CLUSTERS AND ENTREPRENEURSHIP

by

Mercedes Delgado *
Temple University

Michael E. Porter *
Harvard University

and

Scott Stern *
MIT and the NBER

Abstract

This paper examines the role of regional clusters in regional entrepreneurship. We focus on the distinct influences of convergence and agglomeration on growth in the number of start-up firms as well as in employment in these new firms in a given region-industry. While reversion to the mean and diminishing returns to entrepreneurship at the region-industry level can result in a convergence effect, the presence of complementary economic activity creates externalities that enhance incentives and reduce barriers for new business creation. Clusters are a particularly important way through which location-based complementarities are realized. The empirical analysis uses a novel panel dataset from the Longitudinal Business Database of the Census Bureau and the U.S. Cluster Mapping Project (Porter, 2003). Using this dataset, there is significant evidence of the positive impact of clusters on entrepreneurship. After controlling for convergence in start-up activity at the region-industry level, industries located in regions with strong clusters (i.e. a large presence of other related industries) experience higher growth in new business formation and start-up employment. Strong clusters are also associated with the formation of new establishments of existing firms, thus influencing the location decision of multiestablishment firms. Finally, strong clusters contribute to start-up firm survival.

Keywords: Entrepreneurship, Industry Clusters, Dynamic Economies of Agglomeration.

1. Introduction

In this paper, we examine the impact of agglomeration on entrepreneurship at the regional industry level. In particular, we focus on the role of clusters, or agglomerations of closely related industries, in new business formation. Large variations in regional employment growth and in the rate of new firm creation are a striking feature of the US economy (Armington and Acs, 2002; Porter, 2003). While a significant body of work explores why some regions experience more rapid growth than others (Porter, 1990, 1998a; Saxenian, 1994; Glaeser et al., 1992; Barro and Sala-i-Martin, 1995; Fujita, Venables, and Krugman, 1999), there is increasing academic and policy interest in the particular role in growth played by entrepreneurship (Davis et al., 1996; Haltiwanger, et al., 2009). A significant debate is underway regarding the role of the regional economic environment in shaping differences in the rate of regional entrepreneurship and overall economic performance (Porter, 1990, 1998a; Saxenian, 1994; Feldman, 2001; Glaeser and Kerr, 2009).

Any effort to explain region-industry performance growth must account for two countervailing economic forces: convergence and agglomeration (Delgado, Porter, and Stern, 2007). Convergence arises when, due to diminishing returns, the potential for growth is *declining* in the level of economic activity (Barro and Sala-i-Martin, 1991). Agglomeration exerts an opposite force on regional performance. In the presence of agglomeration economies, the potential for growth is *increasing* in the level of economic activity (Glaeser et al., 1992). From an empirical perspective, distinguishing the differential impact of convergence and agglomeration has been problematic. For example, if both convergence and agglomeration effects are present, the impact of the initial level of economic activity on growth will reflect a balancing of the two effects.

We move beyond this traditional impasse by identifying the impact of agglomeration while simultaneously accounting for the impact of convergence. Our key insight is that while forces that give rise to both convergence and agglomeration operate *within* narrow economic units, agglomeration *across* complementary economic units can have a separate and distinctive impact. Building on cluster theory, we focus specifically on economies of agglomeration that arise within clusters of complementary industries related by technology, skills, shared infrastructure, demand and other linkages. The

2

complex interactions that take place within clusters suggest a useful focus on the presence (or absence) of supporting and related industries rather than diversity of industries per se (Porter, 1990, 1998a, 2003; Feldman and Audretsch, 1999; Delgado, Porter and Stern, 2007; Frenken et al., 2007).

This paper examines a particularly important channel through which cluster-driven agglomeration might operate: entrepreneurship. The presence of a cluster of related industries in a location will foster entrepreneurship by lowering the cost of starting a business, enhancing opportunities for innovations, and enabling better access to a more diverse range of inputs and complementary products. (Saxenian, 1994; Porter, 1998a; Feldman, Francis, and Bercovitz, 2005; Glaeser and Kerr 2009). The co-location of companies, customers, suppliers, and other institutions also increases the perception of innovation opportunities while amplifying the pressure to innovate (Porter, 2000). Since entrepreneurs are essential agents of innovation, a strong cluster environment should foster entrepreneurial activity.

The empirical analysis exploits the Longitudinal Business Database (LBD) of the Census Bureau and the cluster definitions from the US Cluster Mapping Project (Porter, 2003). This classification system defines clusters as collections of industries with high levels of co-location in terms of employment. We consider several related measures of the cluster environment surrounding a region-industry, including a measure based on individual clusters, a more encompassing measures that incorporates linkages among related clusters (i.e. "linked" clusters), and a third measure that captures the strength of similar clusters in neighboring regions. We measure entrepreneurship and industrial composition utilizing a dataset that spans the years 1990-2005, includes 177 mutually exclusive Economic Areas (EAs) in the contiguous United States, and incorporates 588 "traded" industries constituting 41 clusters for each EA.[1]

Our empirical work focuses on early stage entrepreneurship, which we measure using two related indicators of start-up activity: the number of new establishments of new firms in a region within a given traded industry (which we refer to as the level of *start-up establishments*); and the employment in these new firms (which we refer to as

---

[1] Traded industries are those that sell products and services across regional and national boundaries. See Porter (2003).

the level of *start-up employment*). We compute the growth rate in start-up establishments and start-up employment in each regional industry, and estimate the impact of the extent of cluster-driven agglomeration while accounting for the impact of convergence. Our core specifications incorporate region and industry fixed effects.

We find striking evidence for the simultaneous yet distinct influences of agglomeration and convergence on the growth rate of start-up establishments and employment. Growth in start-up employment at the region-industry level is declining in the initial level of start-up employment at the region-industry level, consistent with the presence of a convergence effect in start-up activity due to either mean reversion and/or to the potential for local crowding-out of entrepreneurial opportunities within a regional industry. At the same time, however, growth in start-up employment is increasing in measures of the cluster environment surrounding the region-industry. We obtain similar findings for the growth of start-up establishments. By accounting for convergence and the potential for competition within each regional industry, we are able to isolate the positive impact of cluster-related complementarities on entrepreneurship. The results provide support for the hypothesis that strong clusters facilitate growth in entrepreneurship.

While our main focus is the formation of start-up firms, we also look at new establishments of existing firms. We find that stronger clusters are also associated with greater formation of new establishments of existing firms. While more research is needed on the locational decisions of multi-location (multi-establishment and multinational) firms, such firms seem to be opening establishments in locations where strong clusters are present (Enright, 2000; Bresnahan and Gambardella, 2004; Manning 2008). The contribution of established firms to regional entrepreneurship is an important question for future work.

Finally, we examine the role of clusters in the medium-term performance of start-ups. A strong cluster should increase the productivity of participating firms, which raises the bar for survival of new businesses. At the same time, a strong regional cluster can facilitate the growth of start-up firms by providing better access to the necessary inputs to commercialize products and services. We find that a strong cluster environment

improves the level of employment in young (up to 5 year old) start-up firms in a region-industry, suggesting that clusters facilitate medium-term survival of start-ups.

The rest of the paper is organized as follows. We begin by discussing the relationship between entrepreneurship and cluster-driven agglomeration, and develop the main hypotheses. Section 3 presents the empirical framework. Section 4 explains the data, and Section 5 discusses the main findings. A final section concludes.

