from the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement no. SSH-CT-2010-266959.

*: F. Neffke: Harvard Kennedy School, Harvard University, M. Hartog: Urban and Regional research center Utrecht, Utrecht University, R. Boschma: CIRCLE, Lund University and Urban and Regional research center Utrecht, Utrecht University, M. Henning: Lund University.

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 expand not just the scale of their production, but also the scope of production. What is true for firms holds at the aggregate level of the economies of cities (Jacobs, 1969): unless they diversify into new activities, cities will be unable to prosper in a changing competitive landscape.¹ However, unlike a firm, a city and its surrounding region do not act for themselves, but instead must rely 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 structural change in a region. Are the most important economic agents in this process entrepreneurs or existing firms? And, does novelty arise from local entrepreneurs and firms, or is it introduced by actors from outside the region? Up until now, answers to these questions have relied on case studies of the rise of highly visible, often singularly successful industrial clusters, with Silicon Valley as the archetypal example (Saxenian, 1994). By contrast, this study’s aim is to use an explicitly quantitative approach to draw more generalizable conclusions on who are the main agents of structural change in regional economies.

The empirical analyses put forward in this study are informed by the notion that there are commonalities in the development paths of firms and regions. This framework conceives of regions as endowed with bundles of resources, such as physical infrastructure, specialized labor markets and research organizations (Lawson, 1999; Boschma 2004). From these resources, capabilities emerge that are often industry specific in the sense that they cannot be applied to all activities, but nor are these capabilities entirely specialized as groups of related industries utilize the same or similar capabilities.. Taken together, these capabilities constitute what we will call a region’s *capability base*. We argue that,

¹ Detroit, for instance, went through a particularly devastating episode of this kind when the Great Recession hit the city’s automotive industry so hard that it eventually defaulted on part of its debt (Pendall et al, 2010).

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.

These regional capability bases are not static but evolve when firms undertake new economic activities in the region. However, not all new activities transform a region to the same extent. Some draw on the same capabilities as existing activities, whereas others require capabilities that are new to the region. The former, which we term *industrial change*, are therefore easier to develop, but only the latter, which we term *structural change*, expand the region's capability base and open up new opportunities for diversification. Industries are commonly defined as sets of firms that operate on the same market. For instance, web-design and computer games aim at different markets and therefore constitute different industries. However, both economic activities build on similar capabilities. Although a diversification from web-design into computer games would change the industrial composition of a region, this change would not amount to much structural change. In other words, we think of industrial change as merely a change in a region's industrial composition, whereas structural change represents a more profound alteration of the industry composition of a region that coincides with shifts in the region's capability base.

This study contributes to the existing literature in three ways. Firstly, we construct a capability-based theoretical framework for regional diversification. We use this framework to derive hypotheses about which agents induce most structural change, using as a starting point how agents differ in their reliance on and capacity to extract rents from regional capability bases. Secondly, we propose quantitative instruments that can infer how much structural change a region undergoes by looking at 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 that covers every worker in the Swedish economy between 1994 and 2010. These data allow us to assess (1) how much structural change is implied in the observed shifts in regions' industry mixes (2) which economic agents transform a region's capability base the most and (3) which agent types are most persistent in the creation of novelty, meaning that their establishments endure even where adequate capability bases have not been fully developed yet.

We study changes in the capability base of a region by investigating which types of economic agents create or destroy employment in industries that are unrelated to a region's pre-existing industries. In terms of the types of economic agents, we distinguish between 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 and non-local founders.

Although we find substantial reallocation of employment across a region's industries, most of these employment shifts take place among industries that are closely related and thus employ similar capabilities. Our interpretation of this finding is that the volatility of the industrial profile of a region often does not translate into a radical reorientation of the underlying capability base. Moreover, decomposing these changes by agent type shows that the growth, decline and industrial reorientation of existing establishments all tend to reinforce a region's existing capability base (i.e., increase the relatedness among industries in a region), whereas new establishments are often set up in more unrelated activities and hence induce more structural change. However, there are marked differences between the establishments of different founder types. Entrepreneur-owned establishments induce most structural change in the short run, but in the long run, this role is increasingly assumed by new establishments of existing firms. Indeed, whereas the long-term survival rates of entrepreneur-owned establishments (i.e., start-ups), are lower in regions with few related activities, we find no such relation for establishments that belong to a larger firm. Moreover, non-local firms and entrepreneurs set up establishments that are less likely to be related to the other economic activities in the region they locate in than their local counterparts (and therefore represent more structural change). These results are 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 paper is structured as follows. In section 2, we highlight the similarities between regional and firm diversification and derive hypotheses regarding 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) understands 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.² 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 typically associated with sustained competitive advantage; (2) these local resources are often specific, yet also to a certain degree fungible; (3) local resources sometimes grow when they are used more intensively; and (4) given that resources become obsolete with the inevitable changes in technologies or final demand, regions decline if their resource bases are not updated regularly. Observations (1) to (3) suggest that regional diversification is a path-dependent process, while observation (4) suggests that, in parallel to the oft-stated importance of dynamic capabilities in firms, regional capability bases must adapt to new economic realities.

In spite of their commonalities, regional and firm capability bases differ in at least two important respects. First, because firms control their internal capability bases, they can often extract rents from them. In contrast, given that external resources are shared within a region, it is not obvious who will appropriate the rents of a *regional* capability base. Second, regional capability bases do not develop by the volition of a central actor. Instead, a region depends on firms and entrepreneurs to introduce new capabilities and retire 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 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. Firm-external resources have been identified in a variety of literatures, albeit using different vocabularies. 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

² 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 is pivotal to a region's diversification path.

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) or "localized capabilities" (Maskell and Malmberg, 1999), such as inter-organizational networks of knowledge exchange.

Regardless of the terminology used to describe them, regional resources can help local firms compete in global markets if they are valuable, rare, inimitable and non-substitutable (*VRIN*, see Barney, 1991). Many of the regional resources described above would fit this definition. Indeed, that regional resources are often valuable and non-ubiquitous is beyond dispute in all the strands of regional research referred to above. Furthermore, analogous to the inimitability requirement, regional capabilities are often highly localized because many of them are non-tradeable. However, regional capabilities are not necessarily non-substitutable, particularly if establishments can access firm-internal capabilities, a particularity to which we return to at the end of this section.

Apart from fulfilling VRIN conditions, external resources, like firm-internal resources, are often specific to the economic activities they contribute to. 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.

However, at the same time, external resources are often somewhat 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 multiple manufacturing and business services industries.

Finally, external resources often grow the more they are used. For instance, qualified 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: firms that use specialized resources are attracted to regions where these resources are available, while these specialized resources are attracted by the presence of firms willing to pay for them (Duranton and Puga, 2004).

Capability bases and related diversification in regions

To summarize the argument thus far, regional capability bases are often specific, yet fungible, grow as they are being used and can provide competitive advantage to local firms. In the context of firms, resources with these characteristics are often considered preconditions for related diversification (Penrose, 1959; Teece 1982) because firms tend to expand into new activities that 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.³

Although the notion of related diversification has traditionally formed an integral part of the RBV discourse, economic geographers have only recently emphasized the relatedness of regional diversification. That is not to say that the importance of related industries *per se* has remained unnoticed. Pioneering economic geography work on the role of inter-industry relatedness 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 increases the survival rate of manufacturing plants. Similarly, 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. Despite recognizing the importance of various types of inter-industry linkages, the question of how such relatedness affects diversification has not received nearly as much attention in the literature on regional growth as in work on firm growth.

Recently, however, the topic has enjoyed growing attention. For instance, Frenken and Boschma (2007) and Boschma and Frenken (2011) argue that regional development is characterized by a branching process in which new, yet related activities spin out of existing activities. This conjecture has been accruing more and more empirical support. At the national level, Hidalgo *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.*, 2013) and the United States

³ Note that, because a region's capability base is not just accessible to existing firms, this related diversification does not have to rely solely on the actions of existing local firms. It may also be driven by start-ups or by firms in other regions that decide to invest in the region. We will return to this shortly.

(Essletzbichler, 2013; Muneeppeerakul *et al.*, 2013). We argue that important clues about the underlying mechanisms of this phenomenon can be derived from 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: the importance of unrelated diversification

Related diversification is often unable to guarantee the long-term survival of a firm. In the long run, economic environments are not static. Changes in technologies render resources obsolete and erode incumbent firms' competitive advantage (Tushman and Anderson, 1986). Changes in demand or competitive pressures have similar effects. Therefore, a key question in RBV research is how firms manage, not just to diversify into new products, but also to rearrange their underlying resource configurations. This 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, it is not only firms who face resource obsolescence. There is consensus in much of the cluster and economic geography literatures that regions must also adapt to changes in the economic environment (Grabher, 1993; Polder and St. John, 1996; Glaeser, 2005). Thus, once the existing regional capabilities become insufficient for firms to compete at global markets, the regional capability base must be renewed or lose its attraction. In other words, some degree of unrelated diversification may be required. In much the same way as the "new resource configurations" (Eisenhardt and Martin 2000) generated by dynamic capabilities go beyond changing a firm's product portfolio, renewal of the regional capability base goes beyond a mere change in the region's industrial 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 will be denoted by *structural change*.

Rents of regional capabilities

An important difference between firms' resources bases and regions' capability bases concerns the issue of the appropriation of rents. The capability base of a region is, in principle, 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, p. 1203) argue that, within a local cluster, firms are at “competitive parity”. Consequently, if firms can freely enter a 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 local production factors with a relatively inelastic supply, such as labor or land.⁴ However, regardless of who appropriates these rents, a region’s carrying capacity for a given industry will be influenced by the extent to which the regional capability base fulfills that industry’s particular needs. As a consequence, regional diversification will still be predominantly related diversification even if local firms are unable to reap the benefits.

However, the assumption of competitive parity is probably too strong, because not all firms will benefit equally from locally available capabilities. For instance, Alcacer and Chung (2007) argue that the extent to which firms can absorb local knowledge spillovers depends on their own technological sophistication. Moreover, accessing regional capabilities becomes easier as firms grow roots in a region (Grabher, 1993; Pouder and St. John, 1996; Storper and Venables 2004). For instance, preferred access to local suppliers may require long-standing relationships (Ghemawat, 1986). Moreover, not all firms participate equally in the local networks that transmit knowledge (Giuliani 2007). Similarly, given the importance of social networks in job search, local firms will find suitable workers more easily than newcomers will (Sorenson and Audia, 2000). Indeed, Dahl and Sorenson (2012) show that “regional tenure”, *i.e.*, the number of years an entrepreneur has worked in a region, is almost as strong a predictor of a venture’s success as industry tenure.

Access to the regional capability base requires time, and thus entry is restricted. This may enable local incumbents to appropriate at least some of the rents to regional capabilities. It also suggests that not all actors benefit equally from the regional capability base. Consequently, the importance firms attach to regional capabilities, and therewith, the degree to which these capabilities affect corporate strategy, differs by firm. In the next section, we use this insight to develop hypotheses on which agent type induces most structural change in a region.

⁴ Indeed, urban economists often 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.

Agents of structural change

Ultimately, the types of activities developed and, consequently, the capabilities used in a region depend on the production decisions of firms and entrepreneurs. Therefore, the regional counterpart to what the management science literature calls dynamic capabilities (Teece *et al.*, 1997) resides in the ways in which firms and entrepreneurs affect the capability base of a region by creating, expanding and destroying economic activities. We distinguish between two different types of economic agents that can induce change in a region. Firstly, there are the region's existing establishments. When they expand or reduce employment, change industrial orientation or exit the region, existing establishments affect the regional employment structure and, concurrently, the regional capability base. Secondly, new establishments can act as agents of change. New establishments can be set up by existing firms or by entrepreneurs, who themselves originate from either inside (local agents) or outside the region (non-local agents).

Having differentiated between the different types of economic agents, who then are most likely to introduce new capabilities to a region? 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 who can access regional capabilities more easily are more likely to build on existing regional capabilities and, hence, less likely to introduce new capabilities into the region. Because access to regional capabilities is easier if firms have already developed ties 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.

Secondly, several authors (e.g., Storper, 1995; Poudier and St. John, 1996, Lawson and Lorenz, 1999; Gertler, 2003; Boschma, 2004) argue that firms are often strongly conditioned by a locally dominant logic. In contrast, agents that enter the region from elsewhere will lack access to some of the capabilities in their new region, but they may also bring with them the capabilities they developed in other regions, infusing the new region with 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.

Thirdly, 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, which may act as substitutes for regional capabilities. Therefore, these establishments can develop activities that rely on capabilities that do not exist in the region. If these new capabilities spill over⁵ to other firms 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, therefore, induce less structural change than the new establishments of existing firms.

