Inter-industry linkages in local economies

Frank Neffke[*] and Martin Henning[#]

Paper prepared for the 51st ERSA Congress in Barcelona. An earlier version of this paper was titled "Seeds of structural change. The role of entrepreneurs and expanding firms in shaping local path dependencies."

Abstract

We investigate the extent to which a local industry is affected by an overrepresentation of related industries in the local economy. We focus on two types of inter-industry relatedness, namely, the degree to which two industries can employ a similarly skilled labor force and the degree to which two industries are connected in the value chain. We decompose changes in the employment of a local industry into the employment generated or destroyed in incumbent plants and the employment changes due to plant entry and exit. Furthermore, we classify new plants by the type and geographical origins of the plants' founders. We find that entrepreneurs have a stronger tendency than existing firms to set up plants in local industries that can draw on a strong local presence of labor market and value chain linked industries. The same holds for local founders compared to founders from outside the region. In the second part of the paper, we investigate the relative importance of the two relatedness types and whether the two types reinforce each other. We find that, in general, the growth of old plants and the employment generated in new plants is more strongly associated with the relatedness through the labor market. Moreover, for in new plant formation, the two relatedness types indeed tend to reinforce each other. In fact, local value chain linkages seem to be only important if client and supplier firms can also engage in labor sharing.

* Erasmus University Rotterdam, The Netherlands, neffke@ese.eur.nl

Lund University, Sweden, martin@keg.lu.se

1. Introduction

The performance of firms is affected by their proximity to other firms. This contention can be derived from a number of different bodies of research that range from the analyses of clusters (Porter, 1998; 2003) and regional innovation systems (Cooke and Morgan, 1998) to the more general investigation of the externalities associated with economic agglomeration (Glaeser et al., 1992; Henderson et al., 1995; Henderson, 2003). Recent research suggests that an important aspect of the clustering or agglomerating of economic activity is the linkages that exist among firms that belong to different industries. This notion is central in the cluster literature (Delgado et al., 2010), but also scholars that traditionally focused on agglomeration externalities are turning to this issue (Ellison et al., 2010). A third body of research to which the relatedness among industries is pivotal is the research on evolutionary economic geography (Boschma and Frenken, 2006), which views the economic development of regions as a branching process (Frenken and Boschma, 2011). In a recent article, Florida et al. (2011) further emphasize the importance of the presence of interrelated industries in a region, stating that, similar to the economies of scope multi-product that firms achieve, regions can benefit from geographies of scope if they host economic activities that are related.

In this paper, we investigate two types of inter-industry linkages more closely that could yield such geographies of scope. The first type is the capacity of local industries to share the same or a similar pool of workers. The second type is input-output or value chain linkages among local industries. The importance of sharing a trained local labor force and the access to specialized suppliers of intermediate goods is already discussed in the works of Marshall (1890/1920). We raise two questions concerning the role that these two types of inter-industry linkages play in local economies that, to the best of our knowledge, have not been addressed thus far.

First, we note that employment changes in a local industry derive from different sources. The local industry's existing plants may shrink or grow, plants sometimes relocate to another region, new plants are built and old plants are shut down. We use rich data on the Swedish labor market for the years 2004-2007 to study how each of these sources adds (or subtracts) employment in a local industry. We furthermore distinguish between new plants of different types of founders, focusing on the geographical origins of the founder and whether the founder is an existing firm or an entrepreneur that starts a new firm. Our findings indicate that plants of entrepreneurs and of founders from within the direct vicinity of a local industry are most likely to set up new plants that are closely connected to the rest of the local economy through labor market and value chain linkages.

Second, we study whether the growth rate of existing plants and the employment added to a local economy by new plants are associated with the degree to which the local economy is overspecialized in industries that are closely connected through labor market and value chain linkages. Furthermore, we investigate whether the effects of both Marshallian externality channels reinforce each other. We find that growth rates and new plant employment are largest in locations that are overspecialized in activities that maintain labor market linkages to the industry in which the new employment is generated. An overspecialization in value chain linked activities is in general less important. However, an important exception if the value chain linkages are complemented by labor market linkages. If the local economy is overspecialized in such industries, new plant employment is significantly higher. That is, we find that the two Marshallian externality channels reinforce each other.

In the section 2, we discuss the existing literature and theories on the importance of inter-industry linkages for regional economic development and derive hypotheses for our own analyses. In section 3, we present our research strategy. Section 4 discusses the empirical findings and section 5 concludes.

2. Literature review

Marshallian agglomeration forces, clusters and related variety

Economic geographers have been intensively researching agglomeration externalities, i.e., the benefits and costs of the geographical co-location of firms, over the past three decades. According to these scholars, there are a number of ways in which firms that are located in the same region affect each other's performances. The externalities generated among firms that belong to the same local industry are described as localization externalities. The standard treatment of localization externalities refers back to Marshall (1890/1920) and distinguishes three channels through which firms affect one another. The first channel is the sharing of a pool of specialized labor, the second is the demand that is created locally for specialized intermediates producers and the third are serendipitous knowledge spillovers among local firms. Following Glaeser et al. (1992), the interdependences among firms that belong to different industries are usually summarized as either externalities derived from economic diversity (so-called Jacobs externalities) or externalities associated with the size of the local economy (so-called urbanization externalities). Rosenthal and Strange (2004) provide an excellent overview of this literature and De Groot et al. (2009) have pooled a large number of empirical findings in a meta-analysis.

A question that, until recently, has received much less attention is the degree to which firms in related industries generate agglomeration externalities for one another. Authors writing on industrial clusters (Porter 1998, 2003; Maskell, 2005) commonly recognize that inter-industry linkages play an important role. In fact, all three Marshallian agglomeration externality channels play a role in clusters. For instance, strong local buyer-supplier linkages are often credited for the exceptional performance of a geographical cluster of firms (Porter, 1998). Other authors stress the importance of labor market linkages within a cluster (Almeida and Kogut, 1999; McCann and Simonen, 2005; McCann and Faggian, 2006). The literature on Regional Innovation Systems (RIS, Cooke and Morgan, 1998) furthermore highlights local knowledge spillovers as a source of regional competitiveness.

In recent work, scholars in different research traditions have started to gather statistical evidence on the relative importance of different types of inter-industry linkages. Within the traditional research on agglomeration externalities, Ellison et al. (2010) argue that inter-industry linkages can be used to shed light on disentangling the effects of different Marshallian externality channels. Studying localization externalities hitherto often meant eliciting a causal relation running from the degree of a region's specialization in a specific industry to the local performance of that industry. However, this research set-up cannot assess which of the three Marshallian externality mechanisms is responsible for the link between local specialization and performance, because all three mechanisms will be triggered simultaneously if firms belong to the same industry. However, if firms belong to different industries, some channels may be activated, while others may not. For instance, industries that require similarly skilled labor but different intermediate goods would affect each other through the labor market channel, but not through the value chain channel. Ellison and his co-authors exploit this fact to assess which inter-industry linkages cause different industries to locate in the same regions. These authors find that all Marshallian externality channels contribute positively to the co-location patterns of US manufacturing industries. A similarly motivated research strategy can be found in Dauth (2010).

In cluster research, Porter (2003) categorizes all industries into mutually exclusive clusters by drawing on information of the relatedness among industries. This information is derived from co-agglomeration patterns of different industries, mixed with qualitative judgment. Subsequently, this cluster classification is used by Delgado et al. (2010), who show that entrepreneurial activity in a local industry increases in the presence of different industries that belong to the same cluster.