2.   Clusters and entrepreneurship

Numerous mechanisms are associated with entry of new businesses in agglomerated areas. Starting with Marshall (1920), regional studies have highlighted at least three distinct drivers of agglomeration: knowledge spillovers, input-output linkages, and labor market pooling. Over time, an extensive literature has also incorporated additional agglomeration drivers, including local demand characteristics, specialized institutions and the structure of regional business and social networks (Porter, 1990, 1998a, 2000; Saxenian, 1994; Markusen, 1996). While most empirical studies of agglomeration focus on variables such as the overall employment growth, an emerging literature emphasizes the role of new businesses in regional economic growth (Feldman et al., 2005; Acs and Armington, 2006; Haltiwanger et al., 2009; Glaeser and Kerr, 2009). Relative to incumbent firms, entrepreneurs may be more likely to identify opportunities (both in the form of new technologies and new markets) that exploit distinctive sources of regional comparative advantage. However, there is a high rate of churning of young firms in most countries and sectors (Dunne et al., 1988, 2005; Barteisman et al., 2005; Kerr and Nanda, 2009). Start-up firms have greater exit rates than established firms since they often lack experience and resources. Importantly, however, those start-up firms that survive tend to have greater growth potential than established firms. The presence of a strong cluster environment, which reduces barriers to entry and growth and enhances regional comparative advantage, should be a central driver of entrepreneurial vitality.

The precise mechanisms by which the regional economic environment impacts entrepreneurship are numerous and subtle. Chinitz (1961) hypothesizes that a key requirement for entrepreneurship is the presence of a network of smaller suppliers, and attributes differences in the rate of entrepreneurship between New York and Pittsburgh at

that time to differences in the structure of suppliers. Building on earlier studies, a rich (though mostly qualitative) literature has emerged examining the relationship between entrepreneurship and regional economic performance. For example, Saxenian (1994) attributes the success of Silicon Valley to the culture of entrepreneurship and a more decentralized organization of production (relative to Route 128). An extensive literature also highlights the broader relationship between entrepreneurship and the regional innovation system (e.g. Audretsch, 1995; Feldman, 2001; Acs et al., 2009). Recently, Glaeser and Kerr (2009) test for the impact of specific Marshallian economies of agglomeration on new firm entry; while their analysis does not specifically evaluate the impact of clusters, they provide complementary evidence that the presence of small suppliers and workers in relevant occupations is associated with a higher level of new business creation.

At the same time, a small but growing literature within regional and international business studies examines key interactions between the internal organization of the firm (start-ups, multi-location, small or large, young or old, etc.) and the agglomeration benefits realized within a geographical location. One dimension that has received special attention is the role of small (and younger) firms in extracting and generating economies of agglomeration, and spurring additional new business creation (Henderson, 2003; Rosenthal and Strange, 2003, 2009; Glaeser, Kerr and Ponzetto, 2009). In contrast, other regional and cluster studies highlight the importance of so-called "anchor" firms (including multi-establishment and multinationals) that induce spin-offs and attract firms from related industries (Enright, 2000; Agrawal and Cockburn, 2002; Klepper, 2007; Greenstone et al., 2008). Further research is needed to understand the locational strategy of multi-location firms and their contribution to entrepreneurship in a particular region.

The main goal of this paper is to test whether the presence of related economic activity in a region facilitates the growth of start-up establishments and start-up employment in regional industries. There are numerous channels that may facilitate economies of agglomeration (see the review by Rosenthal and Strange, 2004). Here we suggest that important agglomeration forces are due to the presence of clusters -- supporting and related industries -- instead of industry diversity per se (Porter, 1990, 1998a; Feldman and Audretsch, 1999; Delgado, Porter and Stern, 2007). Thus, we use

Porter's (2003) empirical cluster framework to explore agglomeration across sets of industries related by the full array of technology, skills, demand, or other linkages.

Our first hypothesis focuses on the relationship between the growth of start-up activity and the initial level of start-up activity within a region-industry. We expect a convergence effect (i.e. the coefficient on the initial *level* of start-up activity will be negative). This convergence in start-up activity can be interpreted in terms of mean-reversion (arising even from measurement error) and diminishing marginal returns to entrepreneurial opportunities arising from crowding-out effects.[2] Mean reversion implies that a region-industry that has a relatively high level of start-up activity at t_0 (compared to the average start-up activity in the industry in other regions with similar size and economic composition) is more likely to experience a lower (stochastically determined) growth rate of start-up activity between t_0 and t_1 (Barro and Sala-i-Martin, 1991; Quah, 1996). At the same time, it is possible that the returns to entrepreneurial activity can be diminishing in the level of entrepreneurial activity as the result of input scarcity. For example, if the price of specialized (labor or capital) inputs is increasing in the intensity of competition among start-up firms, there will be diminishing returns to entrepreneurship as a result of congestion costs (Sorenson and Audia, 2000). As a result, a high level of entrepreneurship in a region-industry at a point in time may result in diminished near-term opportunities for entrepreneurship in that region-industry. Thus, *our first hypothesis is that there will be convergence in entrepreneurship at the narrowest unit of analysis: the region-industry growth rate of start-up activity will be declining in the initial level of region-industry start-up activity.*[3]

Our remaining hypotheses focus on the impact of related economic activity on the growth rate of start-up activity. Conditional on the convergence effect, the relationship between related economic activity and entrepreneurship will depend on how the presence of particular types of economic activity impacts entrepreneurial incentives. The

---

[2]We draw on the convergence concept used by the cross-sectional growth literature to study economic activity across countries, regions and regional industries (Barro and Sala-i-Martin, 1991; Henderson et al., 1995; Bostic et al., 1997). However, the analysis of convergence is broader, and has historically explored firm growth based on Gibrat's Law (Geroski and Gugler, 2004; Sutton, 1997). It is useful to note that the presence of a negative relationship between the start-up activity growth rate and its level is consistent with convergence to a limiting distribution of start-up activity growth rates.

[3] While we test for convergence at the region-industry level, convergence may also operate at the industry and region levels. In this paper we use region and industry fixed effects to control for these issues.

incentives for start-up entry in a particular location will reflect the nature of competition (Porter, 1980; Bresnahan and Reiss, 1991). While the returns to entrepreneurship are lower in the presence of intense price-based competition, rivalry may also lead to higher pressure to innovate and greater entrepreneurial incentives. At the same time, the presence of complementary economic activity – specialized suppliers, a local customer base, producers of complementary products and services – increases the pool of inputs available and enhances the range of profitable entry opportunities and so improves entrepreneurial incentives. The empirical relationship between entrepreneurship and particular types of pre-existing economic activity will therefore depend on whether these activities are substitutes or complements (Bulow et al., 1985).

It is useful to distinguish, then, between the level of specialization of a region in a particular industry and the strength of the cluster environment around that industry. On the one hand, the relationship between industry specialization and entrepreneurship growth is ambiguous. While industry specialization in a location may enhance opportunities for learning, innovation, and entrepreneurial spawning (Audretsch, 1995; Gompers et al., 2005; Glaeser and Kerr, 2009), a large presence of established firms (relative to the size of the national industry) intensifies local competition, dampening incentives for entrepreneurial entry. *Our second hypothesis is, then, that the empirical relationship between industry specialization and the growth rate of entrepreneurship in that industry is ambiguous, and will depend on the precise nature of competition (cost-based or innovation-based) and the pattern of strategic interaction between entrant and established firms.[4]*

In contrast, a strong cluster environment surrounding a particular region-industry enhances the incentives and potential for entrepreneurship. The firms within a geographically concentrated cluster share common technologies, skills, knowledge, inputs, consumers, and institutions, facilitating agglomeration across complementary and related industries. A strong cluster environment enhances growth at the region-industry level by facilitating operational efficiency and raising the returns to business expansion,

---

[4] In the paper, we do not test how the nature of competition affects start-up activity. Instead, we test whether complementarities with established firms (which would yield a positive relationship between industry specialization and start-up activity) outweigh the potential for (priced-based) competition between start-ups and established firms (yielding a negative effect).

capital investment, and innovation, thereby increasing job creation and productivity (see e.g. Porter, 1990, 1998a; Saxenian, 1994; Feldman and Audretsch, 1999; Delgado, Porter and Stern, 2007; Bresnahan and Gambardella, 2004; Delgado, 2005; Cortright, 2006).