At the same time, there is a long history of thought that associates entrepreneurship with structural change. Indeed, at least 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 person. Given these contradictory considerations, both (opposing) hypotheses are justifiable:

Hypothesis 3a: New establishments of entrepreneurs are less likely to induce structural change in the region than those of existing firms.

Hypothesis 3b: New establishments of existing firms are less likely to induce structural change in the region than those of entrepreneurs.

3. Data

The data we use to test these hypotheses are derived from the administrative records of Sweden, access to which was provided by Statistics Sweden (SCB). These records contain yearly information on individuals' workplaces and incomes for the country's entire workforce. The income information distinguishes between income derived from wages and from a private business, allowing us to identify entrepreneurs. Each establishment has a firm identifier that is shared by all establishments owned by

⁵ Capabilities spillovers 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.

the same parent firm. Furthermore, we have information on establishments' location and industry affiliation.

We aggregate the individual level data to the region-industry level to analyze employment dynamics in 110 labor market regions in Sweden between 1994 and 2010. Industries are defined at the 4-digit level of the European NACE classification, which distinguishes over 700 different industries and can be harmonized easily for the period 1994 to 2010.

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 the large numbers of customers in urban agglomerations. 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 and restaurants), government activities and natural-resource-based activities (*e.g.*, mining and agriculture).⁶

4. Measurement

To test the hypotheses formulated in section 2, we have to quantify how much each agent type diversifies the regional capability base. The term diversification is often used either in a static sense ("How diversified is a region?") or in a dynamic sense ("How much did the portfolio of economic activities change in a given time period?"). Together with the distinction between industrial and structural change the word "diversification" can therefore actually refer to four different concepts. Firstly, we have the static concept of *industrial diversity* as measured by, for instance, the number of different industries in a region or a Herfindahl or entropy index of the employment distribution across industries. Secondly, the corresponding dynamic, change-of-portfolio, notion of diversification, measures how much the industrial composition of a region changes. This *industrial change* is typically measured by regional entry and exit rates of industries or by, for instance, the cosine distance to a base year of the region's industrial employment distribution as it evolves over time. When we move from industries to capabilities, the static notion of diversification refers to the *coherence* (or lack thereof) of economic activities in a region in terms of overlap in capability requirements.⁷ The dynamic notion,

⁶ See Appendix A. Because all industries contribute to the local capability base, we do take the omitted industries into account when measuring the capability match between an industry and the local capability base.

⁷ 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

structural change, refers to a change in the portfolio's implied capability base. Table 1 summarizes these distinctions along the static-dynamic and the industry-capability axes.

TABLE 1 DIVERSICATION MATRIX

Although we cannot observe regions' capability bases directly, capabilities manifest themselves in a region's industry mix.⁸ That is, we can infer that a region that produces cars must host car-making capabilities. Obviously, little is gained by deducing for each product X in a region that the region must have X-making capabilities. However, we can make progress by using information on the relatedness among economic activities. We say that industries are related if they require similar resources or capabilities (Farjoun, 1994; Teece *et al.*, 1994; Bryce and Winter, 2009). Consequently, whenever a region diversifies into an industry that is unrelated to its current portfolio of industries, it expands its capability base.

Take, for instance, a region with a traditional focus on shoe-making. 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 in the motion picture business for shoemakers' leather stitching skills. Therefore, a successful launch of a local motion picture industry would significantly change the regional capability structure, paving the way for a whole set of new industries related to film-making to emerge.

This line of reasoning suggests that regional coherence and structural change can be quantified by (1) determining how related industries are to one another in terms of their capability requirements. This industry-to-industry relatedness can then be used to calculate (2) how related an industry is to the basket of industries that constitute a region's industry mix. We call this the *regional capability match*, or simply the *match* of an industry to a region. Next, (3) regional coherence is quantified as the average

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.

⁸ 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.

regional match of all industries in the region. Finally, (4) structural change is defined as the match of the current industry mix to the region's past capability base. This procedure is summarized in Table 2 and explained in the following sections in detail.

TABLE 2 DEFINITIONS OF QUANTITIES

Inter-industry relatedness: skill relatedness

Inter-industry relatedness can be measured in several ways (for an overview, see Neffke and Henning, 2013). In the main text, we focus on relatedness in terms of similarities in professional skills or *skill relatedness*. We do this for two reasons. Firstly, the skills embedded in a firm's human capital are among its most valuable resources (Grant 1996; Grant and Spender 1996) and have been shown to condition a firm's diversification path (Porter, 1987; Neffke and Henning, 2013). Secondly, human capital can and is shared between firms in a region. It therewith acts as an important channel of local knowledge exchange and local externalities (Almeida and Kogut, 1999).

Neffke and Henning (2013) show that similarities in human capital requirements can be quantified using information on cross-industry labor flows. The logic behind this is that workers are in general reluctant to switch to jobs where their current skills are not valued. Moreover, firms are less willing to hire workers without relevant work experience. Therefore, industries that demand similar skills typically display large labor flows among them and inter-industry skill relatedness should thus reveal itself in excessive inter-industry labor flows.⁹ Using the simplified index proposed in Neffke *et al.* (2013), we measure the skill relatedness between two industries, i and j , as the ratio of observed to expected worker flows, where expectations are based on overall mobility rates in both industries:

$$SR_{ij} = \frac{F_{ij}}{(F_{i,F_j})/F_{..}} \quad (1)$$

In this equation, 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 $F_{..} = \sum_{i,j} F_{ij}$.¹⁰

⁹ Even though we believe that skill relatedness is a particular relevant type of relatedness, as a robustness check we repeat all analyses for input-output and classification-based relatedness (see Appendix D).

¹⁰ 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

The term $(F_i F_j) / F_{..} = F_i \frac{F_j}{F_{..}}$ 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 that j receives from any industry in the economy. SR_{ij} values greater than one signal that industries are skill related, whereas values between zero and one indicate that industries are unrelated. The resulting skill-relatedness index is highly predictive of corporate diversification (Neffke and Henning, 2013), is 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. However, the degree of relatedness of a given industry to a regional economy is an industry-region relationship. We quantify this relationship as how much employment in the region is related to the industry. The more related employment there is, the stronger the industry's match with the region's capability base is supposed to be. Let E_{irt}^{rel} be all employment in industries related to industry i in region r in year t :¹¹

$$E_{irt}^{rel} = \sum_j E_{jrt} * I(SR_{ij} > 1) \quad (2)$$

where E_{jrt} 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 . That is, the capability match between an industry and a region is derived from the location quotient of related employment:

$$LQ_{irt}^{rel} = \frac{E_{irt}^{rel} / E_{rt}}{E_{i,t}^{rel} / E_{..t}} \quad (3)$$

where E_{rt} is the total employment in the region in year t , $E_{i,t}^{rel}$ the total employment in related industries in the country, and $E_{..t}$ the overall employment in the country. If LQ_{irt}^{rel} is greater than one, the employment share of related industries in the region exceeds their share in the national economy. If

average over the entire period to reduce measurement errors as explained in Neffke *et al.* (2013). The exact procedure is described in Appendix B.

¹¹ Because they add to the regional capability base, related employment includes all employment in non-traded, public sector and natural-resource-based industries. Furthermore, please note that the industry's own employment contributes to the related employment in the region. However, including or excluding an industry's own employment does not substantively alter any of our findings. In our analyses of agent types, we do exclude the *agent's* own employment.

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.¹² This asymmetry complicates calculating averages. We therefore transform LQ_{irt}^{rel} as follows:

$$\widetilde{LQ}_{irt}^{rel} = \frac{LQ_{irt}^{rel}-1}{LQ_{irt}^{rel}+1} \quad (4)$$

$\widetilde{LQ}_{irt}^{rel}$ 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 $LQ_{irt}^{rel} = 2$, $\widetilde{LQ}_{irt}^{rel} = \frac{1}{3}$, whereas $LQ_{irt}^{rel} = \frac{1}{2}$ implies $\widetilde{LQ}_{irt}^{rel} = -\frac{1}{3}$.

Regional coherence and structural change

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

$$C_{rt} = \sum_i \frac{E_{irt}}{E_{rt}} \widetilde{LQ}_{irt}^{rel} \quad (5)$$

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 given region r :

$$C_{rt}^{base} = \sum_i \frac{E_{i,t}}{E_{.t}} \widetilde{LQ}_{irt}^{rel} \quad (6)$$

where $\frac{E_{i,t}}{E_{.t}}$ is industry i 's share in total national employment.

¹² 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 a base year, T :

$$S_{rt,T} = \sum_i \frac{E_{irt}}{E_{.rt}} \widetilde{LQ}_{irT}^{rel}, \text{ where } T < t \quad (7)$$

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 local industries with high capability match values, agents reinforce the focus of that capability base. When agents destroy employment in such industries, the capability base’s focus shifts. Similarly, for local industries with low capability match values, employment creation shifts the capability base and employment destruction deepens it. To study structural change by agent type, we divide all establishments by whether they create or destroy employment. Incumbent establishments are divided into three groups: growing, declining and exiting incumbents. Furthermore, incumbents that switch industries create employment in the industry they enter and destroy employment in the industry they leave.¹³ Therefore, we split industry switchers into two artificial types: “out-switching” incumbents and “in-switching” incumbents. In the new establishments, we distinguish among those that belong to local expanding firms, to non-local expanding firms, to local entrepreneurs and to non-local entrepreneurs. Table 3 provides an overview of all agent types.

TABLE 3 AGENT DEFINITIONS

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, an establishment is said to belong to a *local* firm if in the year prior to the new establishment’s founding, the parent firm employed most of its employees outside the new establishment’s labor market area. However, if the founding of the 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. If the entrepreneur sets up an establishment in the same labor market area as where he or she was employed

¹³ In principle, firms may also move to another region. However, such events are very rare and, therefore, left unexplored.

in the previous year, the entrepreneur is considered local, whereas all others are regarded 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. We drop establishments whose origins we cannot determine.

The structural change an agent type induces in a region is derived from the average capability match of the employment this agent type creates or destroys within a given time period to the region's original economic structure. With T as the base year, we can write the structural change an agent induces until year t as:

$$A_{rt,T}^a = \sum_i \frac{\Delta E_{irt,T}^a}{\Delta E_{rt,T}^a} \widetilde{LQ}_{ir,T}^{rel} \quad (8)$$

where $\frac{\Delta E_{irt,T}^a}{\Delta E_{rt,T}^a}$ is the employment created (or destroyed) by all establishment of type a in industry i and region r between the base year T and the current year t ($\Delta E_{irt,T}^a$) as a share of the total employment created (destroyed) by all of agent a 's establishments in any industry in the region ($\Delta E_{rt,T}^a$). Substantively, $A_{rt,T}^a$ shows how strongly an agent type's new (or destroyed) employment is related to the local economy of year T . To facilitate interpretation, we further subtract the average relatedness of existing industries in year T (i.e., a region's base year coherence):

$$\tilde{A}_{rt,T}^a = A_{rt,T}^a - C_{rT} \quad (9)$$

Positive values of $\tilde{A}_{rt,T}^a$ now indicate that the agent's activities are more related to the region than the region's pre-existing activities, whereas negative values indicate the agent's activities are less related.

5. Results

Diversity and industrial change in Swedish regions

Figure 1 shows how the diversity of Swedish regions has evolved. For each year, it depicts the employment entropy of regions' industry mixes averaged over all regions.

FIGURE 1 DIVERSITY

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 industries¹⁴ in 2010, appeared after 1994 and that 27% of the local industries in 1994 had disappeared by 2010. Moreover, Figure 3 shows that not only is there significant turnover in local industries, the employment composition of regions become increasingly dissimilar to their 1994 compositions.

FIGURE 2 CHURNING

FIGURE 3 COSINE DISTANCE

Coherence and structural change

Figure 4 shows the coherence of regions and how it evolves over time. The average coherence significantly exceeds its proportional employment baseline in every single year, showing that local industries are more closely related to the economic structure of their region than the Swedish economy as a whole. In terms of the theoretical framework in section 2, this finding suggests that the industry composition of a regional economy draws on a relatively narrow set of regional capabilities, in much the same way as a firm's product portfolio is often organized around some core competences. Given the observed industrial change, one would expect the capability base of regions to change as well. However, the average coherence fluctuates only marginally between 0.02 and 0.05, without any statistically significant shifts. Regardless, the economy could have moved away from the 1994 structure in a concerted shift of employment across many industries that kept the internal relatedness among activities constant. The downward-sloping line in Figure 5 shows that to some extent this is indeed happening. The relatedness of a region's industry mix to the baseline economy of 1994 progressively weakens. The slope in Figure 5 is significantly negative at -0.0029 (t-statistic: -3.76). This means that structural change unfolds very slowly. The point estimate implies that it would take the average region over 50 years to move one standard deviation – which is somewhat less than the average distance to the national economy – away from its base year economy.

FIGURE 4 COHESION

FIGURE 5 STRUCTURAL CHANGE

¹⁴ A local industry is defined as a region-industry combination, such as for instance shipbuilding-in-Gothenburg.