A third strand of literature that focuses on the relatedness of industries is evolutionary economic geography (Boschma and Frenken, 2006). According to this evolutionary perspective, innovation arises when previously existing ideas are combined into new ones. Regions that host a large number of different industries should, therefore, be fertile building grounds for innovation. This is reminiscent of Jacobs's (1969) assertion that cities need economic diversity to grow, although Jacobs originally identified the deeper divisions of labor in diversified cities as the source from which novelty would spring, rather than the opportunity for a Schumpeterian type of recombinant innovation. This Schumpeterian twist to Jacobs's work seems to have been introduced by Glaeser et al. (1992) and is later picked up in the evolutionary literature by Frenken et al. (2007). The key insight in Frenken et al. (2007) is that the cross-industry spillovers that are associated with Jacobs externalities must be much stronger if industries are related. Unrelated industries can at best have a risk-diversification effect, but the occurrence of knowledge spillovers among them should be considered to be exceptional. This leads these authors to assert that not variety, but related variety is most conducive to productivity growth in local economies. In more recent work, Boschma and Frenken (2009) and Frenken and Boschma (2011) add another argument for the importance of relatedness among industries for understanding regional development. These authors depict regional diversification as a branching process in which new activities in a region do not arise out of thin air, but build on local assets. In other words, new activities are often related to already existing activities in the region. At the national level, Hidalgo et al. (2007) show that countries indeed diversify their export portfolios according to such a branching logic. In a study on the long term evolution of Swedish regions, Neffke et al. (2011a) show that similar processes are at work at the regional level.

In sum, different research traditions identify benefits of the co-location of related industries. In this article, we investigate the effects of two of the three Marshallian externality channels: local labor market pooling and input-output linkages. The third channel, knowledge spillovers, is omitted, because it is notoriously difficult to measure. Moreover, labor mobility is often regarded as a particularly important mechanism for knowledge diffusion (Almeida and Kogut, 1999). In this sense, labor market linkages between industries will subsume some of the less serendipitous knowledge spillovers that occur within a region.

Sources of changes in employment of local industries

A first issue on which we aim to shed some light, is the fact that it is unclear whether the two Marshallian channels we investigate matter equally for all sources of growth of local industries. For instance, the degree to which incumbent firm growth depends on the presence of related industries maybe different from the employment generated by the entry of new establishments into a local economy. Henderson (2003), for instance, finds that single plant firms benefit more from agglomeration externalities than corporate plants do. Neffke et al. (2011b) show that agglomeration externalities affect the survival rates of corporate plants in different ways compared to stand alone plant. Delgado et al. (2010), however, show that entry rates grow faster in the presence of related industries both for start-up firms and for new plants of existing firms.

To investigate the differences among different sources of employment change in a local industry, we decompose the employment growth (or decline) of local industries into three components: the growth and decline of incumbent firms, the entry of new and the exit of old establishments and the in- and outmigration of establishments from a region.[1] The literature on regional lock-in and renewal (Grabher 1993; Martin and Sunley, 2006) considers actors from outside the region as important for generating an influx of new activities in a region. Furthermore, in the evolutionary literature, entrepreneurs are typically regarded as agents that create novelty in a region, but on other occasions it is argued that entrepreneurs often reinforce regional industrial structures rather than transforming them (Staber, 2005). To investigate these issues, we distinguish among different types of new plants according to whether they are founded by existing plants or by entrepreneurs, and in terms of the founder's geographic origin.

A second way in which we aim to deepen the understanding of the two Marshallian channels under investigation is by studying whether both channels reinforce each other. This interaction is implicitly suggested in much of the literature on clusters and regional innovation systems. For instance, Cooke and Morgan's (1998) description of the car manufacturing cluster in Baden-Württemberg highlights innovation efforts between car producers and their main component suppliers. The need to synchronize innovation efforts along the value chain may, therefore, explain why proximity of specialized intermediates to final goods producers is crucial to the performance of a cluster. As long as transportation costs are low, it is not the exchange of goods but the exchange of knowledge between a supplier and a client firm that requires spatial proximity. Intertwined innovation efforts in different industries makes it also likely that the industries will be able– and even need– to share some of their labor force. This suggests that the importance of value chain linkages increases in the presence of labor market linkages.

[1] A fourth source of employment change in a local industry is plants that switch their production to another industry. However, it is difficult to assess how many of these plants really change their industrial orientation and in how many cases we are observing merely a coding error. Therefore, we do not further discuss this fourth source.

Theory and hypotheses formation

The research literature on clusters, agglomeration externalities and evolutionary branching processes, all suggest that local industries benefit from each other's presence if they are somehow related to one another. Therefore, related industries are likely to co-locate and we expect that an industry is larger in locations with an overrepresentation of the industries to which the industry maintains labor market or value chain linkages.

Hypothesis 1a: The size of a local industry is positively associated with the extent to which industries with which the industry can engage in labor market pooling are overrepresented in the local economy.

Hypothesis 1b: The size of a local industry is positively associated with the degree to which industries that are closely connected through the value chain are overrepresented in the local economy.

Turning to the changes in employment, matters are less straightforward. On the one hand, both cluster theories and agglomeration externalities accounts attribute positive effects to the clustering of firms that are engaged in similar activities. However, at the same time, the fiercer competition for local resources may drive up the prices of these resources. Therefore, the balance of costs and benefits of spatial clustering may turn negative. However, arguably negative externalities are particularly strong for firms that are direct competitors of one another. Therefore, a large presence of firms in the same industry will be negatively associated with the growth rate of a local industry. Firms that are not in the same industry, in contrast, often do not directly compete with one another. Also these firms may drive up local resource prices. However, there are reasons to believe they will do so to a lesser extent. On the one hand, if the business cycles of related industries are not perfectly synchronized, the pooling of labor markets and suppliers of intermediates may be more effective. The reduction in employment risk and demand fluctuations resulting from such asynchronous business cycles offers a more a stable environment for both labor and intermediates suppliers (Rosenthal and Strange, 2004). This stability is attractive for workers and may even lead to lower wage demands (Diamond and Simon, 1990). Similar arguments hold for suppliers of non-tradable intermediates.

A second argument for why the cost-benefit balance of clustering of firms may be different for firms that are not in the same, but in related activities, is brought forward in evolutionary economic geography. According to Frenken et al. (2007), a local environment with a rich variety of related industries will exhibit faster learning dynamics. This claim builds on the notion found in the management literature that there exists an optimal cognitive distance for collaborating partners (Nooteboom, 2000). In order to learn from one another, firms should be sufficiently similar to one another to be able to understand each other's technologies, but they should simultaneously be sufficiently different from one another to give access to ideas that the other firm did not already possess itself. Along these lines, Boschma et al. (2011) put forward that firms benefit most from labor flows that originate in firms from related industries. Such reasonings underscore the superior benefits of a clustering of firms in related industries as compared to a clustering of firms in the same industry. In a survival analysis of Swedish manufacturing plants, Neffke et al. (2011b) show that proximity to firms in the own industry does not lead to significantly higher survival rates, whereas a local presence of firms in related industries do increase a plant's survival chances. These considerations lead to the following hypotheses:

Hypothesis 2a: The growth of a local industry in incumbent plants is negatively associated with the size of the local industry.

Hypothesis 2b: The growth of a local industry in incumbent plants is positively associated with the extent to which industries with which the industry can engage in labor market pooling are overrepresented in the local economy.

Hypothesis 2c: The growth of a local industry in incumbent plants is positively associated with the extent to which industries that are closely connected through the value chain are overrepresented in the local economy.

New plants are often set up either by local firms, or by employees leaving these local firms. Therefore, the larger the local industry, the more firms and entrepreneurs are available for setting up new plants and the larger the employment generated by these new plants. Unlike the growth of the local economy in incumbent plants, we therefore expect that new plant formation not only depends positively on the presence of related industries, but also on the presence of the own industry.

Hypothesis 3a: The new plant employment in a local industry is positively associated with the size of the local industry.

Hypothesis 3b: The new plant employment in a local industry is positively associated with the extent to which industries with which the industry can engage in labor market pooling are overrepresented in the local economy.

Hypothesis 3c: The new plant employment in a local industry is positively associated with the extent to which industries that are closely connected through the value chain are overrepresented in the local economy.

As discussed in the previous section, founder types may differ in the degree to which they set up plants in industries that are well-embedded in local economies. However, it is difficult to arrive at clear-cut hypotheses. On the one hand, entrepreneurs and outside actors are often considered to be important sources for generating novelty in a region. If this holds true, these founder types would be more likely to enter industries that are not strongly embedded in the local economy. However, entrepreneurs are also more dependent on the local environment to complement their limited assets and may therefore choose to locate in places with a strong presence of related industries. Similarly, outside actors who willingly choose to leave their home region and forfeit the advantages provided by their current local network in that home region will only do so if they are experience compensating advantages from having many firms in related activities in close vicinity.[2] We are therefore undecided about the differences among founder types:

Hypothesis 4a: The new plant employment for plants set up by entrepreneurs in a local industry may either be more or less positively associated with the extent to which industries with which the industry can engage in labor market pooling are overrepresented in the local economy than the employment of new plants of existing firms.