More specifically, clusters facilitate new business formation and the growth of successful start-ups by lowering the costs of entry (e.g. by providing ready access to suppliers or low-cost access to specialized inputs, offering an environment in which the costs of failure may be lower), enhancing opportunities for innovation-based entry (as a stronger cluster environment will allow local entrepreneurs to develop and commercialize new technologies more rapidly) and allowing start-up firms to leverage local resources to expand new businesses more rapidly. Finally, strong clusters are often associated with the presence of innovation-oriented local consumers, thus providing increased opportunities for entry into differentiated market segments. As a result, entrepreneurship is a particularly important channel for cluster-driven agglomeration, and may therefore be crucial for the role of clusters in enhancing regional performance (Porter, 1998a; Saxenian, 1994; Swann et al., 1998; Feldman, 2001; Feldman, Francis, and Bercovitz, 2005; Feser, Renski, and Goldstein, 2008; Wennberg and Lindqvist, 2008). *Thus, our third hypothesis is that, after controlling for the convergence effect, the growth rate of entrepreneurship will be increasing in the strength of the cluster environment in the region.*

It is useful to also consider the impact on entrepreneurship of similar clusters in neighboring regions. On the one hand, strong clusters in neighboring regions enhance the opportunities and lower the costs of entrepreneurship (e.g. by providing access to suppliers and customers, by allowing firms to leverage local technology and institutions, etc). Indeed, Delgado, Porter and Stern (2007) find that clusters that are co-located in nearby regions benefit from inter-regional spillovers. At the same time, the presence of a strong cluster in a neighboring region is a source of locational competition, particularly for capital investment and demand. *Therefore, our hypothesis is that the impact of the strength of the cluster environment in neighboring regions on the growth rate of entrepreneurship is ambiguous, and will depend on the relative salience of inter-regional spillovers versus locational competition.*

While we focus on the formation of start-up firms to measure pure entrepreneurship, for robustness we also examine all "new" establishments (including new establishments of existing firms). The opening of new establishments is an important channel by which entrepreneurial firms grow and contribute to regional performance. The location decisions of multi-establishment (and multinational) firms will depend on their portfolio of industries and locations (see e.g. Dunne et al., 2005; McCann and Mudambi, 2005), and so these firms location decisions may differ from those of start-ups. Overall, we expect a strong cluster environment will also attract new establishments of incumbent firms since such firms benefit from the specialized resources of particular locations (Bresnahan and Gambardella, 2004; Manning 2008).

3. Econometric model

To test our hypotheses, we need to evaluate the distinct impacts of agglomeration and convergence forces on start-up activity. We measure start-up activity in two related ways: the number of establishments by new firms in a region within a given traded industry (i.e. *start-up establishments)*, and the employment in these new firms (i.e., *start-up employment).* We are interested in separating out the role played by industrial clusters in start-up activity, while controlling for the economic activity within a region-industry, as well as broader factors such as the overall growth of a region or industry. To do so, we evaluate how the growth in start-up activity at the region-industry level is impacted by the initial level of start-up activity, industry specialization, the strength of the cluster environment surrounding the region-industry, and region and industry fixed effects. Our core econometric specification is therefore:

$$
\ln\left(\frac{\text{Start-up Activity}_{i,c,r,2002-2005}}{\text{Start-up Activity}_{i,c,r,1991-1994}}\right) = \alpha_0 + \delta \ln(\text{Start-up Activity}_{i,c,r,1991-1994})
$$

$$
+ \beta_1 \ln(\text{ Industry Spec}_{i,c,r,1990})
$$
$$
+ \beta_2 \ln(\text{Cluster Spec})_{i,c,r,1990}^{\text{outside i}}
$$
$$
+ \beta_3 \ln(\text{Linked Clusters Spec}_{c,r,1990}^{\text{outside c}})
$$
$$
+ \beta_4 \ln(\text{Cluster Spec in Neighbors}_{c,r,1990})
$$
$$
+ \alpha_i + \alpha_r + \varepsilon_{i,c,r,t}.
$$

(1)

The dependent variable is the growth rate in start-up activity of industry i in cluster c at region (EA) r, where the base period is the mean level of start-up activity during 1991-1994, and the end period is the mean level of start-up activity during 2002-2005 (the Data Section includes a detailed explanation of the variables). The explanatory variables include the level of start-up activity in the EA-industry, industry specialization and measures of the strength of related economic activity: cluster specialization, the strength of linked clusters, and the strength of similar clusters in neighboring regions. Our hypotheses are that the growth rate in start-up activity is subject to a convergence effect ($\delta < 0$), is increasing in the strength of clusters and linked clusters ($\beta_2 > 0$, $\beta_3 > 0$), and has an ambiguous relationship with industry specialization (β_1) and the strength of clusters in neighboring regions (β_4).

Our main econometric specification also accounts for other differences across regions and industries that affect start-up growth through the inclusion of industry (α_i) and region fixed effects (α_r). Our analysis thus controls for unobserved factors (such as idiosyncratic demand shocks, regional policies, etc) that might be correlated both with our explanatory variables and the start-up growth rate.

To account for correlation across industries within a regional cluster, the standard errors are clustered by region-cluster. Finally, since nearby regions tend to specialize in like clusters, there might be spatial autocorrelation of the performance and/or unobserved attributes of a region and its neighbors. We account for this by including in our main specifications the strength of similar clusters in neighboring regions (and region fixed effects). Additionally, we test for spatial correlation and find that this is not an issue.[5]

4.   **Data**

To estimate equation (1), we combine data from the Longitudinal Business Database (LBD) of the Census Bureau with cluster definitions drawn from the US Cluster Mapping Project (Porter, 2001, 2003). Before turning to the variable definitions, it is useful to provide an overview of these two data sources.

---

[5]   Based on LeSage (1999, p.171), we test for spatial autocorrelation by running a first-order spatial autoregressive (FAR) model on the residuals of our core OLS specifications. The estimated coefficients are highly insignificant (coefficients of 0.001 with p-values above 0.98), suggesting non relevant spatial correlation.

The LBD provides annual observations of the universe of US establishments with payroll from 1976 onward (see Jarmin and Miranda, 2002).[6] Importantly, the LBD distinguishes between entrepreneurship – the initial entry of a new firm in its first establishment – and business expansions by existing firms through the opening of new establishments. We aggregate this establishment-level data to the region-industry level and the region-cluster level, using four-digit SIC codes as the primary industry unit and economic areas (EAs) as the geographic unit.[7]

The cluster definitions are drawn from the US Cluster Mapping Project (USCMP). While the measurement of complementary economic activity in a consistent and unbiased manner is a considerable challenge,[8] the USCMP develops a methodology for grouping four-digit (and some three-digit) SIC codes into cluster and linked cluster groupings.[9] The methodology first distinguishes between three "types" of industries with different patterns of spatial competition and locational drivers: traded, local, and natural resource-dependent. To focus our analysis on those industries most closely linked to our hypotheses, we focus exclusively on the traded industries. These industries consist of 588 (mostly) four-digit SIC codes that are associated with service and manufacturing industries that sell products and services across regional and national boundaries.[10]

Porter (2003) assigns each traded industry into one of 41 mutually exclusive traded clusters (referred to as "narrow clusters"), where the set of industries within each cluster primarily reflects pairwise correlations of industry employment across locations. This approach to define cluster boundaries captures multiple types of agglomeration

---

[6] We drop establishments with very noisy average wages (below half of the minimum wage or above $2 million USD). Our findings are robust to including these observations.

[7] There are 179 EAs covering the United States, and they reflect meaningful economic regions (see Johnson and Kort, 2004). To minimize concerns about differences in the definition of neighbors, we exclude the Alaska and Hawaii EAs.

[8] A small literature considers alternative classification schemes. Ellison and Glaeser (1997) study the coagglomeration of industries, creating an index reflecting "excess" concentration. Feldman and Audretsch (1999) group those industries that have a common science and technological base, using the Yale Survey of R&D Managers. Other studies define linkages between industry activities in terms of their technological proximity (Jaffe et al., 1993). Finally, Ellison et al. (2007) find that input-output linkages and labor pooling are key drivers of the co-agglomeration of industries. This reasoning is consistent with the methodology developed in Porter (2001, 2003). *See also* Feser and Bergman (2000).

[9] To use industry data back to 1990, the analysis employs SIC system. By construction, recent NAICS-based data can be translated (with some noise) into the older SIC system.