Agents of structural change

Table 4 summarizes how many establishments are associated with each agent type. We focus on those new establishments that were created between 1994 and 2000. Taking entrepreneurs and expanding firms together, they account for over 100,000 new jobs, or about 17,000 a year, which is around one fifth of the yearly employment created by growing incumbents. Moreover, establishments of non-local origin tend to create new local industries more often. The same holds for firm-owned establishments (regardless of their origins) compared to entrepreneur-owned establishments. However, some new local industries may be very related to the pre-existing industries in the region. Therefore, in the remainder of this section, we investigate whether agents' differential rates of industrial change translate into differences in structural change.

TABLE 4: EMPLOYMENT / ESTABLISHMENTS / NEW INDUSTRIES

Short-term structural change

Figure 6 summarizes how much structural change is implied in the employment that each agent type creates (or destroys) within a one-year time period.¹⁵ Agent types are listed along the horizontal axis. On the vertical axis, we plot an agent type's average capability match to its region (the agent's $\tilde{A}_{r,t,T}^a$ value), together with a 95% confidence interval. To facilitate interpretation, the figure contains a second vertical axis that maps the $\tilde{A}_{r,t,T}^a$ -axis into percentiles of the overall match distribution of existing local industries in 1994. This secondary (right-most) axis tells us which percentage of the existing 1994 employment is matched to its region at least as weakly as the corresponding match value on the left-most axis. For instance, $\tilde{A} = -0.10$ means that the corresponding agent type on average creates employment in local industries in the 23rd match-percentile, placing it in the bottom quarter of all existing Swedish employment.

Agent types that generate employment are depicted by a green, upward-pointing arrow, those that destroy employment by a red, down-ward pointing arrow. The markers' sizes vary with the total employment that agents represent. Positive values of \tilde{A} indicate that an agent type is generally found in

¹⁵ Incumbents are defined as 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 2000. Next, we record the structural change induced one year after they are founded. That is, for new establishments, we pool $\tilde{A}_{r94,95}^a$, $\tilde{A}_{r95,96}^a$, $\tilde{A}_{r96,97}^a$, $\tilde{A}_{r97,98}^a$, $\tilde{A}_{r98,99}^a$, and $\tilde{A}_{r99,00}^a$.

industries that match the region more strongly than the (employment-weighted) average local industry. If the agents generate employment in such strongly matched industries, the capability base becomes more focused. If they destroy employment in such industries, the focus weakens. Negative values of \tilde{A} correspond to industries with below-average match values. In these industries, the reverse holds: employment creation diversifies the capability base, whereas employment destruction narrows it.

FIGURE 6 AGENTS 1 YR

Figure 6 shows that different agents change the region in different ways. Incumbent establishments tend to reinforce current specializations. If they grow, they do so predominantly in above-averagely matched local industries. If they shrink or close down, on average they reduce employment in below-averagely matched local industries. Moreover, incumbents that switch industries tend to move to industries that fit the region better: on average, they abandon industries in the 40th and enter industries in the 47th match-percentile.

In contrast, new establishments tend to diversify a region's capability base. Indeed, in support of Hypothesis 1 (incumbents induce less structural change than new establishments), almost all new-establishment types display below-average \tilde{A} -values.¹⁶ The only exception is new establishments of expanding local firms, which tend to enter industries that are closely related to the region's industry mix. However, new establishments of local firms simply represent incumbent growth that needs to be accommodated in new facilities. It is therefore not surprising to find that these establishments behave much like growing incumbents.

New establishments belonging to existing firms occupy on average more strongly matched industries (42nd match-percentile), than those of entrepreneurs (29th match-percentile). This supports Hypothesis 3b over Hypothesis 3a: entrepreneurs induce more structural change than expanding firms. However, we already noted that the founding of new establishments by local firms is some ways similar to the growth of incumbent establishments, which tended to reinforce the local economic structure. Non-local existing firms may thus play a very different role in structural change compared to local existing firms.

¹⁶ Our findings are related to those in Dumais *et al.* (2002), who study the degree to which industries are agglomerated in a limited number of regions. Consistent with our findings, that study shows that new establishments have a deagglomerating effect on industries. Furthermore, Dumais and his colleagues find that exits lead to a strengthening of existing agglomeration patterns, which is similar to our finding that exits reinforce existing capability structures. However, we find that the growth and decline patterns of incumbents strengthen existing specializations, whereas Dumais and co-authors find these patterns to weaken existing agglomeration.

We find that this is indeed the case. Regarding the agents' geographical origins, we find that agents from outside the region induce much more structural change than agents from within the region. On average, local expanding firms create employment in the 59th match-percentile, against the 33rd for non-local firms. The difference between local (32nd match-percentile) and non-local entrepreneurs (22nd match percentile) is smaller, but still economically and statistically significant. Therefore, we conclude that there is strong support for Hypothesis 2.

Long-term structural change

Structural change is typically associated with a much longer time horizon than the one year changes depicted in Figure 6. To induce long-lasting structural change, it is not enough for establishments in unrelated industries to get started. To change the region's productive structure, these establishments need to survive and grow. Therefore, Figure 7 repeats the analyses of Figure 6, but now over a time period of 10 years. Again, the 1994 cumulative distribution of match values is provided on the right-most vertical axis. Furthermore, the 1-year match values of Figure 6 have been retained as a reference and can be recognized by their dotted confidence intervals.

FIGURE 7 AGENTS 10 YR

Long-term structural-change patterns are similar to short-term ones. On a ten-year horizon, incumbents reinforce a region's focus at best weakly, whereas new employment in unrelated industries is mostly created by new establishments.

Differences between the 10-year and the 1-year match figures arise from differences in long-term survival and growth rates at different points in the match distribution. For instance, the fact that match values for out-switching and exiting incumbents shift up (*i.e.*, occur at higher match values) suggests that unrelated activities are abandoned in relatively early stages through establishment closures and adjustments in industry orientation. Indeed, this in turn implies that incumbents in high-match local industries outperform incumbents in low-match industries and matches previous work showing that firms benefit from nearby related economic activity (Delgado *et al.*, 2010; Neffke *et al.*, 2012).

Turning to new establishments, entrepreneur-owned establishments' \tilde{A} -values also shifted upward. This means that growth and/or survival also for these establishments has been overly concentrated in higher-matched industries. In contrast, the new establishments of existing firms either remained at the same match-value (local firms) or even moved down (non-local firms). Apparently, establishments of

non-local firms in low-match industries grew more and/or survived longer than those in high-match industries. The result of these opposite survival dynamics and/or growth dynamics is that new establishments of (non-local) expanding firms end up in match values below those of entrepreneurs, although confidence intervals of non-local entrepreneurs and firms still partially overlap. This means that non-local firms draw equal with or even surpass entrepreneurs as the main agents of structural change.

Plant survival

The long-term structural change analyses suggest that entrepreneurs are less likely to survive in regional economies with few related activities than the establishments of larger firms. To assess this more carefully, we investigate whether the presence of related industries is associated with higher establishment survival rates. For each new establishment between 1995 and 2000, we create a dummy that is valued at one if it survives for at least 10 years and at zero otherwise. Next, we estimate linear probability models, that is we regress this dummy variable on a set of founder type dummies and their interactions with the natural logarithm of related employment in the region.¹⁷ We include entry-year, region and industry dummies in the regressions to isolate the effect of the founder type. However, we do not control for any other establishment characteristics, such as start-up size, because we are interested in *how* survival rates differ by agent type, not necessarily in *why* they do so. Table 5 summarizes the results.

TABLE 5: SURVIVAL

The unconditional average survival rate for new establishments is 0.201. The model in Column (1) of Table 5 contains only a dummy for whether or not an establishment's founder comes from outside the region. The negative coefficient on this dummy shows that establishments of non-local founders have a 2.1 percentage-point lower survival rate than those of local founders. Column (2) adds interactions with the amount of related employment in the region. Plants that enter regions with a large amount of related employment tend to survive longer, especially if their founders are local.¹⁸ The estimated

¹⁷ In this regression-based set-up, effect sizes of log-transformed employment figures are easier to interpret than those of our match variable, while industry and region effects can still be absorbed using adequate sets of dummies.

¹⁸ A t-test reveals that the lower point estimates for the interaction of related employment with the non-local-founder dummy compared to the one with the local-founder dummy is statistically significant with a p-value of 0.026.

coefficients suggest that a doubling in related employment translates into a 1.2 percentage-point increase in survival rates for local establishments and a 0.9 percentage-point increase for non-local establishments.¹⁹ When we split founders into entrepreneurs and existing firms (Columns (3) and (4)), even larger differences emerge. Whereas, consistent with the findings in Balasubramanian (2011), firm-owned establishments generally have higher survival rates than entrepreneur-owned establishments, only the entrepreneur-owned establishments have significantly higher survival rates in regions with more related employment. Column (5) further subdivides establishments by their geographical origin. Regardless of whether an establishment was founded by local or non-local entrepreneurs, survival rates of entrepreneur-owned establishments are always lower than those of firm-owned establishments. However, whereas local roots are associated with higher survival rates among entrepreneurs (the omitted category consists of local entrepreneurs), the opposite holds for firm-owned establishments: here, non-local origins are associated with *higher* survival rates. Furthermore, Column (6) suggests that related employment in the region only matters for entrepreneur-owned, not for firm-owned establishments.²⁰

These findings are consistent with the theoretical framework of section 2. Firstly, the finding depicted in Column (4) – that only entrepreneur-owned establishments display significantly higher survival rates in regions with related employment – is in line with the idea that entrepreneurs depend more strongly on local capabilities than establishments that are part of larger firms. Secondly, the hypothesis that entrepreneurs cannot draw on a parent firm to compensate for the lack of access to local capabilities outsiders typically have would explain why, in Column (5), we find higher failure rates for non-local entrepreneurs but not for non-local firms. However, such a causal interpretation is hazardous, because the decision to enter a region is endogenous, even conditional on industry and region fixed effects. For instance, the fact that firm-owned establishments seem unaffected by the local amount of related employment could alternatively mean that they are more careful when choosing a location. In that case, the absence of an association with higher survival rates is due to the fact that firms make fewer mistakes (or take less risk) when deciding where to locate, not because they draw fewer benefits from the local environment.

¹⁹ The effect size of raising related employment by a factor ξ is calculated as: point estimate $\times \ln(\xi)$.

²⁰ The observed differences between the interactions with $\ln(\text{rel. emp.})$ of local and non-local entrepreneurs are statistically insignificant.

Diffusion of capabilities

The finding in Figures 6 and 7 that non-local agents widen the capability base of a region has an important interpretation when it comes to questions of how economic activities diffuse across regions, namely that non-local agents 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 6 shows that this is indeed the case: for both entrepreneurs and even more so for firms from outside the region the capability match to their home regions is much higher than to their host regions. This implies that the relocation of firms and entrepreneurs is an important vehicle through which capabilities spread out across regions.

TABLE 6 DIFFUSION

6. Conclusion

Summary and robustness

There are many parallels between the RBVs depiction of firm growth and the way in which regional economies develop. In both, firms and regions, growth typically does not only involve enlarging the scale but also the scope of production. Moreover, for both firms and regions, this expansion of scope is typically achieved through related diversification. We can understand these parallels if we think of regions as endowed with capability bases. Although regional capability bases are specific to the economic activities that use them, their use is typically not limited to single products. However, unlike firms, which can exclude others from using their resources, access to regional capabilities is less restricted.

Economic agents will differ in their reliance on and access to these regional capabilities. Establishments of larger firms can substitute their parent firms' resources for regional capabilities, whereas entrepreneurial ventures cannot. Furthermore, local firms and entrepreneurs are often better positioned to access local capabilities than their counterparts from outside the region.

We have tested the implications of this framework using micro data for the Swedish economy between 1994 and 2010. The results of this exercise can be summarized as follows:

- 1) structural change has unfolded slowly in Sweden and occurred in such a way that regions have maintained the coherence of their industry mix;
- 2) existing establishments tend to deepen a region's capability base by destroying employment in unrelated industries and creating employment in related ones, whereas most new establishments create employment in unrelated industries, thereby shifting the region's focus;
- 3) entrepreneur-owned establishments induce more structural change in the short run than in the long run, whereas the reverse holds for the new establishments of existing firms;
- 4) consistent with finding 3), whereas entrepreneur-owned establishments tend to survive longer in regions with more related employment, no such association is found for firm-owned establishments;
- 5) moreover, being local is associated with higher survival rates for entrepreneurs, whereas the opposite holds for firm-owned establishments;
- 6) non-local agents induce significantly more structural change than agents from within the region;
- 7) what is more, non-local agents' home regions typically have more related employment than their new host regions.