Hypothesis 4b: The new plant employment for plants set up by entrepreneurs in a local industry may either be more or less positively associated with the extent to which industries that are closely connected through the value chain are overrepresented in the local economy than the employment of new plants of existing firms.

[2] Sorenson (2003) argues that social networks are particularly important reason for entrepreneurs to set up firms in their home regions.

Hypothesis 5a: The new plant employment for plants set up by local founders in a local industry may either be more or less positively associated with the extent to which industries with which the industry can engage in labor market pooling are overrepresented in the local economy than the employment of new plants founders from outside the region.

Hypothesis 5b: The new plant employment for plants set up by local founders in a local industry may either be more or less positively associated with the extent to which industries that are closely connected through the value chain are overrepresented in the local economy than the employment of new plants by founders from outside the region.

The importance of having local suppliers of intermediate goods and services in a region is often supported by the argument that spatial proximity between trading firms can reduce transactions costs (Martin and Ottaviano, 2001). Transactions costs are notoriously high for trades that involve the exchange of knowledge (Arrow, 1962; Teece, 1982). Exchanging knowledge, especially knowledge that is mostly tacit (Polanyi, 1966), can be prohibitively difficult without an (at least temporary) exchange of labor (Gertler, 2003). However, if two trading firms exchange labor, they are not just trading partners but also engage in labor sharing. Therefore, the highest gains of spatial proximity between a client and a supplier firm should be expected to occur if the firms also require a similar pool of labor, leading to the following hypothesis:[3]

Hypothesis 6a: The growth of a local industry is positively associated with the extent to which industries are overrepresented in the local economy with which the industry not only maintains value chain linkages but can also engage in labor market pooling.

Hypothesis 6b: The new plant employment of plants set up by entrepreneurs in a local industry is positively associated with the extent to which industries are overrepresented in the local economy with which the industry not only maintains value chain linkages but can also engage in labor market pooling.

[3] A good illustration is the already mentioned example of two firms that orchestrate joint innovation efforts along the value chain (Cooke and Morgan, 1998).

3. Research methodology

Data

In our empirical analyses, we use detailed data on the Swedish economy for the years 2004 to 2007. The data are derived from the national employment registers and contain information on all individuals living in Sweden. For each individual, we have information on their wage, their income from a private enterprise, and the establishment in which she or he works. We aggregate these data to generate a database of all plants in Sweden. The database contains information on the industry in which the plant is active, its municipality, and, if the plant is part of a larger firm, which the other Swedish establishments of that firm are. Because we have four years of data, we can determine plant entry and exit patterns.

An interesting feature of our data is that they contain information on both wage income and on income derived from a private business. The firm identifier of a plant allows us to assess which new plants belong to larger firms and which new plants constitute entrepreneurial ventures, that is, which plant formations lead to the founding of new firms. In each entrepreneurial plant, we identify the individual with the highest income derived from a private business and assume that this person is likely to be the new firm's entrepreneur. By investigating the employment history of this individual, we are able to determine whether the entrepreneur is a local entrepreneur or whether he or she came from outside the region. Similarly, it is possible to attribute a geographic origin to plants that are set up by existing firms, by determining the location where the firm had most of its employment prior to the establishment of the new plant. In our descriptive analyses, we investigate whether plants with different origins differ with respect to their embeddedness in the local economy. We will, however, first define what we mean by "embeddedness" in this context.

Local embeddedness

The unit of analysis in this paper is the local industry. Local industries are combinations of a municipality and an industry, for instance the shipbuilding-in-Malmö industry. Industries are defined at the 4-digit level of the European NACE Rev 1.1 industrial classification system, and we take the smallest spatial units that are available in our dataset, the Swedish municipality. We define embeddedness with respect to the Marshallian channels through which two industries are linked to one another. As noted above, we restrict ourselves to two channels, the input-output or value chain linkages and the labor sharing among industries.

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

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

Input-output data are only available at the 2-digit level. Because our analyses make use of industries defined at the 4-digit level, we must assume that the input-output linkages that exist between two 2-digit sectors are representative for the linkages that exist among the 4-digit industries that belong to these sectors.

The degree to which two industries can share the same labor is more difficult to assess. We draw on a method presented in Neffke and Henning (2011) that uses cross-industry labor flows to assess the degree of skill-relatedness among industries. The intuition behind this measure is that individuals that switch jobs will try to prevent making their acquired work experience obsolete by switching to jobs where the skills associated with the individual's previous jobs are still appreciated. Therefore, industries among which one observes large labor flows are likely to be strongly skill-related. Neffke and Henning (2011) compare these observed cross-industry labor flows to predictions of the labor flows based on regression based predictions using sizes, growth rates and wage levels of the industries involved as explanatory variables. If F_{ij} is the observed labor flow between industry i and j, and $\hat{F}_{ij}$ is the fitted labor flow in a regression of F_{ij} on the above mentioned industry characteristics, then the skill-relatedness from i to j, SR_{ij} is defined as:

2) $SR_{ij} = \frac{F_{ij}}{\hat{F}_{ij}}$

A value of one indicates that the labor flows are exactly matching our baseline predictions, indicating that the industries are neither related nor unrelated. Values over one arise if the labor flows are larger than expected and such values thus indicate that the industries are skill-related. Values under one indicate the industries are not skill-related. To avoid circularities in our research design, we calculate this index, excluding all labor flows into newly established plants.

We are thus able to quantify the strength of the two different types of inter-industry linkages for each combination of two industries and, for each local industry, we can calculate how large the local presence of related industries is. We call two industries skill-related if the SR index exceeds 1. Industries are input-output related if the IOR index exceeds 0.026206, the value at which there are about as many input-output related pairs of industries as there are skill-related industry pairs. Next, for each local industry, we calculate the amount of related employment according to both relatedness measures found in the municipality. If E_{mi} denotes the employment of industry i in municipality m, the related employment is given by:

3) $E_{mi}^{rel} = \sum_{k \neq i} I^{rel}(i,k) E_{mk}$

rel can be either SR or IOR and denotes the relatedness type. $I^{rel}(i,k)$ is an indicator function that assumes the value one if industry k is related to industry i according to relatedness type rel. In other words, E_{mi}^{rel} is the sum of the local employment in industries that are related to industry i.

In particular:

4) $I^{SR}(i,k) = \begin{cases} 1 \text{ if } SR_{ik} > 1 \\ 0 \text{ otherwise} \end{cases}$

5) $I^{IOR}(i,k) = \begin{cases} 1 \text{ if } IOR_{ik} > 0.026206 \\ 0 \text{ otherwise} \end{cases}$

It is important to note that industry i is not considered to be related to itself and the own employment is thus excluded when calculating related employment. Next, we calculate the location quotient of the related employment:

6) $LQ_{mi}^{rel} = \frac{E_{mi}^{rel}/E_m}{E_i^{rel} \ E}$

In equation (6), we have adopted the convention that the omission of a subscript indicates the summation over all elements represented in the subscript. E_m is thus the total employment in the region, E_i^{rel} is the total employment in related industries in the whole of Sweden and E is total employment in Sweden summed across all industries.

The location quotient of industries related to industry i in municipality m measures the degree to which the municipality is overspecialized in these related industries compared to the national average. The higher this number, the better the relative fit is between industry i and the rest of the local economy with regard to the Marshallian externality channel that is captured by the relatedness measure. Because of this property, we will refer to LQ_{mi}^{IOR} as the value chain embeddedness and LQ_{mi}^{SR} as the labor market embeddedness of industry i in municipality m.

Skill-relatedness and input-output relatedness are not mutually exclusive. That is, using the threshold levels in (4) and (5), some industries are both skill-related and input-output related. The variable $LQ_{mi}^{IOR\&SR}$ measures a municipality m's overspecialization in industries that are both input-output related and skill-related to industry i. We furthermore define $LQ_{mi}^{IOR\ ex\ SR}$ and $LQ_{mi}^{SR\ ex\ IOR}$ as the overspecialization in industries that are, respectively, exclusively input-output related or exclusively skill-related to industry i.