[10] In contrast, local industries do not agglomerate and focus on local demand.

economies (technology, skills, shared infrastructure, demand and other linkages).[11] For example, in the information technology cluster, 9 individual industries are incorporated, including electronic computers (SIC 3571) and software (SIC 7372). These cluster definitions form our measures of complementary economic activity (see Porter (2003) for a list of the 41 clusters).

4.1. Variable definitions and sample description

4.1.1. Entrepreneurship.

Our two main measures of entrepreneurial activity are start-up employment and start-up establishments by new firms with payroll within a given EA-industry.[12] Specifically, *Start-up Employment* is defined as the level of employment in new firms during their first year of operation (with payroll); and *Start-up Establishments* is the count of these new firms. Consistent with prior work (e.g. Glaeser and Kerr 2009), we computer 4-year averages for these annual start-up activities. Using a multi-year span (and including a Census-year in the base and terminal periods) allows for a better signal of the true level of entrepreneurial activity and also reduces the number of EA-industries in which we observe zero entrants during a given period.[13]

One main goal of this paper is to evaluate how the cluster environment impacts the growth rate of entrepreneurship. While this focus allows us to evaluate the role of the cluster environment on regional dynamics, the most straightforward approach to evaluating growth -- taking ln($Start\text{-}up\ Activity_{i,r,2002\text{-}2005}$ / $Start\text{-}up\ Activity_{i,r,1991\text{-}1994}$) -- must account for the fact that there are many EA-industries in which there is a zero level of employment or, relatedly, a zero level of start-up activity in the study period. In either

---

[11] It is possible that industries with high co-location may have little economic relationship. Thus, in the USCMP two adjustments are made to the cluster definitions to eliminate spurious correlations. First, the SIC industry definitions and list of products and services are used to reveal logical links. Second, the National Input-Output accounts are used to look for meaningful cross-industry flows (see Porter, 2003). While industries that have meaningful interactions tend to co-agglomerate in space, there are *non-geographical* dimensions of proximity (e.g., institutional and organizational factors) that could also facilitate the interactions among industries and their firms (Boschma 2005; Torre, 2008). Our cluster definitions cannot directly capture these non-geographical dimensions.

[12] While we recognize that there are many new businesses without employees (Fairlie, 2009), new firms with employees have greater growth potential in terms of employment and innovation, capturing an important aspect of entrepreneurship.

[13] In the LBD data the inflow of new establishments may be recorded with some delay, with Census-years being most accurate in recording new establishments.

of these cases, we are required to either exclude those observations or impose a positive lower bound on the level of start-up activity.

We focus our analysis on a sample of EA-industries that have a non-zero level of employment during 1990 (53,213 observations). Then, to include in the analysis EA-industries where we observe zero start-up activity, we set a minimum level of start-up activity.[14] Alternatively, we show that our results are robust to the subsample of EA-industries with a positive level of start-up activity in both the baseline and terminal period (11,981 observations).[15]

Finally, we further account for the large number of zeros by examining the impact of the cluster environment on the *level* of start-up activity (versus growth), using all EA-industry pairs. These analyses include probit specifications that evaluate the probability of *any* start-up activity, and count models that account for the skewed distribution of start-up activity.

While we primarily focus on start-up establishments, we also examine total new establishments (including new establishments of existing firms). To do so, we compute two *entry* indicators: *Entry Employment* is defined as the level of employment in all new establishments within a given EA-industry; and *Entry Establishments* is the count of these new establishments. We then compute the growth rate in entry employment and establishments (see Table 1).

Finally, to test for the role of clusters in the medium-term performance of start-ups, we examine the level of employment in young (up to 5-year-old) start-ups in regional industries. The dependent variable is the (log of) average employment in 2004-2005 of start-up firms borne during 2001-2003 in the EA-industry (*ln Employment in Start-up Survivors$_{i,r}$*).

---

[14] We scale the start-up activity indicators by adding the minimum annual start-up employment (1) and start-up establishments (0.01) in the sample, which is a standard procedure to scale variables.

[15] This high skewness of start-up activity is documented in other studies (e.g. Glaeser and Kerr, 2009).

4.1.2. Industry and cluster specialization.

Our main empirical task is to examine the impact of the industry and different aspects of the cluster environment on the growth rate in start-up activity. As such, we require measures of industry and cluster specialization, as well as the strength of related and neighboring clusters. We draw on prior work which uses location quotients (LQ) as a measure of regional specialization (e.g. Glaeser et al., 1992; Porter, 2003). Specifically, the employment-based industry specialization in the base year (1990) is measured by the share of regional employment in the industry as compared to the share of US total employment in the national industry:

$$\text{INDUSTRY SPEC}_{\text{Employ},i,r,90} = \frac{\text{employ}_{i,r}/\text{employ}_r}{\text{employ}_{i,US}/\text{employ}_{US}},$$

where *r* and *i* indicate the region (EA) and the industry, respectively. This indicator captures to what extent the industry is "over-represented" in the EA. In the data, the employment-based industry specialization of EAs has a mean of 2.01 (Table 1).[16]

We utilize an analogous procedure to define cluster specialization. For a particular EA-industry, the specialization of the EA in cluster *c* is measured by the share of regional employment within the cluster (outside the industry) as compared to the share of US total employment in the national cluster (outside the industry):

$$\text{CLUSTER SPEC}_{\text{Employ},icr,90} = \frac{\text{employ}_{c,r}^{\text{outside } i}/\text{employ}_r}{\text{employ}_{c,US}^{\text{outside } i}/\text{employ}_{US}}.$$

It is useful to note that, with the inclusion of EA and industry fixed effects, the independent variation in industry specialization is driven by variation in employment in the EA-industry, and the variation in cluster specialization comes from the employment within a given EA-cluster (outside the industry).

Table 2 illustrates key attributes of the top regional clusters based on cluster specialization in 1990.[17] Top clusters tend to have a higher level of start-up activity, larger establishments and a higher presence of older firms than other clusters. Interestingly, on average over 30% of all the establishments in a top cluster belong to

---

[16] The specialization indicators are based on employment (count of establishments) in the start-up employment (establishment) models.

[17] The top EA-clusters are the top-10 EAs by cluster specialization for each cluster. This criterion is complemented with a minimum share of the US cluster employment (above the 20th percentile value).

firms that have establishments in more than one geographical market (EA). This suggests that a cluster may establish linkages with other locations through the presence of these regionally diversified firms.[18]

4.1.3. Strength of linked and neighboring clusters.

We additionally develop measures of the strength of "linked" clusters and the presence of clusters in neighboring regions. The measure of linked clusters is developed using the set of "broad" cluster definitions in Porter (2003). Specifically, while the narrow cluster definition (used for the earlier measures) classifies each industry into a unique cluster, in the *broad* cluster definition each industry may be associated with multiple clusters (based on locational correlation of employment). To develop a measure based on linkages to cluster c, we identify those broad clusters that have at least 1 of cluster c's narrow industries in common. For example, in the case of automotive, the linked clusters include production technology, metal manufacturing, and heavy machinery; among others. Having identified the set of clusters linked to a cluster (C^*), we then measure the degree of overlap between each pair of clusters (c, j) using the average proportion of (narrow) industries that are shared in each direction:

$\omega_{c,j} = \text{Avg}\left(\dfrac{\text{shared industries}_{c,j}}{\text{total industries}_c}, \dfrac{\text{shared industries}_{j,c}}{\text{total industries}_j}\right)$. The presence in a region of clusters

linked to cluster c is then defined by a weighted sum of the location quotients associated with each linked cluster:

$$\text{LINKED CLUSTERS SPEC}_{\text{Employ}_{c,r}} = \frac{\sum\limits_{j \in C_c^*}^{C^*} (\omega_{c,j} * \text{employ}_{j,r})}{\sum\limits_{j \in C_c^*}^{C^*} (\omega_{c,j} * \text{employ}_{j,US})} \Big/ \frac{\text{employ}_r}{\text{employ}_{US}}.$$

For instance, based on this weighting which emphasizes the degree of overlap between clusters, our measure of the strength of linked clusters for industries within the automotive cluster will weigh the presence of the metal manufacturing cluster more heavily than the presence of the furniture cluster (see Delgado, Porter and Stern, 2007).