Although these results rely on skill relatedness to measure the amount of related employment in a region, other relatedness indicators yield similar outcomes. In Appendix D, we replicate findings 1) to 7) by constructing Figures 4, 5, and 7 and Tables 5 and 6 using an input-output-based measure of relatedness and a relatedness index based on the industry classification system.

Our findings shed light on the interplay between firms and their local environments. For instance, economic geographers often put forward entrepreneurs as sources of industrial renewal (Appold 2000; Boschma 2004; Feldman et al. 2005). Our findings support this view, but also suggest that it needs to be qualified. In comparison to the new establishments of larger firms, entrepreneurs are more likely to introduce activities unrelated to a region's existing industry mix, however findings 3 and 4 suggest they are more likely to fail when doing so. Furthermore, unrelated diversification is mostly driven by firms and entrepreneurs from outside the region (finding 6). As a consequence, non-local agents induce more structural change than local agents. Moreover, the home regions of these new-comers are often better endowed with related industries than their new host regions are (finding 7). This suggests that the mobility of firms and entrepreneurs is an important channel through which industries and their capabilities diffuse.

Caveats

In our study, we only investigate the sources of structural change, not the consequences. What those consequences are depends on, among other things, the time horizon. In the short run, it is often beneficial to leverage existing capabilities but in the long run, regions will have to adapt to new economic realities. However, long-run structural change can be accomplished through a series of small steps, in a process of related diversification that moves the region away from its traditional capability base. What the exact balance and speed of related versus unrelated diversification should be remains a topic for further research.

Related to this is the fact that, by focusing on the new establishments that enter an economy, we have mostly highlighted the diversification aspect of structural change. However, although our analyses show that incumbent exit and decline typically take place in unrelated industries, there are well-known examples in which the core industries of a region decline (e.g., Grabher, 1994). In these cases, structural change occurs because of the loss of a central industry. It is therefore important to note that structural change may occur through the collapse of a region's core industries, in which case it signals the loss of important local capabilities.

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 more prevalent than others. For instance, entrepreneurs set up far more establishments than existing firms do: the new establishments of local entrepreneurs outnumber those of non-local entrepreneurs 5-to-1 and those of non-local firms 20-to-1. Therefore, although the intensity with which they shift a region's capability base to new areas is lower, as a group, local entrepreneurs constitute an important factor in this shift.

Furthermore, we quantified the capability match as the overrepresentation of related employment in the region. However, alternative approaches exist, such as 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. We leave exploring such alternatives to future research.

Finally, our analyses answer the question of who introduces unrelated economic activities in a region. In essence, this question is descriptive, not causal. We therefore remain agnostic about whether the reported differences among agent types reflect different intrinsic capacities for structural change or, for instance, differences in location choices. Similarly, in the survival analyses, we cannot distinguish spatial

sorting of high quality establishments from agglomeration externalities, an issue that has attracted considerable attention in urban economics (e.g., Combes et al., 2008).

Implications and future research

We have differentiated the industries a region currently hosts from the capabilities that allow these industries to thrive. Indeed, Figures 1-5 are taken as evidence that, although the industry mix of a region may fluctuate strongly, its capability base changes much more slowly. This is an important finding for local policy-making, because it highlights that the current constellation of industries in a region is just one manifestation of how the local capabilities can be put to work. Local policy makers could use the tools suggested in this paper to shift their focus from the industries that *are* present in the region to those that *could* be present. More fundamentally, the 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. Similarly, the capability match values we calculated could help managers and entrepreneurs identify regions with a suitable capability base for their new activities.

Our study also raises a number of new questions. Firstly, in finding that new subsidiaries of existing firms are better able to grow and survive in unrelated environments than stand-alone establishments, we come to the question of why this is the case. Our proposal – that firm-owned establishments draw on their parent firms' capabilities – remains to be proven, as do the questions of how and across what distance multi-establishment firms can accomplish this. Secondly, the fact that firms switch affiliations from low-match to high-match industries in the region suggests that firm strategies interact with regionally available capabilities in ways that are still poorly understood. We hope that the analyses in this paper will prove useful in approaching these and other questions on how regional economies and their capability bases co-evolve with the firms they host.

REFERENCES

- Alcácer J, Chung W. 2007. Location strategies and knowledge spillovers. *Management Science* **54**(5): 760-776.
- Almeida P, Kogut B (1999) Localization of knowledge and the mobility of engineers. *Management Science* **45**(7): 905–917
- Andersson J, Arvidson G. 2006. Företagens och arbetsställes dynamik (FAD). Memo, Statistics Sweden.
- Appold SA. 2000. The control of high-skill labor and entrepreneurship in the early US semiconductor industry, *Environment and Planning, A* **32**, 2133–2160.
- Balasubramanian N. 2011. New Plant Venture Performance Differences Among Incumbent, Diversifying, and Entrepreneurial Firms: The Impact of Industry Learning Intensity. *Management Science* **57**(3): 549-565.
- Barney J. 1991. Firm resources and sustained competitive advantage. *Journal of Management* **17**: 99–120.
- Bathelt H, Malmberg A, Maskell P. 2004. Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography* **28**: 31-56.
- Boschma RA. 2004. Competitiveness of regions from an evolutionary perspective, *Regional Studies* **38**(9): 1001-1014.
- Boschma R, Frenken K. 2011. Technological relatedness and regional branching, in: H. Bathelt, M.P. Feldman and D.F. Kogler (eds.), *Beyond Territory. Dynamic Geographies of Knowledge Creation, Diffusion and Innovation*, Routledge, London and New York, pp. 64-81.
- Boschma RA, Minondo A, Navarro M. 2013. The Emergence of New Industries at the Regional Level in Spain: A Proximity Approach Based on Product Relatedness. *Economic Geography* **89**(1): 29-51.
- Bryce DJ, Winter SG. 2009. A general inter-industry relatedness index. *Management Science* **55**: 1570-1585.
- Cooke P.H., Morgan K. (1998). *The Associational Economy. Firms, Regions, and Innovation*. Oxford: Oxford University Press.
- Combes P-Ph, Duranton D, Gobillon L. 2008. Spatial wage disparities: Sorting matters! *Journal of Urban Economics* **63**(2): 723-742.
- Cramer JS, Hartog J, Jonker N, Van Praag CM. 2002. Low risk aversion encourages the choice for entrepreneurship: an empirical test of a truism. *Journal of Economic Behavior & Organization* **48**(1): 29–36.
- Dahl MS, Sorenson O. 2012. Home sweet home: Entrepreneurs' location choices and the performance of their ventures, *Management Science* **58**(6): 1059–1071.
- Dauth, W. (2010). The mysteries of the trade: employment effects of urban interindustry spillovers. *IAB Discussion Paper 15/2010*: 1-25.
- Delgado M, Porter ME, Stern S. 2010. Clusters and entrepreneurship. *Journal of Economic Geography* **10**(4): 495-518.
- Delgado M, Porter ME, Stern S. 2013. Defining clusters of related industries, HBS working paper.
- Dumais G, Ellison G, Glaeser EL. 2002. Geographic concentration as a dynamic process. *Review of Economics and Statistics*, **84**(2): 193-204.
- Duranton G, Puga D. 2004. Micro-foundation of urban agglomeration economies, in Henderson J. and Thisse J.-F. (Eds) *Handbook of Urban and Regional Economics*, vol 4, pp. 2065-2118. Elsevier, Amsterdam.
- Eisenhardt KM, Martin JA. 2000. Dynamic capabilities: What are they? *Strategic Management Journal* **21**: 1105–1121.
- Ellison G., Glaeser E. L., Kerr W. R. (2010) What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review*, **100**(3): 1195–1213.

- Essletzbichler J. 2013. Relatedness, Industrial Branching and Technological Cohesion in US Metropolitan Areas. *Regional Studies*, 10.1080/00343404.2013.806793
- Faggian A. McCann Ph. (2006). Human capital flows and regional knowledge assets: a simultaneous equation approach. *Oxford Economic Papers* 52: 475–500.
- Farjoun M. 1994. Beyond industry boundaries: Human expertise, diversification and resource-related industry groups. *Organization Science* 5(2): 185–199.
- Feldman, M., J. Francis, & J. Bercovitz. 2005. Creating a Cluster While Building a Firm: Entrepreneurs and the Formation of Industrial Clusters. *Regional Studies*, 39(1): 129-141.
- Florida R, Mellander C, Stolarick K. 2012. Geographies of scope: an empirical analysis of entertainment, 1970–2000. *Journal of Economic Geography*, 12(1): 183-204.
- Frenken K, Boschma RA. 2007. A theoretical framework for economic geography: industrial dynamics and urban growth as a branching process. *Journal of Economic Geography* 7 (5), pp. 635-649.
- Gertler, M.S. 2003. Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there), *Journal of Economic Geography*, 3: 75-99.
- Ghemawat P. 1986. Sustainable advantage, *Harvard Business Review*, 64(5): 53–58.
- Giuliani E. 2007. The selective nature of knowledge networks in clusters: evidence from the wine industry, *Journal of Economic Geography* 7: 139-168.
- Glaeser EL. 2005. Reinventing Boston: 1630–2003, *Journal of Economic Geography* 5(2): 119-153.
- Glaeser, E.L., Kallal, H.D., Scheinkman, J.A., Schleifer A. (1992) Growth in Cities, *Journal of Political Economy*, 100 (6): 1126-1152.
- Grabher, G. (1993) The weakness of strong ties: the lock-in of regional development in the Ruhr area, in Grabher, G. (ed.) *The Embedded Firm*. London: Routledge.
- Grant RM. 1996. Toward a knowledge-based theory of the firm. *Strategic Management Journal* 17: 109–122.
- Grant RM, Spender J-C. 1996. Knowledge and the firm: overview. *Strategic Management Journal* 17: 5–9.
- Helfat CE, Peteraf MA. 2003. The dynamic resource-based view: capability lifecycles, *Strategic Management Journal*, 24(10): 997–1010.
- Henderson, J.V., Kuncoro, A., Turner, M. (1995) Industrial development in cities, *Journal of Political Economy*, 103:1067-1085.
- Henderson R, Cockburn I. 1994. Measuring Competence? Exploring Firm Effects in Pharmaceutical Research, *Strategic Management Journal* 15: 63–84.
- Hidalgo CA, Klinger B, Barabási A-L, Hausmann R. 2007. The Product Space Conditions the Development of Nations. *Science* 317: 482-487.
- Jacobs J. 1969. *The Economy of Cities*. New York: Vintage Books.
- Lawson C. 1999. Towards a competence theory of the region, *Cambridge Journal of Economics* 23(2): 151-166.
- Lawson C, Lorenz E. 1999. Collective learning, tacit knowledge and regional innovative capacity, *Regional Studies* 33: 305-317.
- Maskell P. 2005. Towards a Knowledge-Based Theory of the Geographical Cluster. In: Breschi S, Malerba F (Eds.). *Clusters, Networks, and Innovation*. Oxford: Oxford University Press. Chapter 15: 411-432.
- Maskell P, Malmberg A. 1999. Localised learning and industrial competitiveness, *Cambridge Journal of Economics* 23(2): 167-185.
- McCann Ph., Simonen J. (2005). Innovation, knowledge spillovers and local labour markets. *Papers in Regional Science* 84(3): 465-485.
- Montgomery CA, Wernerfelt B. 1988. Diversification, Ricardian Rents, and Tobin's q, *RAND Journal of Economics* 19(4): 623-632.