The distributions of location quotients are strongly right-skewed. The index's interval corresponding to relative underspecialization runs from zero to one, whereas the interval that corresponds to overspecialization runs from one to infinity. This reduces the meaningfulness of averages of location quotients. To cope with this problem, we use the following function to transform location quotients:

7) $LQ' = f(LQ) = \frac{LQ-1}{LQ+1}$

This function maps the values of the location quotient on the interval [-1,1] in such a way that overspecialization and underspecialization are treated symmetrically. Regional overspecialization corresponds with LQ values between zero and one; regional underspecialization corresponds with LQ values between minus one and zero. That is, if one compares two regions, one where the industry's share of regional employment is a times the industry's national share, and one in which it is $\frac{1}{a}^{th}$ of the industry's national share, the transformed location quotients have the same absolute values, but opposite signs:[4]

[4] For example, a location quotient of 3 maps into $LQ' = \frac{1}{2}$, whereas a location quotient of $\frac{1}{3}$ maps into $LQ' = -\frac{1}{2}$.

8) $f(a) = \frac{a-1}{a+1} = \frac{1-{}^{1}/_{a}}{1+{}^{1}/_{a}} = -\frac{{}^{1}/_{a}-1}{{}^{1}/_{a}+1} = -f({}^{1}/_{a})$

We refer to the transformed values of the location quotients of different types of related employment, $LQ\ {}_{mi}^{rel}$, as *embeddedness indices*.

Sources of local employment change

Changes in the level of employment in a local industry results from changes in *incumbent employment*, i.e., changes in the employment sizes of existing establishments, and from establishments entering or exiting the local industry. Plants that enter or exit a local industry are subdivided by founder type and by geographical origin. First, plants belonging to firm that existed already in the year prior to the plant's entry are regarded as an expansion of the parent firm's activities. The geographical origin of the new plant is the plant's parent firm's main location, i.e., the municipality in which the parent firm had most of its employment prior to the establishment of the new plant. If we find no previously existing parent firm in our data, the formation of the new plant is tantamount to establishing a new firm. Therefore, we call such plant entries *entrepreneurial entries*. The geographical origin of an entrepreneurial plant entry is assessed by examining which of the plant's employees has the largest income from a private business. Assuming that this is the new firm's entrepreneur, we use the municipality where the individual worked in the year prior to the plant entry as the geographical origin of the new plant.

Incumbent employment changes and plant exit and entry are by far the most important sources of the changes that occur in a local industry's employment. However, to get a complete overview, we also determine whether plants relocate into or out of the municipality. Table 1 provides an overview of the different sources of local industry employment changes.

Each source of employment change can affect the employment in a local industry. For instance, if local firms set up new car manufacturing plants in Gothenburg, the car-industry-in-Gothenburg experiences employment growth due to expanding local firms. Furthermore, each local industry is characterized by its set of embeddedness indices. By calculating weighted averages of these embeddedness indices for all local industries in a municipality, we arrive at a set of variables that capture the degree to which an employment source matches the industrial structure of the local economy.

9) $M_m^{S,rel} = \sum_i \frac{E_{mi}^S}{E_m^S} LQ\ {}_{mi}^{rel}$

E_{mi}^{S} is the employment change associated with source S in municipality m and industry i and E_{m}^{S} is the total employment change from source S in municipality m. Thus, the weights $\left(\frac{E_{mi}^{S}}{E_{m}^{S}}\right)$ in (9) reflect the local industry's share of the total employment change in the region that can be attributed to employment source S. This procedure gives one match variable for each source and each relatedness type in each municipality. We also calculate the match of a municipality's present industrial employment structure with itself, and the match between a hypothetical industrial employment structure that assumes that the share of each municipality's employment in an industry is proportional to its overall size. These match indices are given in equations (10) and (11) and provide a baseline against which other match indices can be compared:

10) $M_{m}^{own,rel} = \sum_i \frac{E_{mi}}{E_m} LQ_{mi}^{rel}$

11) $M_{m}^{prop,rel} = \sum_i \frac{E_{mi}^{prop}}{E_{m}^{prop}} LQ_{mi}^{rel}$

In these equations, E_{mi} denotes the employment of municipality m in industry i and $E_{mi}^{prop} = \left(\frac{E_m}{E}\right) E_i$, maintaining the convention that omitting a subscript indicates the summation across the subscript's domain.

Export industries

Not all industries are equally interesting for our study. For instance, retail activities will abound in cities with large populations. Indeed, the geographical distribution of a substantial amount of economic activity predominantly follows the distribution of population across the country. Porter (2003) dubs industries in such activities *local* industries. We use a Ellison and Glaeser's (1997) raw index of an industry's geographical clustering to assess to what extent an industry's geographical location pattern follows the spatial distribution of overall economic activity.[5] Industries that are heavily concentrated in a few locations are likely to export their products to other parts of the country and beyond. In contrast, industries that are spread out across space will often serve local markets. To emulate Porter's (2003) classification procedure, we use a cut-off according to which about two thirds of all local employment is

[5] We do not use the Ellison-Glaeser index, which controls for an industry's plant size distribution. The reason is that for our definition, it does not matter whether or not an industry concentrates because there are only few active plants, it only matters that they *are* spatially concentrated.

considered local. The remaining industries are denoted export industries. Another group of industries that is severely constrained in terms of its location choice are resource extraction industries. For instance, the fishing industry has to locate in coastal areas, and a location's suitability for ore mining depends on its geological conditions.

Given their strong dependence on either the local population or on natural resources, both local industries and extraction industries are not well suited for the analyses in this paper. We therefore omit these industries from the analyses in the following sections.[6] However, although the omitted industries are not primarily affected by Marshallian externalities, these industries may still generate externalities for other industries. Mining skills may be related to skills used in mechanical engineering and iron ore is an input in steel making. Therefore, we do use the employment in these industries when calculating the embeddedness and match indicators described above.

- Table about here -

4. Empirical findings

Descriptive analyses

In the period 2004-2007, the Swedish labor market consists of about 4.1 million workers. This economic activity is spread out across 290 different municipalities, the economic size of which ranges from about 750 employees to over 500,000 employees for the country's capital Stockholm. We distinguish among a total of 415 different industries, 259 of which are considered as non-natural resource based export industries. The export industries employ together about 1.2 million employees. We restrict the analyses to the 198 export industries for which we can calculate all three embeddedness indices, LQ_{mi}^{SR}, LQ_{mi}^{IOR} and $LQ_{mi}^{IOR\&SR}$.

Growth, entry and exit rates are available for the years 2004, 2005 and 2006. However, a substantial amount of new plants (about one third) enter and exit in the same year. These plants simultaneously reflect positively and negatively on the success of a local industry and, therefore, obscure our findings. We choose to interpret these plants as turbulence and omit them from our analyses. However, this

[6] In addition to extraction industries and industries meeting the threshold condition for local industries, we also omit a small number of other industries from our analyses, namely: 5221 (retail sale of fruit and vegetables), 5222 (retail sale of meat and meat products), 5223 (Retail sale of fish, crustaceans and mollusks), 5522 (camping sites), 5523 (other provision of lodging), 7521 (foreign affairs), 7522 (defense activities), 9003 (Sanitation, remediation and similar activities), 9253 (Botanical and zoological gardens and nature reserves activities) and 9900 (Extra-territorial organizations and bodies).

means that we must observe a new plant for at least two years to be counted as a real entry or exit. As a consequence, we observe plant entry only in the years 2004 and 2005 and plant exit only in the years 2005 and 2006. The different employment sources add on average about 123,000 employees and remove about 110,000 employees from the local export industries a year. Table 1 provides definitions for the different employment sources and for two baselines reflecting the industrial composition of the Swedish economy, along with the average local employment change associated with each employment source.

Next, we calculate the match variables as defined in equation (9) for each municipality. The match variables reflect the degree to which industries that are related to the industries in which a source generates or destroys employment are overrepresented in the local economy. Table 2 provides averages across all municipalities for each of the different match variables. If an employment source is not adding or destroying employment in a municipality, this municipality does not contribute to the average match of the source in that year.