---

[18] We cannot distinguish between employment in primary and headquarters activities and employment to serve local markets (e.g. sales offices). Hence we may overstate "true" cluster employment (Porter, 2003).

Finally, we develop a measure of the presence of like clusters in neighboring regions (cluster specialization in neighbors variable). In part, we include this measure based on the empirical observation that specialization in a particular cluster tends to be spatially correlated across neighboring regions. To explore the role of neighboring clusters in start-up growth in a region-industry, we compute the (average) specialization of adjacent EAs in the cluster (including the focal industry).

5. Results

Our analysis begins in Table 3 which shows the average start-up growth rates for EA-industries, based on their *initial* levels of start-up activity and cluster specialization. We divide all EA-industries into four categories based on whether they have low or high (below or above the median) level of start-up employment and cluster specialization (for their industry). We find that there is a significant decrease in the average start-up employment growth rate (between the baseline (1991-1994) and the terminal period (2002-2005)) when moving from *low* to *high* initial level of start-up employment, consistent with convergence in start-up activity (e.g., the growth rate decreases from 25 to -33% for the EA-industries with low cluster specialization). Importantly, regardless of the initial level of start-up employment, there is a significant increase in the growth of start-up activity when one moves from an EA-industry with a *low* level of cluster specialization to one with *high* level of cluster specialization (e.g., the growth rate increases from 25 to 36% or from -33 to -9%). In other words, those regional industries that are located in a relatively strong cluster experience much higher growth rates in entrepreneurship.

While the sharp contrasts in Table 3 are intriguing, it is possible that alternative factors are driving these results. We therefore turn in Table 4 to a more systematic regression analysis. The dependent variable is the start-up employment growth rate. In (4-1), we include only the initial levels of start-up employment, industry specialization and cluster specialization. The results provide evidence for the two main findings of this paper. First, there is a large convergence effect – the initial level of start-up employment is associated with a decline in the expected growth rate of start-up activity. At the same time, the presence of complementary economic activity in the form of clusters also has an

important positive influence on the growth rate of entrepreneurship. Both industry and cluster specialization (outside the industry) are associated with higher growth in start-up activity.

These results are reinforced in (4-2), where we incorporate the strength of linked clusters and the strength of the cluster in neighboring regions, and control for the total employment in the region. Both cluster specialization and the presence of linked clusters have a positive influence on the start-up growth rate, while strength of clusters in neighboring regions is actually associated with a lower growth of start-up activity. This latter finding is consistent with the hypothesis that neighboring clusters may also attract entrepreneurs and so provide a substitute for growth within a particular EA. Interestingly, the employment size of the region contributes to the growth of start-up activity of its regional industries, countervailing the convergence forces that take place at the region-industry level.

In (4-3) and (4-4), we implement the core specifications, in which we include region and industry fixed effects. The main results concerning convergence and the impact of cluster specialization are robust. The only meaningful change in the estimates concerns the impact of neighboring clusters. The estimated coefficient depends on whether we control for industry and region heterogeneity; not surprisingly, given that the expected effect is ambiguous.

Finally, in (4-5), we use the subsample of EA-industries with positive start-up activity during both the base and terminal periods (this allows us to avoid the scaling adjustments to the dependent variable and the convergence effect measure that we discussed in the Data Section). The main results are robust. To illustrate the size of the effects, an increase in the industry specialization variable (cluster and linked clusters specialization variables) from the first to the third quartile value is associated with a 3.4% (1.5%) increase in the *annual* start-up employment growth rate.

The core findings concerning the convergence effect and the impact of clusters persist when we examine growth in *start-up establishments*. In Table 5, the dependent and independent variables are now based on counts of establishments, and the structure of the specifications mirrors the logic of Table 4. Interestingly, the only significant difference in the results concerns the effect of industry specialization on entrepreneurial

growth. Whereas the coefficient on industry specialization was positive in the start-up employment models, the coefficient is negative in models (5-1) and (5-2). The coefficient becomes positive when we include region and/or industry fixed effects. The heterogeneity of this parameter across specifications is consistent with the fact that the expected effect is ambiguous. In contrast, the coefficient on the impact of clusters is positive across all specifications.[19]

In Table 6, we consider an alternative measure of "new" economic activity by examining the growth in employment in all "new" establishments (including new establishments by existing firms) and counts of all new establishments. Each of the specifications in Table 6 includes EA and industry fixed effects, and only vary in the measures of related economic activity that are included and whether the dependent and explanatory variables are measured based on employment (6-1 and 6-2) or based on counts of new establishments (6-3 and 6-4). The results are robust across all the specifications – the growth rate of entry declines with the initial level of entry activity (convergence), and increases with the presence of related industries within the cluster and linked clusters. Importantly, these findings hold when we focus only on new establishments of existing firms. In the data, these new establishments often belong to firms that operate in like clusters in other locations (EAs). While more research is needed, the findings suggest that multi-location firms seem to be opening establishments in strong clusters.

Moving beyond our growth framework, in Table 7 we examine the probability of *any* start-up activity and the *level* of start-up activity using all EA-industry pairs. In (7-1) we study the probability of any start-up activity during 2002-2005 using a probit model that includes a dummy equal to one if the EA-industry experienced *any* start-up activity in an earlier period (1990-1996), the measures of the strength of the industry and the

---

[19] It is useful to emphasize that our core findings on start-up activity growth (Tables 4-5) are robust to a variety of sensitivity checks. We have included in the model a dummy indicator equal to one for EA-industries with *any* start-up activity in the base period to further capture unobserved factors that influence entrepreneurship. We have varied the length and precise dates of the base and terminal periods. Specifically, we consider start-up activity during the base period (1991-1995) and the terminal period (1997-2001). Finally, we use a larger (noisier) sample that includes establishments with wage outliers (average wage below half of the minimum wage or above $2 million USD), which are dropped from our initial sample. In each of these cases the main results are robust.

overall cluster environment, and EA and industry fixed effects. We find that the probability of any start-up activity is increasing in the strength of the cluster.[20]

In (7-2) we examine the count of start-up establishments during 2002-2005 using a fixed effects negative binomial model, and the same independent variables than in the probit model (Table 7, Model 1). The estimated incidence-rate ratios suggest that an increase in the initial strength of the cluster (industry) has more than a 33% (29%) boost on the subsequent count of start-up establishments. Finally, these findings are confirmed when we study the *annual* count of start-up establishments in 1997 and 2002 (Census years) using a negative binomial model with year, EA and industry fixed effects (Table 7, Model 3).[21] The findings in Table 7 reinforce the estimates from the growth models: the probability of *any* start-up activity and the *level* of start-up establishments are increasing in the strength of the cluster environment.

Finally, while the focus of this paper is the formation of new businesses, we also examine the role of clusters in the medium-term performance of start-ups. In particular, in Table 8 we study the level of employment in young (up to 5 year old) start-ups in an EA-industry using our core econometric specifications (and samples). We find that a strong cluster positively affects the level of employment in young start-ups, suggesting that clusters facilitate survival and growth. In a related paper we explore more carefully the role of clusters in new business growth and the attributes of successful start-ups.

6.   Conclusion and Extensions

This paper finds striking evidence for the simultaneous yet distinct influences of agglomeration and convergence on the growth in the number of new firms and employment by new firms in regional (traded) industries. The growth in start-up activities at the region-industry level is declining in the initial level of start-up activity at the region-industry level due to convergence forces. After controlling for convergence, however there is strong evidence that the presence of a strong cluster surrounding a region-industry accelerates the growth in start-up activities. We find that industries

---

[20] For EA-industries with zero employment, we replace (log) industry specialization with the minimum value of this variable.

[21] Drawing on Blundell et al. (1995), we include an indicator of pre-existing start-up activity to control for unobservable attributes of EA-industries that influence entrepreneurship.

located within a strong cluster or that can access strong related clusters are associated with higher growth rates in the formation of new firms and start-up employment.