- Muneepeerakul R, Lobo J, Shuttters ST, Gómez-Liévano A, Qubbaj MR. 2013. Urban Economies and Occupation Space: Can They Get “There” from “Here”? *PLoS ONE* **8**(9): e73676.
- Neffke FMH, Henning M. 2013. Skill relatedness and firm diversification. *Strategic Management Journal*. **34**(3): 297-316.
- Neffke FMH, Henning M, Boschma RA. 2012. The impact of aging and technological relatedness on agglomeration externalities: a survival analysis. *Journal of Economic Geography* **12**(2): 485-517.
- Neffke FMH, Henning M, Boschma RA. 2011. How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. *Economic Geography* **87**: 237-265.
- Neffke FMH, Hidalgo C, Otto A, Weyh A. 2013. Inter-industry labor flows. mimeo.
- Palich LE, Cardinal LB, Miller C. 2000. Curvilinearity in the diversification-performance linkage: An examination of over three decades of research. *Strategic Management Journal* **21**: 155–174.
- Pendall R, Foster KA, Cowell M. 2010. Resilience and regions: building understanding of the metaphor, *Cambridge Journal of Regions, Economy and Society*: 3, 71–84
- Penrose ET. 1959. *The Theory of the Growth of the Firm*. Blackwell: Oxford, UK.
- Peteraf MA. 1993. The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal* **14**: 179–191.
- Porter ME. 1987. From competitive advantage to corporate strategy. *Harvard Business Review* **65**(3): 43–59.
- Porter ME. 1990. *The Competitive Advantage of Nations*. Macmillan, London.
- Porter, M.E. (1998). Clusters and the new economics of competition. *Harvard Business Review* 76 (November-December): 77-90.
- Porter ME. 2000. Location, Competition, and Economic Development: Local Clusters in a Global Economy, *Economic Development Quarterly* **14**: 15-34.
- Porter M. 2003. The Economic Performance of Regions. *Regional Studies* 37(6-7): 549–578.
- Pouder R, St. John CH. 1996. Hot Spots and Blind Spots: Geographical Clusters of Firms and Innovation, *Academy of Management Review* 21(4): 1192-1225.
- Rosenthal S.S., Strange W.C. (2004). Evidence on the Nature and Sources of Agglomeration Economies. In: *Handbook of Urban and Regional Economics* (4). Henderson J., Thisse J.-F. (eds.). Amsterdam: Elsevier.
- Saxenian A. 1994. *Regional advantage: Culture and competition in Silicon Valley and Route 128*, Cambridge MA: Harvard University Press.
- Schumpeter, J. A. (1942) *Capitalism, Socialism and Democracy*. New York: Harper.
- Sorenson O, Audia PG. 2000. The social structure of entrepreneurial activity: Geographic concentration of footwear production in the United States, 1940–1989, *American Journal of Sociology* **106**(2): 424-462.
- Storper M. 1995. The resurgence of regional economies, ten years later: The region as a nexus of untraded interdependencies, *European Urban and Regional Studies* **2**: 191-221.
- Storper M, Venables T. 2004. Buzz: face-to-face contact and the urban economy *Journal of Economic Geography*, 4 (4). 351-370.
- Teece DJ. 1982. Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior and Organization* **3**: 39–63.
- Teece DJ, Pisano G, Shuen A. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* **18**: 509–533.
- Teece DJ, Rumelt R, Dosi G, Winter S. 1994. Understanding corporate coherence. Theory and evidence. *Journal of Economic Behavior and Organization* **23**: 1-30.
- Tushman ML, Anderson Ph. 1986. Technological discontinuities and organizational environments, *Administrative Science Quarterly* **31**(3): 439-465.
- Wernerfelt B. 1984. A resource-based view of the firm. *Strategic Management Journal* **5**: 171–180.

Zhao H, Seibert SE. 2006. The Big Five personality dimensions and entrepreneurial status: A meta-analytical review. *Journal of Applied Psychology*, 91(2): 259-271.

Tables and figures

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

	Static	Dynamic
Industries	<p><i>Diversity</i></p> <p>Measured by: entropy</p> <p>Underlying question: How many different industries are there and how equal is their size distribution?</p>	<p><i>Industrial change</i></p> <p>Measured by: cosine distance</p> <p>Underlying question: How fast are new industries introduced and how much does the size distribution of activities change?</p>
Capabilities	<p><i>Coherence</i></p> <p>Measured by: see section 3</p> <p>Underlying question: How similar are the capabilities required by the various industries in the region? That is, how related are the industries in a region to one another?</p>	<p><i>Structural change</i></p> <p>Measured by: see section 3</p> <p>Underlying question: To what extent does the capability base change due to changes in the region's industries? That is, how related are current industries to the industry mix in the base year?</p>

Table 2: Definitions and relationships among quantities

quantity	unit of analysis	definition	description	normalization	range
labor low	industry-industry	F_{ij}	How many people change jobs from industry i to j ?		$[0, \infty)$
skill relatedness	industry-industry	$SR_{ij} = \frac{F_{ij}}{F_i F_j}$	How related are two industries to one another?		$[0, \infty)$
employment	industry-region	E_{irt}	How many workers does industry i employ in region r in year t ?		$[0, \infty)$
related employment	industry-region	$E_{irt}^{rel} = \sum_j E_{jrt} I(SR_{ij} > 1)$	How much related employment to industry i exists in region r in year t ?		$[0, \infty)$
capability match	industry-region	$LQ_{irt}^{rel} = \frac{E_{irt}^{rel}/E_{rt}}{E_{i,t}^{rel}/E_{.t}}$	How overrepresented are related industries in the region?	$\widetilde{LQ}_{irt}^{rel} = \frac{LQ_{irt}^{rel}-1}{LQ_{irt}^{rel}+1}$	$[0, \infty)$ norm.: $[-1,1)$
coherence	region	$C_{rt} = \sum_i \frac{E_{irt}}{E_{.rt}} \widetilde{LQ}_{irt}^{rel}$	How related are a region's industries on average to the regional economy as a whole?		
coherence baseline	region	$C_{rt}^{base} = \sum_i \frac{E_{i,t}}{E_{.t}} \widetilde{LQ}_{irt}^{rel}$	How related are the industries in the national economy to the regional economy?		$[-1,1)$
structural change	region	$S_{rt,T} = \sum_i \frac{E_{irt}}{E_{.rt}} \widetilde{LQ}_{irt}^{rel}$	How related are a region's current activities to the region's industry mix of year T ?		$[-1,1)$
structural change by agent type	agent-region	$A_{rt,T}^a = \sum_i \frac{\Delta E_{irt,T}^a}{\Delta E_{.rt,T}^a} \widetilde{LQ}_{irt}^{rel}$	How related are the industries in which a given agent type creates or destroys employment to the region's industry mix of year T ??	$\tilde{A}_{rt,T}^a = A_{rt,T}^a - C_{rT}$	$[-1,1)$ norm.: $(-2,2)$

Table 3: Agent types

Agent type	Description	ΔE	Effect on capability base if match is	
			below average	above average
incumbent establishments	<i>existing establishments that ...</i>			
growing	expand their workforce	+	diversify	specialize
shrinking	reduce their workforce	-	specialize	diversify
closing	close down	-	specialize	diversify
industry switchers	<i>existing establishments that ...</i>			
into the industry	switch into the industry	+	diversify	specialize
out of the industry	switch out of the industry	-	specialize	diversify
New establishments	<i>new establishments set up by ...</i>			
local expanding firms	pre-existing firm with main employment concentration inside the region	+	diversify	specialize
non-local expanding firms	pre-existing firm with main employment concentration in another region	+	diversify	specialize
local entrepreneurs	new firm created by entrepreneur from inside the region	+	diversify	specialize
non-local entrepreneurs	new firm created by entrepreneur from outside the region	+	diversify	specialize

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 region (*i.e.*, if the match is below or over the region's coherence).

Table 4: Agent types: employment, number of establishments and new local industries

Agent type	# establishments			employment		% creating new local industries
	entry yr	after 1 yr	after 10 yrs	after 1 yr	after 10 yrs	
Growth, decline and exit						
Incumbent growth		17,507	9,933	75,851	122,359	
Incumbent decline		12,494	8,031	46,577	77,776	
Incumbent exit		10,420	45,268	29,794	270,030	
Industry switching						
Entered industry		1,708	3,643	32,629	107,652	
Exited industry		1,708	3,643	30,812	93,492	
New establishments						
All expanding firms	2,249	1,809	666	38,419	21,449	4.09%
All entrepreneurs	51,806	35,307	10,206	63,166	37,992	2.38%
Local expanding firms	557	435	152	13,263	7,562	1.97%
Non-local expanding firms	1,692	1,374	514	25,156	13,887	4.79%
Local entrepreneurs	42,993	29,617	8,644	53,741	32,798	2.01%
Non-local entrepreneurs	8,813	5,690	1,562	9,425	5,194	4.20%

Table 5: Establishments' 10-year survival rates

Dep. var.: ≥ 10 yr survival (0/1)	(1)	(2)	(3)	(4)	(5)	(6)
Non-local agent	-0.021*** (0.004)	0.02 (0.018)				
Firm agent			0.073*** (0.009)	0.209*** (0.047)		
Local firm					0.041** (0.016)	0.148 (0.094)
Non-local firm					0.074*** (0.01)	0.191*** (0.056)
Non-local entrepreneur					-0.037*** (0.004)	-0.009 (0.019)
<hr/>						
Local X ln(rel. emp.)		0.017*** (0.005)				
Non-local X ln(rel. emp.)		0.013*** (0.005)				
Firm X ln(rel. emp.)				0.004 (0.007)		
Entrepreneur X ln(rel. emp.)				0.017*** (0.005)		
Local Firm X ln(rel. emp.)						0.006 (0.01)
Non-local firm X ln(rel. emp.)						0.004 (0.007)
Local entrepr. X ln(rel. emp.)						0.017*** (0.005)
Non-local entrepr. X ln(rel. emp.)						0.014*** (0.005)
Constant	0.157*** (0.005)	-0.053 (0.057)	0.153*** (0.005)	-0.057 (0.057)	0.156*** (0.005)	-0.042 (0.057)
Entry-year FEs	yes	yes	yes	yes	yes	yes
Industry FEs	yes	yes	yes	yes	yes	yes
Region FEs	yes	yes	yes	yes	yes	yes
<hr/>						
Number of observations	54,055	54,055	54,055	54,055	54,055	54,055
R ²	0.0254	0.0257	0.0265	0.0270	0.0280	0.0284

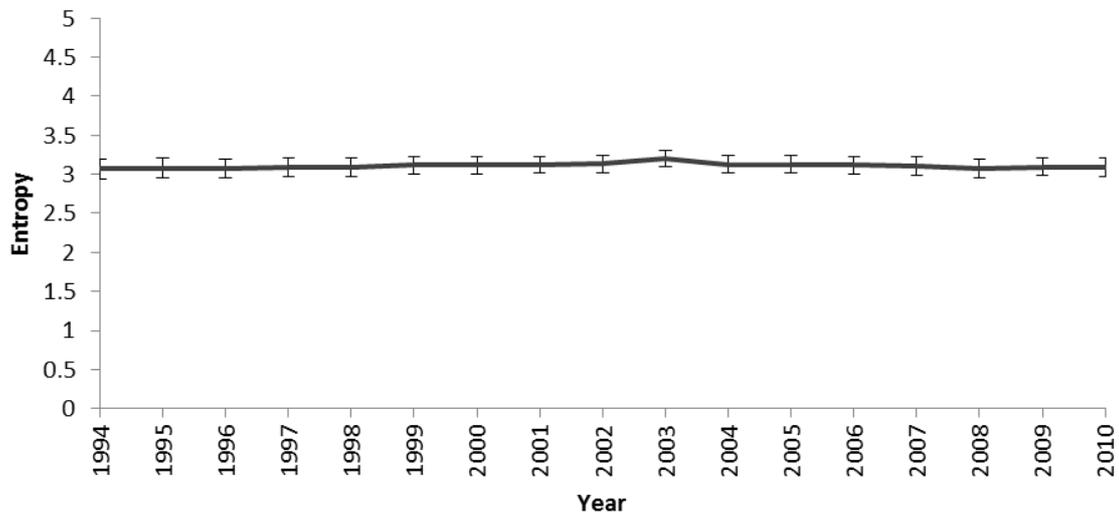
The reported outcomes are of linear probability models of 10-year survival rates for new establishments that enter the Swedish economy in traded, private sector, non-resource based industries between 1994 and 2000. Robust standard errors are shown in parenthesis below the point estimates. Stars indicate significance levels: ***: p<.01, **: p<.05, *: p<.10. The unconditional average 10-year survival rate for the establishments in the sample is 0.201.

Table 6: Knowledge diffusion by outside agents

Agent type	Capability match to:		p-value
	home region	host region	
Non-local expanding firms	0.072 (0.004)	-0.019 (0.004)	0.000
Non-local entrepreneurs	0.001 (0.002)	-0.019 (0.002)	0.000

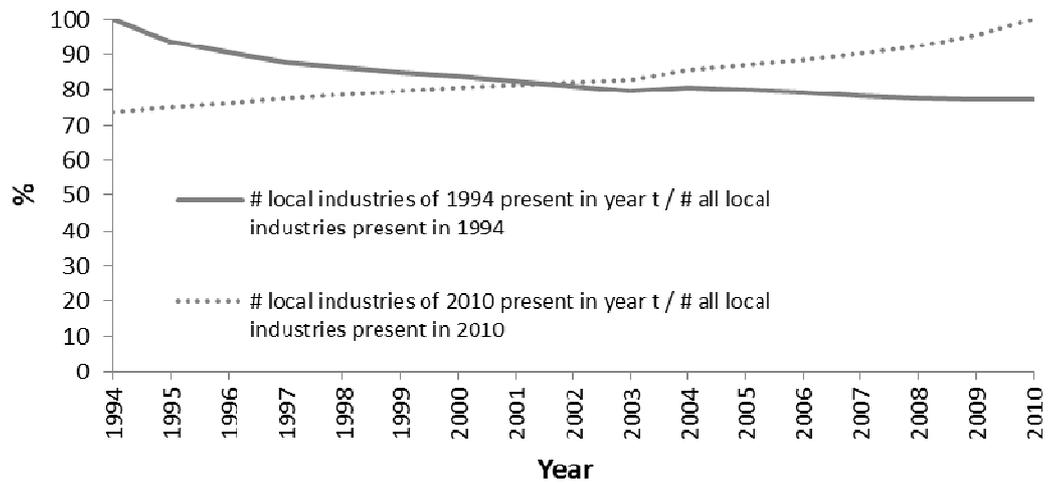
Average capability match of a non-local agent to home and host region (standard error in parentheses). The home region is defined as the region in which the new establishment's parent firm employed most of its workers (non-local firms) or as the region in which the new establishment's entrepreneur was employed in the year prior to opening up the new establishment (non-local entrepreneurs). p-value refers to a test of equal means for the agent's capability match to the home versus to the host region.