A score of zero in Table 2 indicates that, on average, the employment source adds (or destroys) employment in industries for which there is neither over- nor underrepresentation of related industries in the region. To interpret these numbers, we generate the two baselines in equations (11) and (12). The first baseline ("current industrial structure") calculates the match of the current industrial profile of a region against itself. However, against this baseline we cannot judge how coherent the current industrial structures of municipalities are. Therefore, we create a second baseline ("proportional shares") that assumes that each municipality has a share of the country-wide employment in an industry that is proportional to the municipality's total employment size. This baseline can be viewed as reflecting a random distribution of industrial activities across municipalities.

All match variables are created for three different types of linkages among industries: SR (industries are skill-related), IOR (industries are linked through the value chain) and IOR&SR (industries are both skill-related and linked through the value chain). In Figure 1, we represent these outcomes graphically. The figure shows point estimates of the mean of match variables along with a confidence interval that represents a band width of four times the point estimate's standard error. Values have been recentered such that zero represents the proportional shares baseline.

- Table 2 about here -

- Figure 1 about here -

From Table 2 and Figure 1, it is clear that the current industrial structure baseline is well above the proportional shares baseline for both relatedness types. This strongly suggests that hypotheses 1a and 1b (which state that there is a positive association between the size of a local industry and the degree to which industries related through one of the Marshallian channels are overrepresented in the municipality) find support in the data. The same holds for most of the other employment channels that add employment to a local industry. The employment growth in incumbents and most types of new plant entry occur in industries that match the local industrial profile of a municipality better than the proportional shares baseline. However, the only source of new employment that also outstrips the match of the current industrial structure is the growth of incumbent plants and even then only in terms of labor market linkages. In Table 3, we assess the differences among employment channels by testing whether the match values of different employment channels are statistically different from one another.

- Table 3 about here -

Table 3 reports the p-value of a test that the match variable associated with the source (or baseline) in a row exceeds the match variable associated with the sources in a column. The upper panel contains p-values for match variables based on skill-relatedness, the lower for match variables based on value chain linkages. For instance, the p-value of 0.079 reported in row 2 column 3 of the upper refers to the test that, on average, the existing regional industrial are better embedded within the regional structure (in a labor market sense) than the industries that experience net growth.

In both panels, the upper row contains exclusively high values (mostly ones) and the first column low values (mostly zeros). This shows that the different employment sources are, on average, better embedded in both value chain and skill-relatedness terms than one would expect based on a random distribution of industries, although the evidence is stronger for skill-relatedness.[7] However, taking the present industrial structure as a benchmark, we find that, regardless of whether we use skill-relatedness or value chain linkages to calculate embeddedness, the employment added by any type of new plant is in general worse embedded than the existing employment (rows/columns 9 and 11-18). Another way to look at this is that new plants increase the variety of industrial activities in a municipality. Growing incumbent plants (row/column 3), in contrast, are mostly found in local industries that are slightly better

[7] The only real exception is the employment added because firms from outside the region set up new plants (row/column 18).

embedded than the overall existing activity, at least in terms labor market embeddedness. With a p-value of 9.1%, the difference is weakly significant.

Interestingly, not only the sources of employment growth, but also of employment decline are better embedded than the random, proportional shares baseline (rows/columns 6: net decline of incumbent plants, 8: out-migration of plants, and 10: closure of plants). This finding is, however, not surprising. After all, employment can only decline in a local industry if there the local industry had at least some employment. Given the strong match between the existing industrial structure with itself, declining industries are likely to be to some extent well-embedded in the local economy. If we compare the sources of employment decline with their corresponding sources of employment growth, we find that industries that suffer a net decline in incumbent plant employment are less embedded than the industries that gain incumbent plant employment (row 5, column 6). There is less evidence that in-migrating plants are better embedded than out-migrating plants (row 7, column 8), or that new plants are better embedded than plants that are closed down (row 9, column 10). In fact, for the latter, the situation seems to be reversed, with exiting plants being better embedded than newly established plants, although the difference is not statistically significant.

Turning to the founders of plants, we find some interesting differences among founder types. First, entrepreneurs tend to set up plants in industries that are better embedded in the local economy than the industries chosen by existing firms (row 11, column 12). The evidence for this is strong in terms of skill-relatedness, but somewhat weaker in terms of value chain linkages. Moreover, both local firms and local entrepreneurs choose better embedded industries than firms and regions from outside the region. The evidence on regional firms and entrepreneurs is less clear-cut, although these founders seem to choose industries with intermediate levels of embeddedness.

We can draw some support for our hypotheses from these descriptive analyses. First, the fact that the match of the existing industrial structure with itself is stronger than the match of the random proportional shares baseline supports hypotheses 1a and 1b. Similarly, hypotheses 2b and 2c are supported by the fact that industries that experience a net employment growth are stronger embedded than the proportional shares baseline. Indeed, industries that experience a net growth in incumbent plants are even better embedded than the current local industrial structure, suggesting that the growth of incumbent plants further reinforces the industrial cohesion of a region, both in terms of labor market and value chain linkages. New plants are also typically founded in industries that are better embedded than the proportional shares baseline, lending support to hypotheses 3b and 3c. However, new plants add to the industrial variety of a municipality, given that they are set up in industries that match the current industrial structure less well than the existing economic activity. Exit of and employment decline in existing plants also occurs mostly in industries that are under-embedded in the local economy compared to the existing economic activity. Because shrinking and exiting incumbent plants decrease the employment in less embedded industries, these phenomena tend to reduce the local industrial variety. The evidence regarding hypotheses 4a and 4b suggests that entrepreneurs set up their plants in industries that are better embedded than those chosen by expanding firms. This finding is more in line with the notion that entrepreneurs are heavily dependent on the local environment to complement their own resources than with the idea that entrepreneurs introduce novelty in the region. Novelty seems to be brought in mostly by founders from outside the region. Both entrepreneurs and firms from outside the region are more likely to choose industries with low levels of local embeddedness than local founders are. The evidence for hypotheses 5a and 5b therefore suggests that employment generated by outside founders is less embedded in the regional economy than employment generated by local founders.

Regression analyses

The descriptive analyses already show that both Marshallian channels are important factors for understanding the industrial compositions and changes therein of local economies. In this section we investigate the comparative strength of the channels in the growth of incumbent plants, the formation of new plants and the employment in new plants, henceforth referred to as new plant employment.[8]

For this purpose, we run regression analyses that use the local industry, i.e., the combination of an industry and a municipality, as its unit of analysis. The first set of regressions investigates the growth of incumbent plants. Using ordinary least squares (OLS), we estimate the following main equation:

$$12)\ \log\left(\frac{E_{mi}^{inc,2007}}{E_{mi}^{inc,2004}}\right) = \alpha + \beta_1 \log\left(E_{mi}^{inc,2004}\right) + \beta_2 LQ_{mi}^{SR,2004} + \beta_3 LQ_{mi}^{IOR,2004} + \beta_4 LQ_{mi}^{IOR\&SR,2004}$$

$E_{mi}^{inc,t}$ denotes the employment in incument plants in industry i and municipality m at the beginning of year t. Growth is thus defined as the logarithm of the employment change in incumbent plants between

[8] We think that the entry of plants into a region by relocation of existing plants is qualitatively different from establishing new plants. We therefore omit relocating plants when we analyze new plant formation and employment and focus only on plants that are genuinely new.

2004 and 2007. The embeddedness variables are defined as in equation (6) and are measured at the beginning of 2004. The indicators measure the relative overspecialization of the municipality in industries that are skill-related, input-output related, or skill-related and input-output related to industry i. We also add a variable for the incumbent employment in the year 2004 to measure the effect of the industry's own size on its growth rate.

Next, we estimate the effect of the embeddedness indicators on new plant formation and new plant employment. Because these variables are positive and integer valued, we use count data regression models. Compared to the OLS regressions, a further complication arises. New plants can be formed in local industries without any prior (incumbent) employment. The term $\log\left(E_{mi}^{inc,2004}\right)$ would, therefore, be undefined in these cases. We add a dummy that assumes the value one whenever $E_{mi}^{inc,2004}$ equals zero, which allows us to use all possible local industries.[9] Table 5 provides descriptive statistics of the variables we use in the analyses and Table 6 provides a correlation table.