These findings offer an important contribution to the ongoing debate about the impact of related economic activity on entrepreneurship and economic performance. Most notably, building on the cluster framework developed by Porter (1990, 1998a), this paper moves beyond the traditional debate in which the presence of related economic activity simultaneously gives rise to complementarities as well as competition for inputs and customers, clouding the interpretation of any particular empirical finding. By first accounting for convergence and the potential for competition within each industry in a region, we are able to isolate the important impact of cluster-related complementarities on entrepreneurship. In other words, while at a (narrow) industry level firms may compete for a given pool of resources, the cluster environment that surrounds an industry will increase the pool of competitive resources and reduce the barriers of entry for new firms. Strong regional clusters enhance the range and diversity of entrepreneurial start-up opportunities while also reducing the costs of starting a new business.

While our analysis has focused on new business formation, we also examine the medium-term performance of start-ups. We find that clusters contribute to the level of employment in young start-ups in regional industries, suggesting that a strong cluster environment in a region enhances the performance of start-ups. The role of clusters in the dynamics of new businesses is an open research question that we are further investigating.

Our findings support the idea that clusters of related and complementary industries facilitate the growth in the formation of new businesses and the medium-term performance of start-ups in regional industries, even after controlling for the strength of the region-industry as well as region and industry heterogeneity. There is large heterogeneity in the types of entrepreneurship that we could further explore. Start-ups will differ in size, innovative-orientation, and their growth potential. Similarly, we could also examine the role of the attributes of the firms that participate in clusters (e.g., size, age, and region and product diversification). Further understanding the role of clusters in the demography of entrepreneurship will help design more effective entrepreneurship policies.

Finally, we find that clusters also matter for the formation of not just new firms, but of new establishments of existing firms. These new establishments often belong to firms that participate in like clusters in other locations, suggesting that firms may seek complementary regional clusters, benefiting from the comparative advantages of each location. While more research is needed, regional clusters may integrate into national and global value chains often through complex networks of subsidiaries of multi-location firms (Dunning 1998; Porter 1998b; Enright, 2000; Alcacer, 2006). Drawing on international business studies, future work should further study the attributes of clusters that are more attractive for multi-location firms, and how their participation in multiple clusters affect the organization and performance of these firms, and how they contribute to entrepreneurship in a particular regional cluster.

7. References

Acs, Z.J., and C. Armington, 2006, *Entrepreneurship, Geography and American Economic Growth*, New York, NY: Cambridge University Press.

Acs, Z.J., D. Audretsch, P. Braunerhhjem, and B. Carlsson, 2009, "The Knowledge Spillover Theory of Entrepreneurship," *Small Business Economics* 32 (1), pp. 15-30.

Agrawal, A and I. Cockburn, 2003, "The Anchor Tenant Hypothesis: Exploring the Role of Large, Local, R&D-intensive Firms in Regional Innovation Systems," *International Journal of Industrial Organization* 21, pp. 1227-1253.

Alcacer, J., 2006, "Location Choices across the Value Chain: How Activity and Capability Influence Collocation," *Management Science* 52 (10), pp. 1457-1471.

Armington, C., and Z.J. Acs, 2002, "The Determinants of Regional Variation in New Firm Formation," *Regional Studies* 36(1), pp. 33-45.

Audretsch, D.B., 1995, *Innovation and Industry Evolution*, MIT Press, Cambridge, MA.

Barro, R.J. and X. Sala-i-Martin, 1991, "Convergence across States and Regions," *Brookings Papers on Economic Activity* 1, pp. 107-182.

Barro, R.J. and X. Sala-i-Martin, 1995, *Economic Growth,* Cambridge, MA: MIT Press.

Barteisman, E., S. Scarpetta, and F. Schivardi, 2005, "Comparative Analysis of Firm Demographics and Survival: Evidence from Micro-Level Sources in OECD Countries," *Industrial and Corporate Change* 14(3), pp. 365-391.

Blundell, R., R. Griffith, and J. Van Reenen, 1995, "Dynamic Count Data Models of Technological Innovation," *The Economic Journal, 105 (*No. 429), pp. 333-344.

Boschma, R., 2005, "Proximity and Innovation: A Critical Assessment," *Regional Studies,* 39 (1), pp. 61-74.

Bostic, R.W., J.S. Gans, and S. Stern, 1997, "Urban Productivity and Factor Growth in the Late Nineteenth Century," *Journal of Urban Economics* 41(1), pp 38-55.

Bresnahan, T.F., A. Gambardella (eds.), 2004, *Building High-Tech Clusters. Silicon Valley and Beyond*. Cambridge University Press, New York.

Bresnahan, T.F. and P.C. Reiss, 1991, "Entry and Competition in Concentrated Markets," *The Journal of Political Economy* 99 (5), pp. 977-1009.

Bulow, J. L., J.D. Geanakoplos, and P.D. Klemperer, 1985, "Multimarket Oligopoly: Strategic Substitutes and Complements," *Journal of Political Economy* 91(31), pp.488-511.

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

Cortright, J., 2006, "Making Sense of Clusters: Regional Competitiveness and Economic Development," The Brookings Institution, Washington, D.C.

Davis, S., J. Haltiwanger and S. Schuh, 1996, *Job Creation and Destruction*, MIT Press, Cambridge, MA.

Delgado M, M.E. Porter and S. Stern, 2007, "Clusters, Convergence and Economic Performance," manuscript.

Delgado, M., 2005, "Do Clusters Really Matter for Companies' Competitive Strategies? Evidence at the Country Level," manuscript.

Dunne, T, S.D. Klimek, M.J. Roberts, 2005, "Exit from regional manufacturing markets: The role of entrant experience," Int. J. Ind. Organ. 23, pp. 399-421.

Dunne, T., M. Roberts, and L. Samuelson, 1988, "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," RAND Journal of Economics 19, pp. 495-515.

Dunning, J. H., 1998, "Location and the Multinational Enterprise: A Neglected Factor?" *Journal of International Business Studies*, 29 (1), 45-66.

Ellison, G. and E. Glaeser, 1997, "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy* 105, pp. 889-927.

Ellison, G. and E. Glaeser, W. Kerr, 2007, "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," NBER Working Paper, No. 13068.

Enright, M., 2000, "Regional clusters and multinational enterprises: independence, dependence or interdependence?," *International Studies of Management and Organization*, 30 (2), pp. 114-138.

Fairlie, R.W., 2009, "Kauffman Index of Entrepreneurial Activity Report: 1996-2008," at http://www.kauffman.org/uploadedFiles/kiea_042709.pdf.

Feldman, M.P., 2001, "The Entrepreneurial Event Revisited: An Examination of New Firm Formation in the Regional Context," *Industrial and Corporate Change* 10, pp. 861-891.

Feldman, M.P. and D. Audretsch, 1999, "Innovation in Cities: Science-based Diversity, Specialization and Localized Competition," *European Economic Review* 43, pp. 409-429.

Feldman, M.P., J. Francis, and J. Bercovitz, 2005, "Creating a Cluster While Building a Firm: Entrepreneurs and the Formation of Industrial Clusters." *Regional Studies* 39 (1), pp. 129-141.

Feser, E.J. and E.M. Bergman, 2000, "National Industry Cluster Templates: A Framework for Applied Regional Cluster Analysis," *Regional Studies* 34(1), pp. 1-19.

Feser, E., H. Renski, H. Goldstein, 2008, "Clusters and Economic Development Outcomes," *Economic Development Quarterly* 22 (4), pp. 324-344.

Frenken, K., F.G. van Oort & T. Verburg, 2007, "Related Variety, Unrelated Variety and Regional Economic Growth," *Regional Studies* 41, pp. 685-697.

Fujita, M., P. Krugman, and A.J. Venables, 1999, *The Spatial Economy*, MIT Press, Cambridge, Massachusetts.

Geroski, P.A. and K.P. Gugler, 2004, "Corporate Growth Convergence in Europe," *Oxford Economic Papers*, 56(4), pp. 597-620.