Figure 1: Average entropy of the employment composition of labor market regions



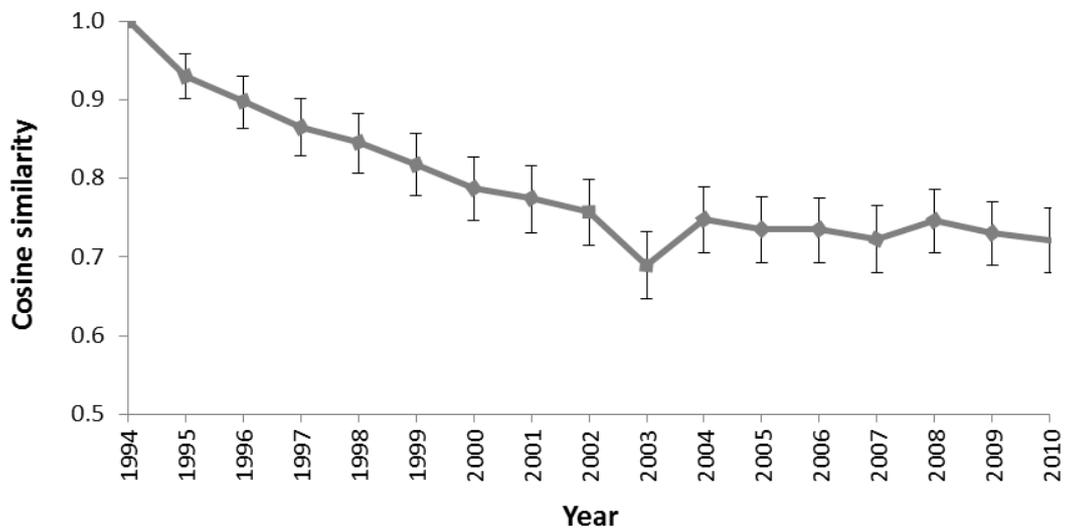
The figure graphs the development of the average employment entropy of Swedish regions over time. Employment entropy is a measure of how diversified a local economy is and is calculated as $entropy_{rt} = -\sum_{i=1}^N \frac{E_{irt}}{E_{rt}} \ln \frac{E_{irt}}{E_{rt}}$, where E_{irt} denotes the employment 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 local industries have equal employment shares. The error bars depict a 95% confidence interval calculated as ± 1.96 times the standard deviation of the entropy's mean across regions.

Figure 2: Turnover of local industries



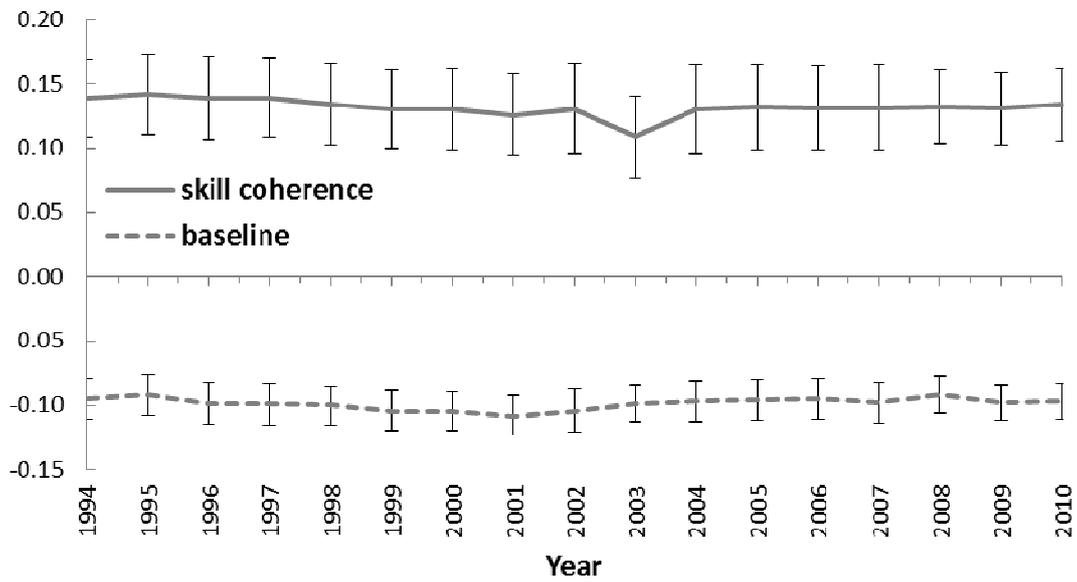
The solid blue line depicts the share of local industries (region-industry combinations) existing (*i.e.*, with non-zero employment) in Sweden in 1994, that survived to at least year $1994 + t$. The dotted red line depicts the share of local industries existing in 2010 that had existed already in year $2010 - t$.

Figure 3: Average cosine similarity to the base year 1994 of regional employment profiles



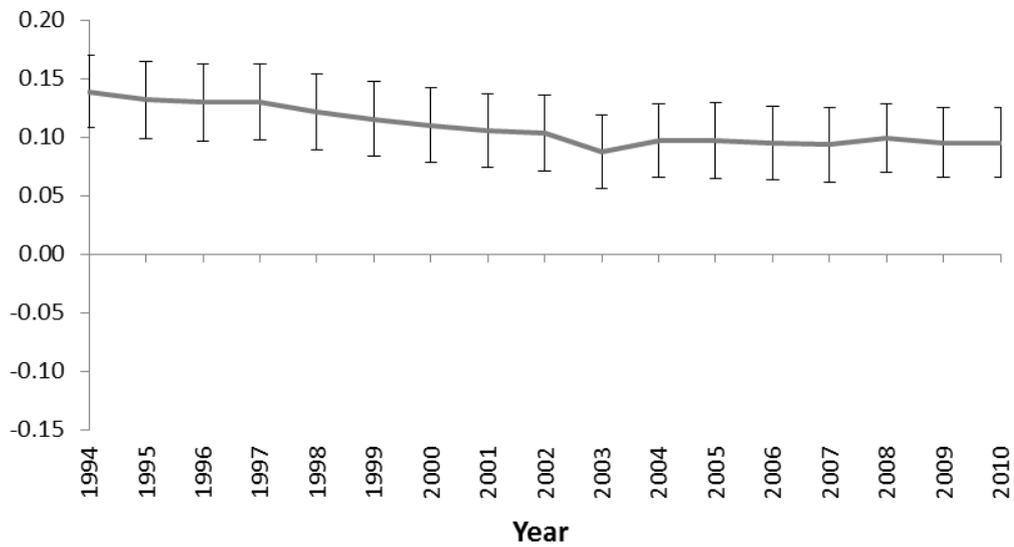
The graph depicts the development of the average cosine similarity between a region's current industrial employment mix and the industrial employment mix of the base year 1994. The cosine similarity measures the similarity of two vectors, in this case, the region's employment profile at two different points in time: $cos_sim_{r,t,T} = \frac{e_{rt} \cdot e_{rT}}{\sqrt{|e_{rt}| |e_{rT}|}}$, where $e_{rt} = (E_{1rt} \dots E_{Nrt})'$ a vector whose elements 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 standard deviation of the mean of the cosine similarity across regions.

Figure 4: Coherence of labor market regions' skill bases



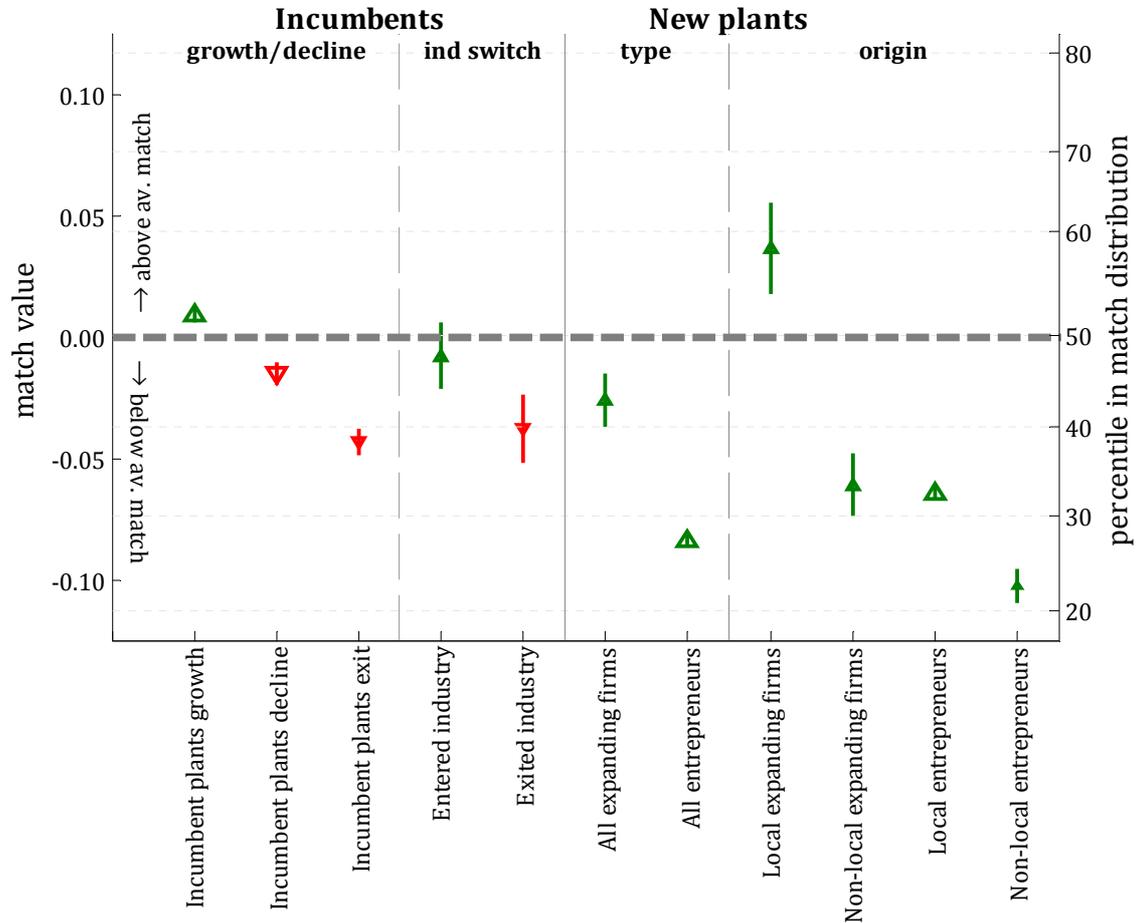
The upper line depicts the development of the average coherence of a regions' capability base. It is measured by its local industries' employment-weighted average capability match to the regional economy as a whole ($C_{rt} = \sum_i \frac{E_{irt}}{E_{rt}} \widetilde{LQ}_{irt}^{rel}$). As a baseline, the lower line depicts the development of the average capability match of Sweden's aggregate, national industries to the region's capability base ($C_{rt}^{base} = \sum_i \frac{E_{it}}{E_t} \widetilde{LQ}_{irt}^{rel}$). The error bars depict a 95% confidence interval calculated as ± 1.96 times standard deviation of the mean across regions.