- Table 4 about here -

- Table 5 about here -

Employment growth in incumbent plants

The outcomes of the first set of analyses are summarized in Table 6. The first column contains only the variable measuring the local industry's employment at the start of the period. In line with hypothesis 2a, incumbent employment growth is negatively associated with the size of the local industry at the start of the period. However, it is unclear whether this effect is due to negative agglomeration externalities incurred from plants in the own industry, or whether this should be simply regarded as a statistical artifact caused by mean reversion. In columns 2 and 3, we add variables for the skill and input-output embeddedness of the local industry. Both variables are strongly and positively associated with employment growth. When we add both embeddedness variables in column (4), we see that the positive association between incumbent growth and value chain embeddedness becomes insignificant. This suggests that the skill embedding of a local industry is more important than the value chain

[9] However, we only investigate new plant formation and incumbent growth in non-resource based export industries. Furthermore, because we omit plants that enter and exit in the same year, the plant formation data refer to plants that are founded during the years 2004 and 2005 and that are still active at the beginning of the years 2006 and 2007 respectively.

embedding. This finding supports hypothesis 3b, but it fails to support hypothesis 3c. In column (5), we test our conjecture that the value chain and the labor market channels reinforce each other. The estimate of the variable that captures a local industry's simultaneous skill and value chain embeddedness is not significant. In column (6), we exclude the employment in industries that are at the same time skill and value chain linked to a local industry from the calculations of the skill embeddedness and the value chain embeddedness indicators. Outcomes are by and large unaffected by this redefinition of variables.

- Table 6 about here -

New plant formation

Table 7 presents the outcomes of negative binomial regressions for the number of newly established plants in a local industry. Column (1) contains only the control variables for the local industry's size at the beginning of the year 2004. In line with hypothesis 3c, but in contrast to our findings for incumbent plant growth, we observe that the size of the local industry at the start of the period positively affects new plant formation. The reported point estimates are incidence rate ratios and values over one represent positive associations. The prior presence of a local industry increases new plant formation by a factor of almost 4.5. The elasticity with respect to the industry's size is $\ln(1.7) = 53\%$. In other words, for every doubling of the employment, we expect a rise in new plant formation of 53%.

Columns (2)-(5) add the various embeddedness indicators. Surprisingly, the skill embeddedness of a local industry shows no statistically significant association with new plant formation. What is more, the value chain embeddedness is negatively associated with the number of newly established plants. Only the embeddedness in terms of the overspecialization in industries that maintain both value chain and labor market linkages with a local industry is positively associated with new plant formation in a local industry. The redefinition of mutually exclusive channels in column (6) provides somewhat puzzling results. The labor market channel is positively associated with new plant formation, whereas the associations with the value chain channel and the combination of both channels are insignificant. Our overall conclusion regarding the hypotheses on new plant formation are, therefore, somewhat mixed. Hypothesis 3a (plant formation is positively associated with the local industry's size) finds strong support. However, hypothesis 3b on the role of skill embeddedness finds weak support at best, whereas hypothesis 3c on the role of value chain embeddedness is even rejected in some specifications.

- Table 7 about here -

New plant employment

Table 8 contains the estimates of negative binomial regression models for new plant employment. Column (1) shows that, like new plant formation, new plant employment positively depends on the prior presence of the local industries in which the new plants are active. Columns (3) and (4) show that skill embeddedness is significantly and positively associated with new plant employment, whereas the association between new plant employment and value chain embeddedness is negative, but insignificant. When we control for the degree to which local industries are simultaneously labor market and value chain embedded in column (5), this negative association turns significant. The effect of the value-chain-cum-labor-market embeddedness is positive and weakly significant with a p-value of 7.5%. The estimates in column (6) confirm these findings. If we use the mutually exclusively redefined embeddedness indicators for the labor market and value chain embeddedness, the association of new plant employment with the combined embeddedness variable is positive and significant, whereas the value chain embeddedness shows a negative, yet insignificant association. The association of new plant employment with labor market embeddedness is throughout all models positive and significant.

- Table 8 about here -

Discussion

Taking stock of the evidence, we find strong support for hypotheses 1a and 1b that the size of local industries is positively associated with the degree to which industries with which the local industries can share labor and maintain value chain linkages are overrepresented in the local economy. This evidence is in line with a finding in Neffke et al. (2011a) that local economies are by and large coherent. In this paper, we find further evidence for this coherence of local economies, both in a value chain and a labor market sense. This finding lends support for Porter's (2003) method of using co-location of industries to assess which industries belong to the same cluster.

Our findings on growth and new plant formation show that, in line with hypotheses 2a and 3a, incumbent growth is negatively associated with an industry's local size, whereas new plant formation and new plant employment are both positively associated with the size of a local industry. The explanation for the latter positive associations is that new plants are often set up by an industry's existing local actors. The larger the local industry, therefore, the more potential new plant founders there are. The association of growth rates and new plant employment with the local embeddedness is less clear-cut. Hypotheses 2b and 2c and 3b and 3c respectively state that incumbent growth rates and new plant employment are positively associated with both the labor market and value chain embeddedness of a local industry. We find strong statistical support for the positive association of labor market embeddedness with incumbent growth rates and with the amount of employment in new plants (hypotheses 2b and 3b). The associations for value chain embeddedness are less pronounced: although they are positive for incumbent growth rates, they are insignificant in most of the models. Moreover, in new plant employment regressions we find a negative association with the value chain embeddedness of a local industry. Therefore we fail to find support for hypothesis 2c and would tend to reject hypothesis 3c. Hypotheses 6a and 6b propose that the value chain and labor market channels reinforce each other. Although we find no support for this hypothesis in the incumbent growth rates, we do find a significant positive association of the combined labor market and value chain embeddedness with new plant employment. Overall, this leads us to a tentative, yet thought-provoking conclusion: value chain linkages among industries in a local economy seem to be irrelevant, unless the industries also share the same labor pool.

Turning to our descriptive analyses of the differences among founder types (hypotheses 4a, 4b 5a and 5b), we conclude that entrepreneurs set up plants in local industries that are in general more embedded than the ones chosen by existing firms. The same holds for local actors compared to actors from outside the region. This means that the new plants of existing firms and outside actors add most novelty to a local economy. Existing firms and founders from outside the region may, therefore, be crucial for the capacity of a region to renew its economy. Although our findings are in line with an extensive literature that argues that attracting outside actors to a region is important to overcome regional lock-ins (Grabher 1993), the renewal potential that is often attributed to entrepreneurs is less evident in our analyses. However, to give a final verdict on these questions, it is important to follow new plants for a longer period. After all, the potential for renewal will only materialize if plants survive for longer periods and steadily increase their employment.

5. Conclusion

In this article we examined new plant formation and growth of local industries. A number of different bodies of research have recently stressed the interdependences among industries along a number of different dimensions. We focused our analyses on two Marshallian channels through which firms benefit from each other's proximity, the possibility to share the same labor and value chain linkages. We found that the labor market channel seems to be the most important one. Incumbent plants show faster growth rates and a higher employment in new plants in industries that are located in municipalities with an overspecialization in industries that have similar labor requirements to the plants' own industries. The overspecialization of value chain related industries is, in contrast, mostly either statistically insignificantly or even negatively associated with growth rates and new plant employment. There is some evidence that, if industries are at the same time related in a labor market sense, an overspecialization in value chain linked industries matters for the formation of and employment in new plants. If this finding can be replicated, it is an important piece of information about the workings of local clusters, which may proof useful for local policy making. Accordingly, the existence of input-output linkages between local industries is insufficient to increase the strength of a cluster. The proximity between client and supplier firms is crucial only if the firms are also able to share some of their labor.

A further finding is that different types of founders choose different industries in which they set up new plants. Entrepreneurs and local actors set up plants particularly often in industries that are well embedded both in a labor market and in a value chain sense. This suggests that structural change and the creation of novel activities in a region are strongly dependent on an influx of plants of existing firms and set up by founders coming from outside the region.