Glaeser E.L., et al., 1992, "Growth in Cities," *Journal of Political Economy* 100(6), 1126-52.

Glaeser, E.L., and W.R. Kerr, 2009, "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?" *Journal of Economics and Management Strategy*, 18(3), pp. 623-663.

Glaeser, E.L, W.R. Kerr, and G.A. M. Ponzetto, 2009, "Clusters of Entrepreneurship," *Journal of Urban Economics*, 67(1), pp. 150-168.

Gompers, P., J. Lerner, and D. Scharfstein, 2005, "Entrepreneurial Spawning: Public Corporations and the Formation of New Ventures, 1986-1999," *Journal of Finance* 60 (2), pp. 577-614.

Greenstone M., R. Hornbeck, and E. Moretti, 2008, "Identifying Agglomeration Spillovers: Evidence from Million Dollar Plants," NBER Working Paper No. 13833.

Haltiwanger, J., R.S. Jarmin, and J. Miranda, 2009, "Who Creates Jobs? Small vs. Large vs. Young," mimeo, NBER.

Henderson, J.V., 2003, "Marshall's Scale Economies," *Journal of Urban Economics* 53, pp. 1-28.

Henderson, J.V., A. Kuncoro and M. Turner, 1995, "Industrial Development in Cities," *Journal of Political Economy* 103, pp. 1067-1090.

Jaffe, A., M. Trajtenberg, R. Henderson. 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly J. Econ.* 108 577-598.

Jarmin, R, and J Miranda, 2002, "The Longitudinal Business Database," CES Working Paper.

Johnson K.P. and J.R. Kort, 2004, "2004 Redefinition of the BEA Economic Areas," at http://www.bea.gov/scb/pdf/2004/11November/1104Econ-Areas.pdf.

Kerr, W.R. and R. Nanda, 2009, "Democratizing Entry: Banking Deregulations, Financing Constraints, and Entrepreneurship, *Journal of Financial Economics* 94, pp. 124–149.

Klepper, S., 2007, "Disagreements, spinoffs, and the evolution of Detroit as the capital of the U.S. automobile industry," *Management Science,* 53(4), pp. 616-631.

LeSage, J.P. (1999), "The Theory and Practice of Spatial Econometrics", Econometrics Toolbox Functions for Matlab**,** at http://www.spatial-econometrics.com.

Manning, S., 2008, "Customizing Clusters: On the Role of Western Multinational Corporations in the Formation of Science and Engineering Clusters in Emerging Economies," *Economic Development Quarterly*, 22(4), pp. 316-323.

Markusen, A., 1996, "Sticky Places in Slippery Space: A Typology of Industrial Districts," *Economic Geography* 72(3), pp. 293-313.

Marshall, A., 1920, *Principles of Economics.* London: MacMillan.

McCann, P. and Mudambi, R. 2005, "Analytical Differences in the Economics of Geography: The Case of the Multinational Firm," *Environment and Planning A*, 37(10), pp. 1857-1876.

Porter, M. E., 1980, *Competitive Strategy: Techniques for Analyzing Industries and Competitors*, New York: Free Press.

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

Porter, M.E., 1998a, Clusters and Competition: New Agendas for Companies, Governments, and Institutions. In M.E. Porter (ed.). *On Competition*. Harvard Business School Press, Boston, pp. 197-299.

Porter, M.E., 1998b, 'Competing Across Locations: Enhancing Competitive Advantage Through a Global Strategy, in Porter, M. (Ed.): *On Competition*, Harvard Business School, Boston, pp. 305-344.

Porter, M.E, 2000, "Location, Competition, and Economic Development: Local Clusters in a Global Economy." *Economic Development Quarterly* 14 (1), pp. 15-34.

Porter, M.E., 2001, "Regions and the New Economics of Competition," in A.J. Scott (ed.), *Global City-Regions. Trends, Theory, and Policy*, New York: Oxford University Press, pp. 145-151.

Porter, M.E., 2003, "The Economic Performance of Regions," *Regional Studies* 37, pp. 549-578.

Quah, D.T., 1996, "Empirics for Economic Growth and Convergence," *European Economic Review* 40 (6), pp. 1353–1375.

Rosenthal, S.S., and W.C. Strange, 2009, "Small Establishments/Big Effects: Agglomeration, Industrial Organization and Entrepreneurship, forthcoming in E. Glaeser (ed.), 2009, *Economics of Agglomeration*, NBER.

Rosenthal, S.S. and W.C. Strange, 2004, "Evidence on the Nature and Sources of Agglomeration Economies," in J.V. Henderson and J.F. Thisse (eds.), *Handbook of Regional and Urban Economics*, vol. 4, Amsterdam: Elsevier North-Holland, pp. 2119-2172.

Rosenthal, S.S. and W.C. Strange, 2003, "Geography, Industrial Organization, and Agglomeration," *Review of Economics and Statistics* 85 (2), pp. 377-393.

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

Sorenson, O., P.G. Audia, 2000, The Social Structure of Entrepreneurial Activity: Geographic Concentration of Footwear Production in the United States, 1940-1989. *Amer. J. Sociology* 106 (2) 424-62.

Sutton, J., 1997, "Gibrat's Legacy" *Journal of Economic Literature*, 35, pp. 40-59.

Swann, G.M. P., M. Prevezer and D. Stout (eds.), 1998, *The Dynamics of Industrial Clustering: International Comparisons in Computing and Biotechnology*, Oxford: Oxford University Press.

Torre, A., 2008, "On the Role Played by Temporary Geographical Proximity in Knowledge Transmission," *Regional Studies,* 42 (6), pp. 869-889.

Wennberg, K. and G. Lindqvist, 2008, "The Effect of Clusters on the Survival and Performance of New Firms," *Small Business Economics*, 34(3), pp. 221-241.

Table 1: Region-industry descriptive statistics (mean and standard deviations)

EA-Industries with positive employment in 1990

Variables	Definition	N=53213	"No-zeros" Positive start-ups in 1991-94 & 2002-05 N=11981
EMPLOYMENT$_{90}$	EA-industry employment in 1990	554.67 (2464.7)	1544.09 (4778.32)
START-UP EMPLOYMENT$_{91-94}$	EA-industry (1991-1994 average) annual employment in start-up establishments	6.37 (86.42)	20.99 (179.05)
ENTRY EMPLOYMENT$_{91-94}$	EA-industry (1991-1994 average) annual employment in *all* new establishments	15.39 (146.70)	43.75 (262.82)
EMPLOYMENT IN START-UP SURVIVORS$_{04-05}$	EA-industry (2004-2005 average) annual employment of start-ups borne in 2001-03	18.28 (131.52)	66.41 (263.86)
ESTABLISHMENTS$_{90}$	EA-industry establishments in 1990	16.16 (71.23)	56.00 (142.01)
START-UP ESTABLISHMENTS$_{91-94}$	EA-industry (1991-1994 average) annual start-up establishments	.47 (2.59)	1.87 (5.21)
ENTRY ESTABLISHMENTS$_{91-94}$	EA-industry (1991-1994 average) annual new establishments	0.82 (5.28)	2.54 (9.39)
Δ START-UP EMPLOYMENT	Growth rate in start-up employment $\ln\left(\dfrac{\text{start-up employ}_{i,r,02-05}}{\text{start-up employ}_{i,r,91-94}}\right)$	.16 (1.02)	.47 (1.73)
Δ START-UP ESTABLISHMENT	Growth rate in start-up establishments	.10 (1.78)	.04 (.91)
Δ ENTRY EMPLOYMENT	Growth rate in entry employment	.14 (1.25)	**.**36 (1.80)
Δ ENTRY ESTABLISHMENT	Growth rate in entry establishments	.07 (1.86)	.02 (.94)
INDUSTRY SPEC$_{Employ, 90}$	Industry employment-based Location Quotient $LQ_{i,r} = \dfrac{\text{employ}_{i,r}/\text{employ}_r}{\text{employ}_{i,US}/\text{employ}_{US}}$	2.01 (6.42)	1.47 (4.26)
CLUSTER SPEC$_{Employ, 90}$	Cluster employment-based LQ (outside the industry)	1.17 (1.95)	1.18 (1.91)
LINKED CLUSTERS SPEC$_{Employ, 90}$	Linked clusters' employment-based LQ (weighted by cluster overlap)	1.04 (.92)	1.01 (.70)
CLUSTER SPEC in NEIGHBORS$_{Employ, 90}$	Neighboring clusters' average employment-based LQ	1.13 (1.19)	1.06 (1.06)
INDUSTRY SPEC$_{Estab, 90}$	Industry establishment-based LQ.	1.78 (3.20)	1.40 (2.57)
CLUSTER SPEC$_{Estab, 90}$	Cluster establishment-based LQ (outside the industry)	1.06 (1.14)	1.13 (1.25)
LINKED CLUSTERS SPEC$_{Estab, 90}$	Linked clusters' establishment-based LQ (weighted by cluster overlap)	.97 (.46)	.98 (.35)
CLUSTER SPEC in NEIGHBORS$_{Estab, 90}$	Neighboring clusters' average establishment-based LQ	1.03 (.80)	1.02 (.81)