Figure 5: Structural change in Sweden's labor market regions



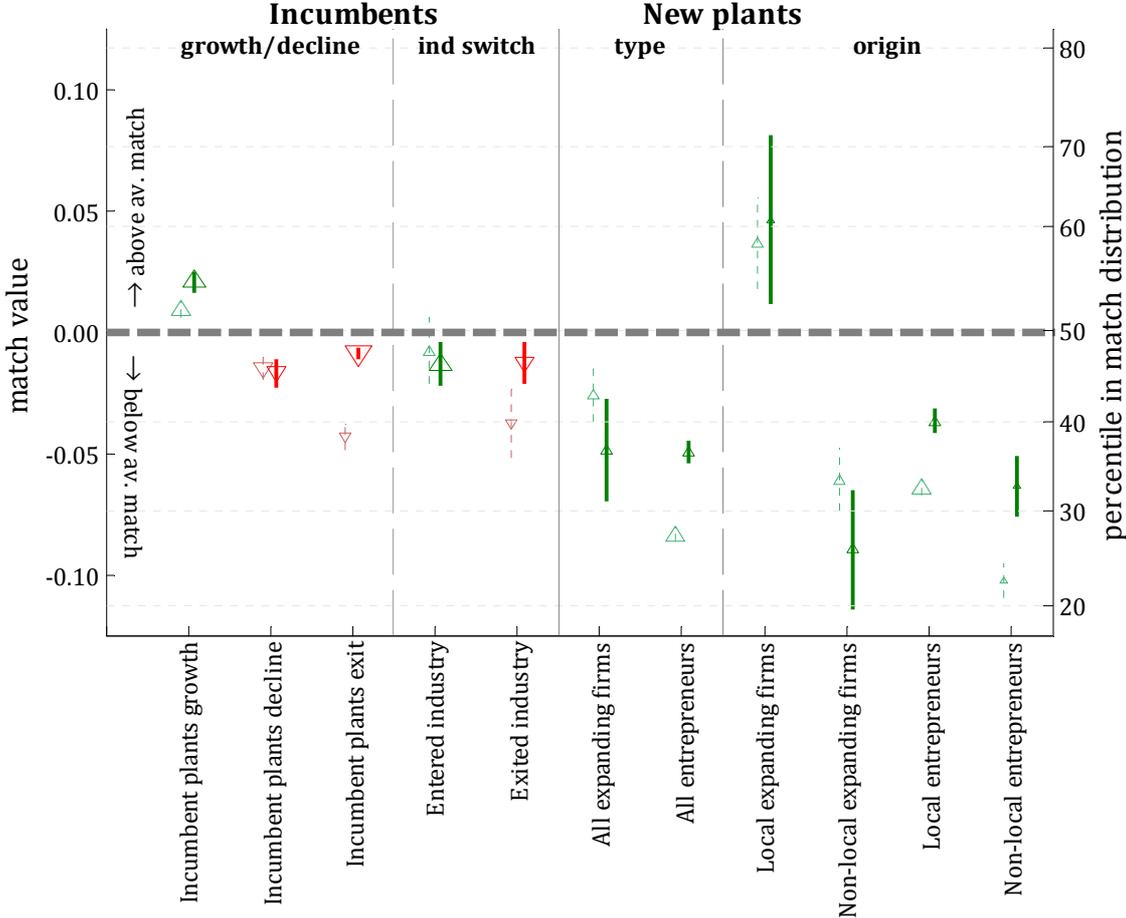
The graph depicts the development of the average capability match of a regions' local industries to the local economy of the base year 1994 ($S_{rt,1994} = \sum_i \frac{E_{irt}}{E_{rt}} \bar{Q}_{ir1994}^{rel}$), including 95% error bars. In the presence of structural change, $S_{rt,1994}$ should fall (diversification) or rise (increasing of existing focus) over time. The error bars depict a 95% confidence interval calculated as ± 1.96 times the standard deviation of the mean 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 of the local industries in which the agent type creates or destroys employment within one year after being founded, averaged across all establishments of the agent minus the region's coherence. That is, the graph shows by how much an agent's average capability match exceeds the region's coherence. As a consequence, employment-created (destroyed) at $\bar{A} > 0$ corresponds to a diversification (further focusing) of the regional capability base. Employment-creating agents are denoted with a green, upward pointing arrow, employment-destroying agents with a red downward-pointing arrow. The error bars depict a 95% confidence interval, based on the standard deviation of the mean capability match across all establishments of a given agent type. To facilitate interpretation of \bar{A} -values, the right most vertical axis provides the corresponding percentiles of a capability match value in the overall distribution of the 1994 economy.

Figure 7: Structural change by agent type over a 10-year horizon



The markers shows for each agent type the employment weighted average capability match of the local industries in which the agent type creates or destroys employment within ten years after being founded, averaged across all establishments of the agent minus the region’s coherence. That is, the graph shows by how much an agent’s average capability match exceeds the region’s coherence. As a consequence, employment created (destroyed) at $\tilde{A} > 0$ corresponds to a diversification (further focusing) of the regional capability base. Employment-creating agents are denoted with a green, upward pointing arrow, employment-destroying agents with a red downward-pointing arrow. The error bars depict a 95% confidence interval, based on the standard deviation of the mean capability match across all establishments of a given agent type. To facilitate interpretation of \tilde{A} -values, the right most vertical axis provides the corresponding percentiles of a capability match value in the overall distribution of the 1994 economy. The one-year figures reported in Figure 6 are provided as a reference with lighter markers and dotted confidence intervals.

Appendix A: Classification of industries

Table A1: Industries included in the analyses

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

Appendix B: Measuring skill relatedness

Our data refer to all individuals between the age of 18 and 65. We measure skill relatedness among industries by assessing the labor flows between industry pairs. In the period 1994 to 2010, about 4.5 million workers in Sweden switched jobs among different 4-digit industries. To avoid problems with missing industries, we keep only those industries that have nonzero employment in each year. First, we use equation (1) to calculate skill relatedness for every year between 1994 and 2010. Letting years be indexed by t and summation over omitted categories indicated by $\cdot\cdot$ this yields:

$$SR_{ijt} = \frac{F_{ijt}}{(F_{jt}F_{it})/F_{\cdot\cdot t}} \quad (B1)$$

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

$$\widetilde{SR}_{ijt} = \frac{SR_{ijt}-1}{SR_{ijt}+1} \quad (B2)$$

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

$$M\widetilde{SR}_{ij} = \frac{1}{16} \sum_{t=1994}^{2009} \widetilde{SR}_{ijt} \quad (B3)$$

Finally, we symmetrize the measure so that $S\widetilde{SR}_{ij} = S\widetilde{SR}_{ji}$:

$$S\widetilde{SR}_{ij} = \frac{M\widetilde{SR}_{ij} + M\widetilde{SR}_{ji}}{2} \quad (B4)$$

The actual condition for two industries to be skill related that we evaluate in the indicator function in equation (2) is therefore $S\widetilde{SR}_{ij} > 0$.

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 a firm identifier (see Andersson and Arvidsson, 2006), which enables us to follow establishments over time regardless of changes in ownership or legal status. Entrepreneurial entries are new establishments that create new firms (both the establishment and the firm identifiers did not exist before 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. Firstly, 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 he or 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 used to work in the same region, we take this region as the geographical origin. If no entrepreneur is found and if 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. 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, 5 and 7 and Tables 5 and 6 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 tables and graphs are shown below.

Table D1: Knowledge diffusion by outside agents (NACE-relatedness)

Agent type	Capability match to:		p-value
	home region	host region	
Non-local expanding firms	0.113 (0.007)	-0.089 (0.007)	0.000
Non-local entrepreneurs	-0.043 (0.003)	-0.081 (0.003)	0.000

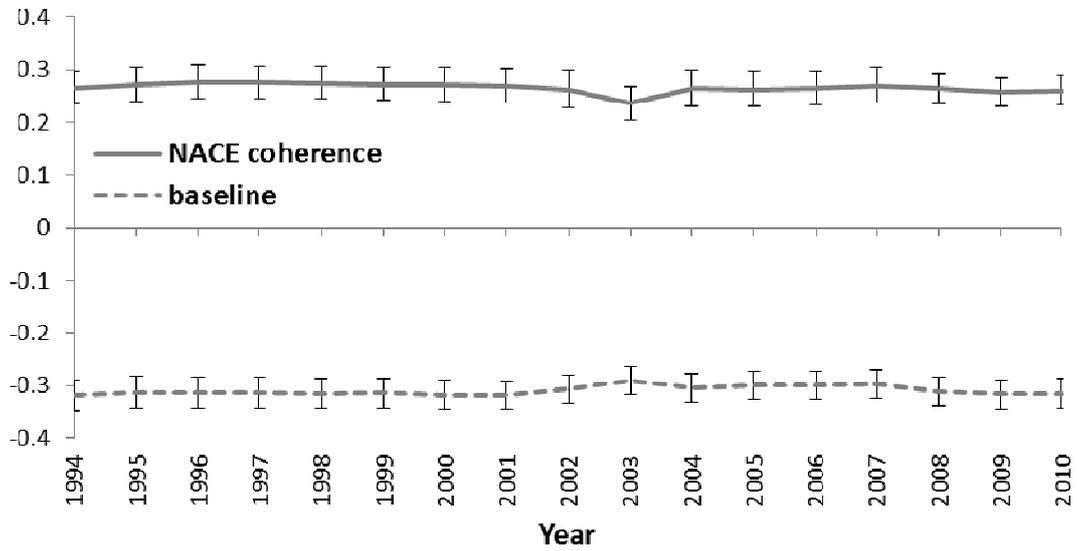
See Table 6

Table D2: Establishments' 10-year survival rates (NACE-relatedness)

Dep. var.: ≥ 10 yr survival (0/1)	(1)	(2)	(3)	(4)	(5)	(6)
Non-local agent	-0.021*** (0.004)	0.015 (0.013)				
Firm agent			0.073*** (0.009)	0.179*** (0.034)		
Local firm					0.041** (0.016)	0.109 (0.075)
Non-local firm					0.074*** (0.01)	0.170*** (0.039)
Non-local entrepreneur					-0.037*** (0.004)	-0.016 (0.013)
<hr/>						
Local X ln(rel. emp.)		0.007*** (0.003)				
Non-local X ln(rel. emp.)		0.002 (0.003)				
Firm X ln(rel. emp.)				-0.007 (0.005)		
Entrepreneur X ln(rel. emp.)				0.007*** (0.003)		
Local Firm X ln(rel. emp.)						-0.002 (0.009)
Non-local firm X ln(rel. emp.)						-0.006 (0.005)
Local entrepr. X ln(rel. emp.)						0.006** (0.003)
Non-local entrepr. X ln(rel. emp.)						0.004 (0.003)
Constant	0.157*** (0.005)	0.088*** (0.027)	0.153*** (0.005)	0.085*** (0.027)	0.156*** (0.005)	0.09*** (0.027)
Entry-year FEs	yes	yes	yes	yes	yes	yes
Industry FEs	yes	yes	yes	yes	yes	yes
Region FEs	yes	yes	yes	yes	yes	yes
Number of observations	54,055	54,055	54,055	54,055	54,055	54,055
R ²	0.0254	0.0257	0.0265	0.0269	0.0280	0.0283

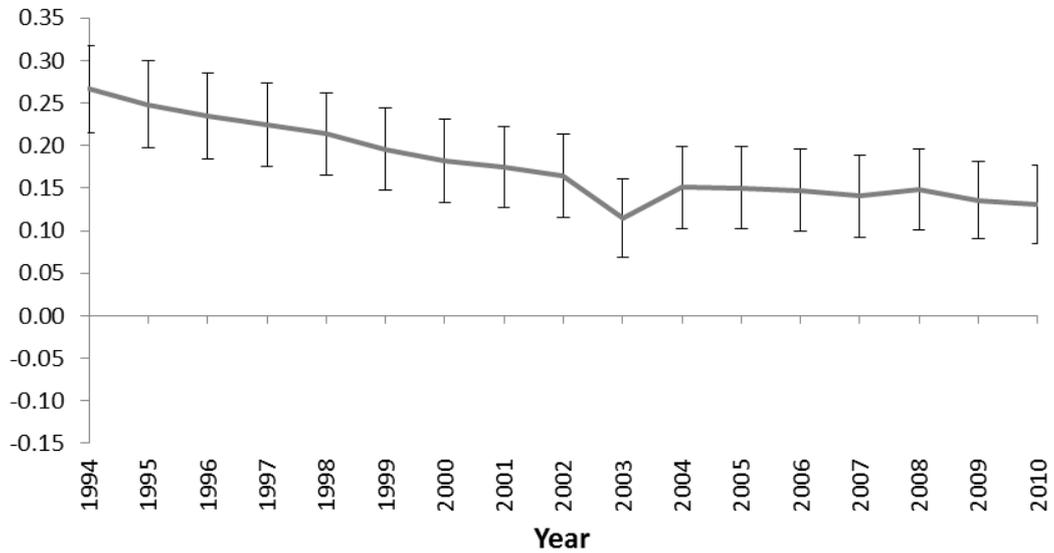
See Table 5.