We want to point out some important caveats. First, we have not the ambition to uncover causal relationships. Finding instruments for estimating causal effects on local industry specific growth and plant formation rates would be an important way to improve upon the research we presented. Second, the differences between founder types and the long run effects of plants set up by different founders cannot be fully established in this article, because of the short time span covered by our data. However, the analyses can be easily extended by following newly established plants over a longer period. Such research should be able to answer a pivotal question in the field of economic geography, namely, which actors are responsible for the long-run transformational capacity of a local economy.

References

Almeida P., Kogut B. (1999). Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science*, 45(7): 905-991.

Arrow K. J., (1962). Economic welfare and the al- location of resources for invention. In National Bureau and Economic Research, *The Rate and Direction of Inventive Activity*: 609-625. Princeton, NJ: Princeton University Press.

Boschma R.A., Frenken K. (2009) Technological relatedness and regional branching, in: Bathelt H., Feldman M.P., Kogler D.F.(eds.), *Dynamic Geographies of Knowledge Creation and Innovation*, Routledge, Taylor and Francis, forthcoming.

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

Cooke P.H., Morgan K. (1998). *The Associational Economy. Firms, Regions, and Innovation*. Oxford: Oxford University Press.

Dauth, W. (2010). The mysteries of the trade: employment effects of urban interindustry spillovers. *IAB Discussion Paper* 15/2010: 1-25.

De Groot H., Poot J., Smit M.J. (2009). Agglomeration Externalities, Innovation and Regional Growth: Theoretical Perspectives and Meta-Analysis. In: *Handbook of Regional Growth and Development Theories*, Capello R., Nijkamp P. (eds.). Cheltenham: Edward Elgar.

Delgado M., Porter M. E., Stern S. (2010). Clusters and entrepreneurship. *Journal of Economic Geography* 10(4): 495-518.

Diamond C. A., Simon C. J. (1990). Industrial Specialization and the Returns to Labor. *Journal of Labor Economics*, 8(2): 175-201.

Ellison G., Glaeser E. L. (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy* 105(5): 889-927.

Ellison G., Glaeser E. L., Kerr W. R. (2010) What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review*, 100(3): 1195–1213

Florida R., Mellander C., Stolarick K. (2011). Geographies of scope: an empirical analysis of entertainment, 1970–2000. *Journal of Economic Geography*. In press.

Frenken K., Boschma R.A. (2011), Economic development as a branching process, in: Buenstorf G. (ed.), Cheltenham: Edward Elgar, in print.

Frenken K., Van Oort F.G., Verburg T., (2007). Related Variety, Unrelated Variety and Regional Economic Growth. *Regional Studies* 41(5): 685-697.

Gertler M. S. (2003). Tacit knowledge and the economic geography of context, or The undefinable tacitness of being (there). *Journal of Economic Geography* 3(1): 75-99.

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: *The Embedded Firm. On the Socioeconomics of Interfirm Relations*. Grabher G., (ed.). London/New York: Routledge.

Henderson J.V., Kuncoro A., Turner M. (1995). Industrial Development in Cities. *Journal of Political Economy* 103(5): 1067-1090.

Henderson J.V. (2003). Marshall's Scale Economies. *Journal of Urban Economics* 53: 1-28.

Hidalgo C.A., 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: Random House.

Marshall A. (1920). *Principles of Economics*. 8th edition. London: Macmillan and Co. Original edition 1890.

Martin, Ph., Ottaviano, G. I. P. (2001). Growth and Agglomeration. *International Economic Review* 42(4): 1468-2354.

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

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.

Faggian A. McCann Ph. (2006). Human capital flows and regional knowledge assets: a simultaneous equation approach. *Oxford Economic Papers* 52: 475–500.

McCann Ph., Simonen J. (2005). Innovation, knowledge spillovers and local labour markets. *Papers in Regional Science* 84(3): 465-485.

Neffke F.M.H., Henning M. (2011). Skill-relatedness and Firm diversification. *Strategic Management Journal*. Forthcoming.

Neffke F.M.H., Henning M., Boschma R.A. (2011a). How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. *Economic Geography* 87(3): 237–265.

Neffke F.M.H., Henning M., Boschma R.A. (2011b). The impact of aging and technological relatedness on agglomeration externalities: a survival analysis. *Journal of Economic Geography* (in press).

Nooteboom B., (2000). *Learning and Innovation in Organizations and Economies*. Oxford: Oxford University Press.

Polanyi M., (1966). *The Tacit Dimension*. London: Routledge.

Porter M. (2010). The Economic Performance of Regions. *Regional Studies* 37(6-7): 549–578.

Porter, M.E. (1998). Clusters and the new economics of competition. *Harvard Business Review* 76 (November-December): 77-90.

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.

Sorenson O. (2003). Social networks and industrial geography. *Journal of Evolutionary Economics* 13: 513–527.

Staber U. (2005) Entrepreneurship as a source of path dependency, in Fuchs G. and Shapira P. (Eds) *Rethinking regional innovation and change: path dependency or regional breakthrough*. Springer, New York.

Teece D.J. (1982). Towards an Economic Theory of the Multiproduct Firm. *Journal of Economic Behavior and Organization* 3: 39-63.

Table 1: Sources of local industry employment changes

Baselines		**Average yearly employment**
Proportional shares	fictitious employment shares that would result if the employment in each industry would mimic the spatial distribution of total employment	1,239,413
Current industrial structure	industries' shares of existing employment in the region	1,239,413
Sources of employment change (aggregate categories)		
Total growth	net employment growth in local industries, 0 if negative	91,496
Total decline	net employment decline in local industries, 0 if negative	78,189
Incumbent growth	net employment growth in incumbent plants, 0 if negative	51,077
Incumbent decline	net employment decline in incumbent plants, 0 if negative	40,238
In-migration	total employment of plants relocating into the municipality	11,630
Out-migration	total employment of plants relocating out of the municipality	11,719
Entry by new plants	total employment in newly established plants	29,857
Exit existing plants	total employment in plants that are shut down	27,897
Sources of employment change by founder type		
Entrepreneur	total employment in plants set up by entrepreneurs	23,580
Firm expansion	total employment in plants set up by existing firms	6,278
Local entrepreneur	total employment in plants set up by local entrepreneurs (distance to entrepreneur's former municipality: <50 km)	19,113
Regional entrepreneur	total employment in plants set up by regional entrepreneurs (distance to entrepreneur's former municipality: 50-250 km)	2,319
Outside entrepreneur	total employment in plants set up by regional entrepreneurs (distance to entrepreneur's former municipality: >250 km)	2,148
Local firm	total employment in plants set up by local firms (distance to parent firm's municipality: <50 km)	3,099
Regional firm	total employment in plants set up by regional firms (distance to parent firm's municipality: 50-250 km)	966
Outside firm	total employment in plants set up by regional firms (distance to parent firm's municipality: >250 km)	2,213

Distance is the road distance between the main population centers of two municipalities and is derived from Google maps.

Table 2: Match of employment source and local economy

	Labor market		Input-output		# obs
Source	**mean**	**std. dev.**	**mean**	**std. dev.**	
Proportional shares	-0.1120	0.079	0.0105	0.119	1,160
Current industrial structure	0.0661	0.145	0.0619	0.145	1,160
Total growth	0.0623	0.152	0.0582	0.151	870
Total decline	0.0397	0.163	0.0496	0.152	870
Incumbent growth	0.0705	0.153	0.0605	0.157	869
Incumbent decline	0.0487	0.168	0.0493	0.168	869
In-migration	0.0014	0.209	0.0179	0.181	777
Out-migration	-0.0042	0.209	0.0150	0.174	801
Entry by new plants	0.0696	0.188	0.0805	0.216	834
Exit existing plants	0.0573	0.196	0.0701	0.216	836
Entrepreneur	0.0185	0.185	0.0339	0.147	577
Firm expansion	-0.0030	0.235	-0.0038	0.252	327
Local entrepreneur	0.0256	0.186	0.0390	0.159	572
Regional entrepreneur	0.0003	0.211	0.0104	0.208	464
Outside entrepreneur	0.0092	0.215	0.0147	0.183	395
Local firm	0.0506	0.218	0.0211	0.241	163
Regional firm	-0.0061	0.241	0.0130	0.238	132
Outside firm	-0.0310	0.229	-0.0357	0.251	204

Employment weighted average location quotient of employment in related industries (left: skill-related, right: input-output related) by source.