Figure D1: Coherence of labor market regions' capability bases (NACE-relatedness)



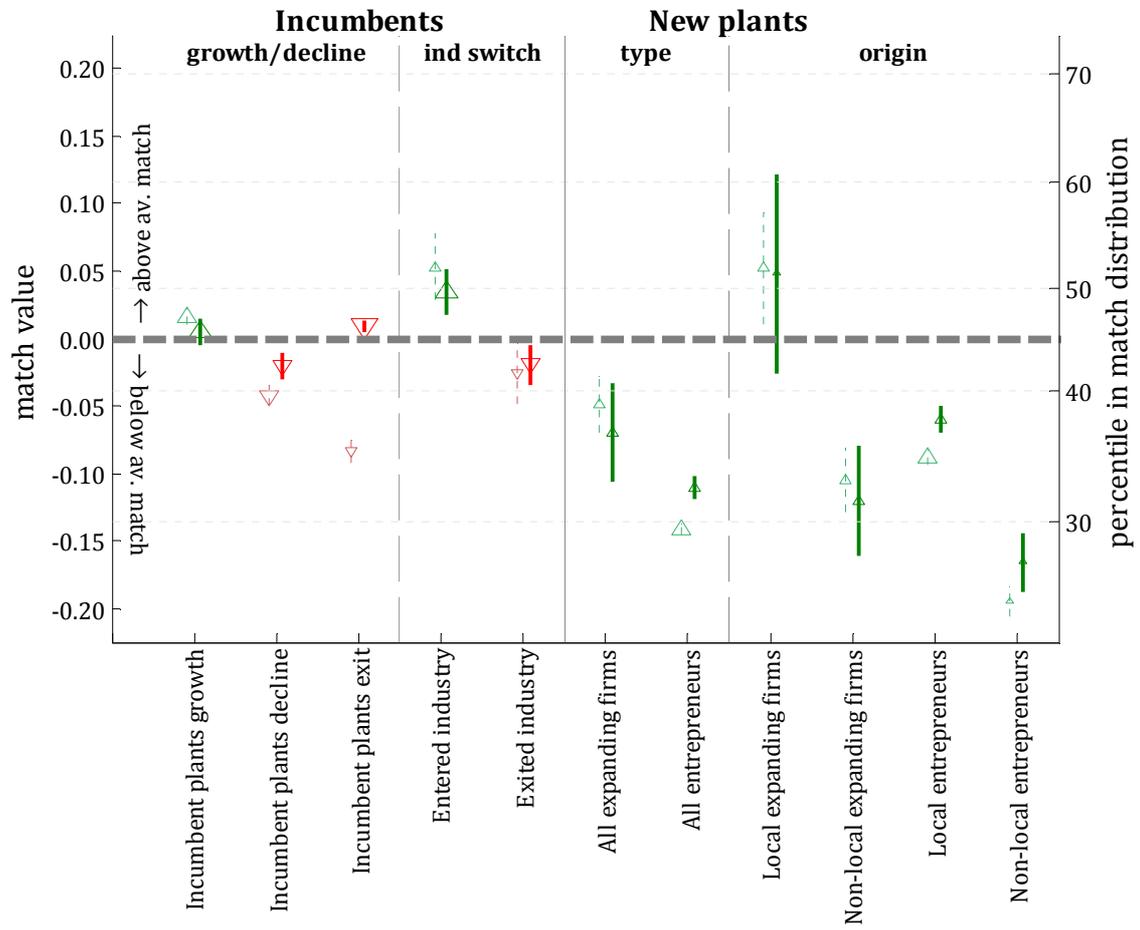
See Figure 4.

Figure D2: Structural change in Sweden's labor market regions (NACE-relatedness)



See Figure 5.

Figure D3: Structural change by agent type (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 every 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:

$$IOR_{ij} = \frac{1}{2} \left(\frac{CF_{ij}}{\sum_k CF_{ik}} + \frac{CF_{ji}}{\sum_l CF_{li}} \right) \quad (D1)$$

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 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:

Table D3: Knowledge diffusion by outside agents (input-output relatedness)

Agent type	Capability match to:		p-value
	home region	host region	
Non-local expanding firms	0.096 (0.007)	-0.075 (0.007)	0.000
Non-local entrepreneurs	-0.029 (0.003)	-0.058 (0.003)	0.000

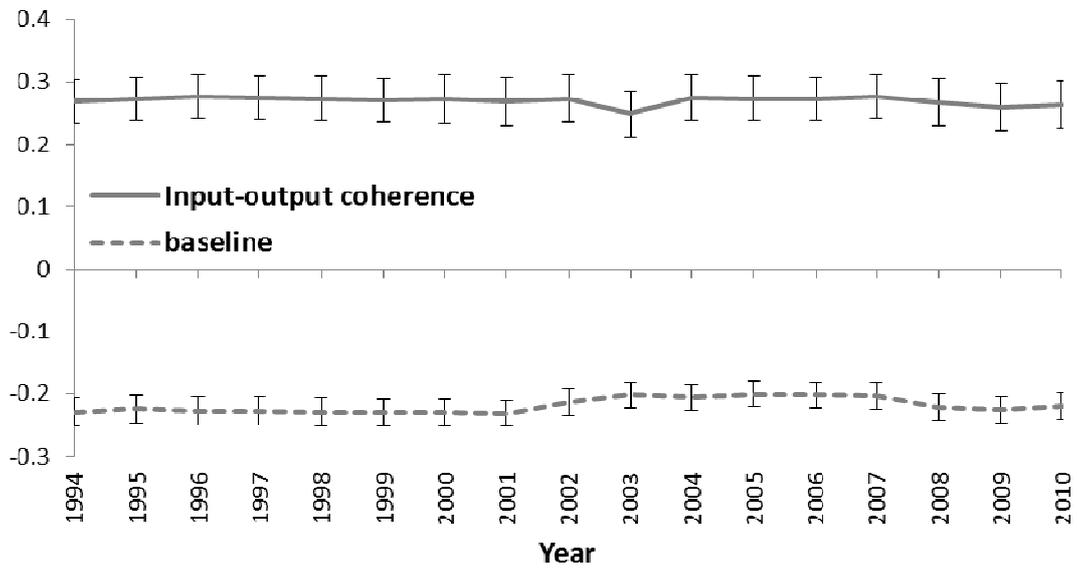
See Table 6.

Table D4: Establishments' 10-year survival rates (input-output relatedness)

Dep. var.: ≥ 10 yr survival (0/1)	(1)	(2)	(3)	(4)	(5)	(6)
Non-local agent	-0.021*** (0.004)	-0.002 (0.012)				
Firm agent			0.073*** (0.009)	0.13*** (0.028)		
Local firm					0.041** (0.016)	0.113* (0.061)
Non-local firm					0.074*** (0.01)	0.108*** (0.032)
Non-local entrepreneur					-0.037*** (0.004)	-0.03** (0.012)
<hr/>						
Local X ln(rel. emp.)		0.008*** (0.003)				
Non-local X ln(rel. emp.)		0.005* (0.003)				
Firm X ln(rel. emp.)				0.001 (0.004)		
Entrepreneur X ln(rel. emp.)				0.008*** (0.003)		
Local Firm X ln(rel. emp.)						-0.001 (0.007)
Non-local firm X ln(rel. emp.)						0.003 (0.004)
Local entrepr. X ln(rel. emp.)						0.007*** (0.003)
Non-local entrepr. X ln(rel. emp.)						0.006** (0.003)
Constant	0.157*** (0.005)	0.073** (0.031)	0.153*** (0.005)	0.067** (0.03)	0.156*** (0.005)	0.078** (0.031)
Entry-year FEs	yes	yes	yes	yes	yes	yes
Industry FEs	yes	yes	yes	yes	yes	yes
Region FEs	yes	yes	yes	yes	yes	yes
Number of observations	54,055	54,055	54,055	54,055	54,055	54,055
R ²	0.0254	0.0256	0.0265	0.0268	0.0280	0.0282

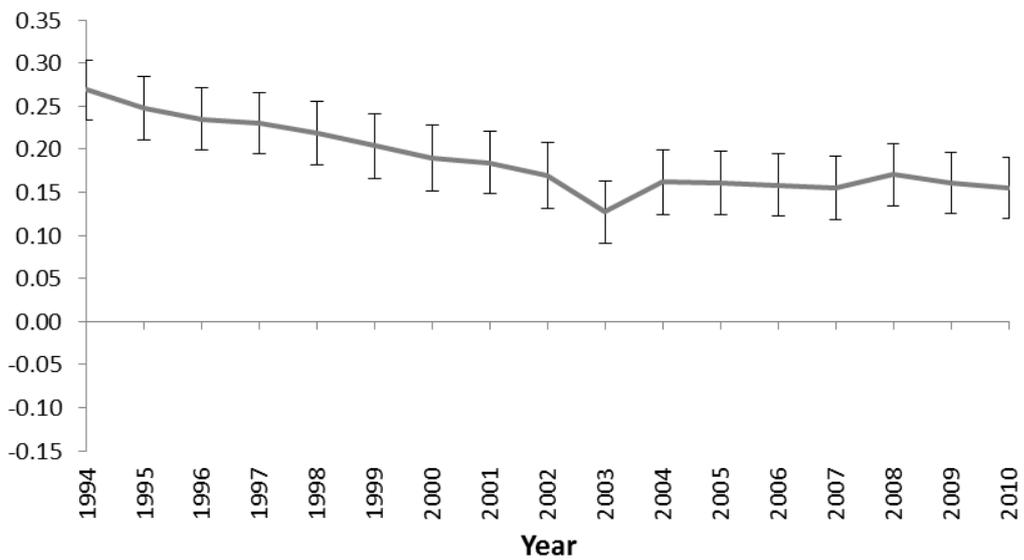
See Table 5.

Figure D4: Coherence of labor market regions' capability bases (input-output relatedness)



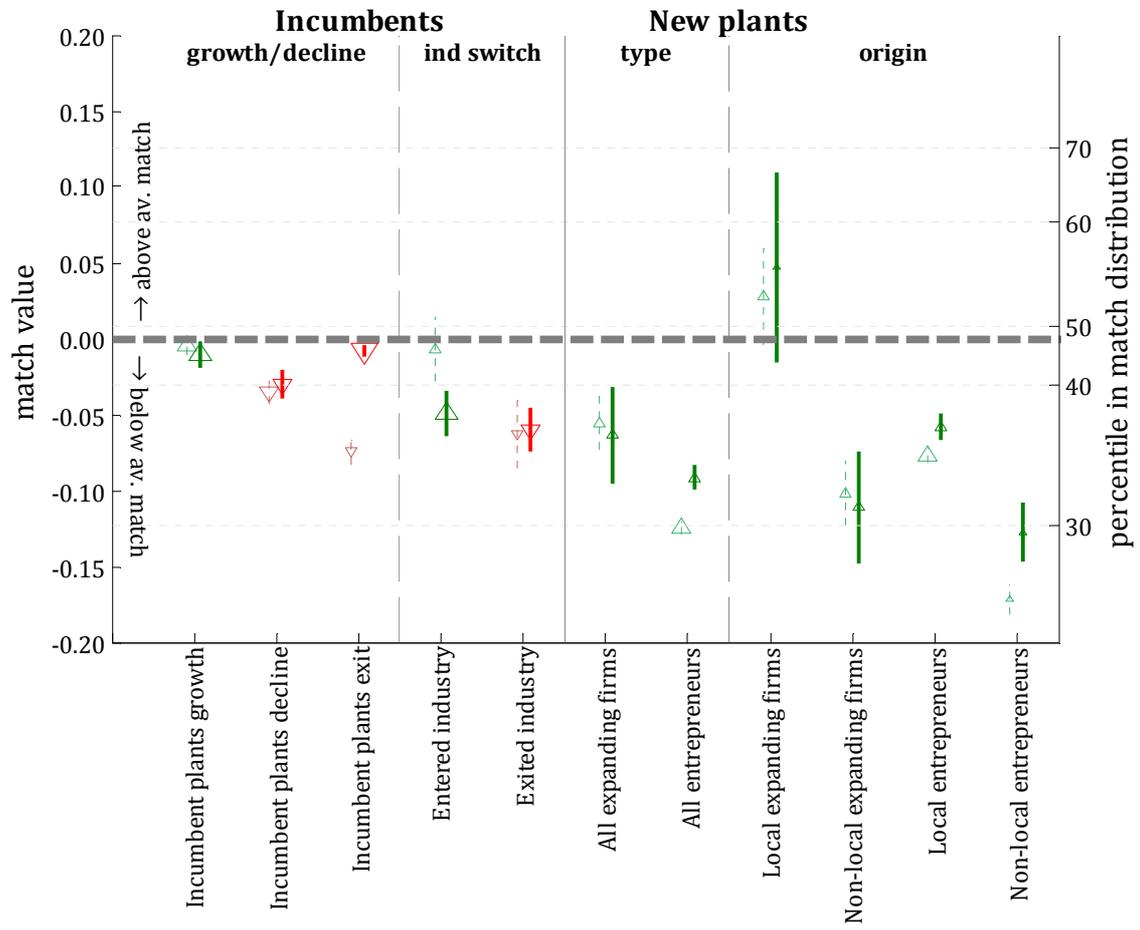
See Figure 4.

Figure D5, Structural change in Sweden's labor market regions (input-output relatedness)



See Figure 5.

Figure D6: Structural change by agent type (input-output relatedness)



See Figure 7.