Table 3: p-values for tests of differences between the matches of employment sources

MATCH IN TERMS OF OVER-SPECIALIZATION IN SKILL-RELATED INDUSTRIES

TEST ROW>COLUMN	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(1) Proportional shares	.	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.997
(2) Current industrial structure	0.000	.	0.079	0.000	0.909	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.080	0.001	0.000
(3) Total growth	0.000	0.921	.	0.000														
(4) Total decline	0.000	1.000	1.000	.														
(5) Incumbent growth	0.000	0.091			.	0.000												
(6) Incumbent decline	0.000	1.000			1.000	.												
(7) In-migration	0.000	1.000					.	0.264										
(8) Out-migration	0.000	1.000					0.737	.										
(9) Entry by new plants	0.000	1.000							.	0.591								
(10) Exit existing plants	0.000	1.000							0.410	.								
(11) Entrepreneur	0.000	1.000									.	0.009						
(12) Firm expansion	0.000	1.000									0.991	.						
(13) Local entrepreneur	0.000	1.000											.	0.000	0.000			
(14) Regional entrepreneur	0.000	1.000											1.000	.	0.216			
(15) Outside entrepreneur	0.000	1.000											1.000	0.784	.			
(16) Local firm	0.000	0.920														.	0.099	0.001
(17) Regional firm	0.001	0.999														0.901	.	0.002
(18) Outside firm	0.003	1.000														0.999	0.998	.

MATCH IN TERMS OF OVER-SPECIALIZATION IN VALUE CHAIN LINKED INDUSTRIES

TEST ROW>COLUMN	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(1) Proportional shares	.	1.000	1.000	1.000	1.000	1.000	0.940	0.882	1.000	1.000	1.000	0.311	1.000	0.687	0.378	0.983	0.774	0.005
(2) Current industrial structure	0.000	.	0.184	0.001	0.402	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.030	0.249	0.000
(3) Total growth	0.000	0.816	.	0.044														
(4) Total decline	0.000	0.999	0.956	.														
(5) Incumbent growth	0.000	0.598			.	0.030												
(6) Incumbent decline	0.000	0.998			0.970	.												
(7) In-migration	0.060	1.000					.	0.433										
(8) Out-migration	0.118	1.000					0.567	.										
(9) Entry by new plants	0.000	1.000							.	0.786								
(10) Exit existing plants	0.000	1.000							0.214	.								
(11) Entrepreneur	0.000	1.000									.	0.054						
(12) Firm expansion	0.689	1.000									0.946	.						
(13) Local entrepreneur	0.000	1.000											.	0.008	0.001			
(14) Regional entrepreneur	0.313	1.000											0.993	.	0.278			
(15) Outside entrepreneur	0.622	1.000											0.999	0.722	.			
(16) Local firm	0.017	0.970														.	0.681	0.041
(17) Regional firm	0.226	0.751														0.320	.	0.033
(18) Outside firm	0.995	1.000														0.959	0.967	.

Table 4: Descriptive statistics

	mean	std. dev.	min	max
incumbent growth	0.0001	0.1445	-3.5264	3.0445
# new plants	0.9265	8.5104	0.0000	1,341.00
new plant employment	2.6488	26.9094	0.0000	3,019.00
presence dummy	0.4753	0.4994	0.0000	1.0000
log(emp(m,i))	1.1520	1.7263	0.0000	9.9869
skill-embeddedness	-0.0656	0.2192	-1.0000	0.9295
IO-embeddedness	-0.0612	0.3113	-1.0000	0.9951
skill&IO-embeddedness	-0.1867	0.4376	-1.0000	0.9915
skill-embeddedness (ex)	-0.0914	0.2355	-1.0000	0.9295
IO-embeddedness (ex)	-0.0840	0.3310	-1.0000	0.9951

Table 5: Correlations among regressors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) incumbent growth	1.00									
(2) # new plants	0.00	1.00								
(3) new plant employment	-0.01	0.69	1.00							
(4) presence dummy	0.00	0.12	0.10	1.00						
(5) log(emp(m,i))	-0.03	0.24	0.23	0.72	1.00					
(6) skill-embeddedness	0.00	0.04	0.04	0.14	0.16	1.00				
(7) IO-embeddedness	0.00	0.02	0.01	0.05	0.07	0.22	1.00			
(8) skill&IO-embeddedness	0.00	0.05	0.04	0.19	0.20	0.47	0.54	1.00		
(9) skill-embeddedness (ex)	0.01	0.05	0.04	0.12	0.13	0.77	-0.03	0.08	1.00	
(10) IO-embeddedness (ex)	0.00	0.00	0.01	0.00	0.01	-0.01	0.79	0.17	-0.08	1.00

Table 6: OLS regressions of the logarithm of incumbent employment growth on embeddedness indicators

INCUMBENT GROWTH	(1)	(2)	(3)	(4)	(5)	(6)
log(emp(m,i))	-0.017***	-0.019***	-0.017***	-0.019***	-0.019***	-0.019***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
skill-embeddedness		0.073***		0.067***	0.071***	
		(0.000)		(0.000)	(0.000)	
IO-embeddedness			0.030***	0.016	0.021	
			(0.003)	(0.129)	(0.084)	
skill-embeddedness (ex)						0.061***
						(0.000)
IO-embeddedness (ex)						0.028***
						(0.005)
skill&IO-embeddedness					-0.007	0.013
					(0.473)	(0.075)
constant	0.055***	0.063***	0.058***	0.063***	0.063***	0.068***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Nobs	17,577	17,577	17,577	17,577	17,577	17,397
Rsq	0.0064	0.008	0.0069	0.0081	0.0082	0.0085

***: p<.01, **: p<.025 *: p<.05. p-values are provided in parentheses beneath the point estimates.

Table 7: Negative binomial regressions of the number of new plants established in a local industry on embeddedness indicators

NEW PLANT ENTRY	(1)	(2)	(3)	(4)	(5)	(6)
industry presence (y/n)	4.46***	4.46***	4.43***	4.44***	4.39***	4.49***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log(emp(m,i))	1.70***	1.70***	1.71***	1.70***	1.70***	1.67***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
skill-embeddedness		0.99		1.05	0.87	
		(0.918)		(0.610)	(0.220)	
IO-embeddedness			0.84**	0.83***	0.67***	
			(0.015)	(0.007)	(0.000)	
skill-embeddedness (ex)						1.55***
						(0.000)
IO-embeddedness (ex)						1.02
						(0.769)
skill&IO-embeddedness					1.34***	1.05
					(0.000)	(0.288)
alpha	3.56	3.56	3.55	3.56	3.55	3.52
Nobs	57,420	57,420	57,420	57,420	57,420	56,260
pseudo log likelihood	-23,106.5	-23,106.5	-23,101.8	-23,101.6	-23,087.2	-22,864.8

***: p<.01, **: p<.025 *: p<.05. p-values are provided in parentheses beneath the point estimates of incidence rate ratios.

Table 8: Negative binomial regressions of the employment in new plant in a local industry on embeddedness indicators

NEW PLANT EMPLOYMENT	(1)	(2)	(3)	(4)	(5)	(6)
industry presence (y/n)	2.93***	3.15***	2.92***	3.15***	3.13***	3.15***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log(emp(m,i))	1.88***	1.81***	1.89***	1.82***	1.81***	1.79***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
skill-embeddedness		3.32***		3.58***	3.18***	
		(0.000)		(0.000)	(0.000)	
IO-embeddedness			0.85	0.71	0.62***	
			(0.393)	(0.051)	(0.008)	
skill-embeddedness (ex)						3.72***
						(0.000)
IO-embeddedness (ex)						0.91
						(0.547)
skill&IO-embeddedness					1.22	1.26**
					(0.075)	(0.012)
alpha	11.55	11.33	11.55	11.30	11.29	11.04
Nobs	57,420	57,420	57,420	57,420	57,420	56,260
pseudo log likelihood	-30,506.2	-30,391.5	-30,502.8	-30,377.6	-30,371.8	-30,011.6

***: p<.01, **: p<.025 *: p<.05. p-values are provided in parentheses beneath the point estimates of incidence rate ratios.

Figure 1: Match of employment source and local economy