Note: The core sample includes EA-industries with positive employment in 1990, resulting in 53213 observations. The "No-zeros" subsample includes EA-industries with positive start-up activity in both the base and terminal periods (1991-1994 and 2002-2005).

Table 2: Demographics of EA-clusters in 1990 (mean and standard deviation)

	Start-up Estab$_{91-94}$ (Annual, level)	Multi-EA Estab (rate)	Avg. Size Estab	Young Estab Age<5 (rate)	Old Estab Age>10 (rate)
Top EA-clusters$_{1990}$	**8.80**	**.313**	**247.761**	**.249**	**.520**
N=410	**(29.16)**	**(.248)**	**(394.20)**	**(.162)**	**(.210)**
Other EA-clusters$_{1990}$	**3.11**	**.193**	**55.661**	**.304**	**.426**
N=6220	**(12.97)**	**(.217)**	**(93.301)**	**(.214)**	**(.240)**
t-test of diff. of means	**5.950**	**10.684**	**28.273**	**-5.095**	**7.717**

Note: The top regional clusters are the top-10 EAs by cluster specialization for every cluster (e.g., for automotive the top-one regional cluster would be in Detroit-Warren-Flint, MI). *Multi-EA Establishments* measures the presence in a local cluster of firms that are active in more than one geographical market (EA) for a given cluster (e.g., Microsoft has establishments in the IT cluster in many EAs). Specifically, this variable is the rate of establishments of multi-EA firms in a given EA-cluster. *Avg. Size Establishment* is the average employment size of the establishments in a given EA-cluster. *Young* (Old) *Establishments* is the rate of establishments with Age<5 (Age>10) in a given EA-cluster.

Table 3: EA-industry average growth rate in start-up employment
(by level of start-up employment and cluster specialization, N=53213)

		START-UP EMPLOYMENT$_{ir, 1991-94}$	
		Low	**High**
CLUSTER SPEC$_{Employ, cr, 1990}$ (Outside the industry)	**Low**	ΔSTART-UP EMPLOYMENT= .25 N= 20507	ΔSTART-UP EMPLOYMENT= -.33 N=6265
	High	ΔSTART-UP EMPLOYMENT= .36 N= 17474	ΔSTART-UP EMPLOYMENT= -.09 N=8967

Notes: Low versus High is based on the median of the variable (for each industry).
The start-up employment growth rate is between the baseline period (1991-1994) and the terminal period (2002-2005). All the averages are significantly different from each other at 1% level.

Table 4: EA-industry growth in start-up employment (N=53213)

	START-UP EMPLOYMENT GROWTH				
					No zeros **N=11981**
	1	2	3	4	5
Ln START-UP EMPLOYMENT$_{91-94}$	**-.299**	**-.358**	**-.680**	**-.684**	**-.827**
	(.012)	**(.011)**	**(.009)**	**(.009)**	**(.011)**
Ln INDUSTRY SPEC$_{Employ, 90}$	**.030**	**.045**	**.112**	**.107**	**.283**
	(.003)	**(.003)**	**(.003)**	**(.003)**	**(.014)**
Ln CLUSTER SPEC$_{Employ, 90}$	**.031**	**.017**	**.025**	**.013**	**.093**
(Outside the industry)	**(.002)**	**(.003)**	**(.003)**	**(.003)**	**(.019)**
Ln LINKED CLUSTERS SPEC$_{Employ, 90}$		*.013*		**.061**	**.110**
		(.007)		**(.007)**	**(.030)**
Ln CLUSTER SPEC in		**-.018**		**.031**	**.076**
NEIGHBORS$_{Employ, 90}$		**(.007)**		**(.007)**	**(.027)**
Ln REGIONAL EMPLOYMENT		**.157**			
		(.005)			
EA FEs	**No**	**No**	**Yes**	**Yes**	**Yes**
INDUSTRY FEs	**No**	**No**	**Yes**	**Yes**	**Yes**
R-Squared	**.084**	**.115**	**.267**	**.269**	**.400**

Notes: Bold and italic numbers refer to coefficients significant at 1% and 10% levels. Robust standard errors clustered by EA-Cluster. The explanatory variables are in logs.

Table 5: EA-industry growth in start-up establishments (N=53213)

	START-UP ESTABLISHMENT GROWTH				
					No zeros **N=11981**
	1	2	3	4	5
Ln START-UP ESTABLISHMENTS$_{91-94}$	**-.375**	**-.406**	**-.863**	**-.865**	**-.654**
	(.007)	**(.006)**	**(.005)**	**(.005)**	**(.012)**
Ln INDUSTRY SPEC$_{Estab, 90}$	**-.107**	**-.066**	**.570**	**.557**	**.295**
	(.008)	**(.009)**	**(.009)**	**(.010)**	**(.012)**
Ln CLUSTER SPEC$_{Estab, 90}$	**.086**	**.047**	**.026**	.007	**.067**
(Outside the industry)	**(.006)**	**(.007)**	**(.006)**	(.006)	**(.017)**
Ln LINKED CLUSTERS SPEC$_{Estab, 90}$		**.092**		**.152**	**.100**
		(.020)		**(.018)**	**(.032)**
Ln CLUSTER SPEC in NEIGHBORS$_{Estab, 90}$		.018		**.050**	.010
		(.018)		**(.014)**	(.025)
Ln REGIONAL ESTABLISHMENTS		**.213**			
		(.010)			
EA FEs	**No**	**No**	**Yes**	**Yes**	**Yes**
INDUSTRY FEs	**No**	**No**	**Yes**	**Yes**	**Yes**
R-Squared	**.176**	**.193**	**.440**	**.442**	**.311**

Notes: Bold numbers refer to coefficients significant at 1% level. Robust standard errors clustered by EA-Cluster. The explanatory variables are in logs.

Table 6: EA-industry growth in entry (all new establishments, N=53213)

	ENTRY EMPLOYMENT GROWTH		ENTRY ESTABLISHMENT GROWTH	
	1	2	3	4
Ln ENTRY $_{91-94}$	**-.708**	**-.712**	**-.872**	**-.875**
	(.007)	**(.007)**	**(.005)**	**(.005)**
Ln INDUSTRY SPEC $_{90}$	**.145**	**.139**	**.621**	**.604**
	(.003)	**(.003)**	**(.010)**	**(.010)**
Ln CLUSTER SPEC $_{90}$	**.032**	**.019**	**.035**	*.011*
(Outside the industry)	**(.003)**	**(.004)**	**(.006)**	*(.007)*
Ln LINKED CLUSTERS SPEC $_{90}$		**.083**		**.185**
		(.009)		**(.019)**
Ln CLUSTER SPEC in NEIGHBORS $_{90}$		**.031**		**.068**
		(.007)		**(.015)**
EA FEs	**Yes**	**Yes**	**Yes**	**Yes**
INDUSTRY FEs	**Yes**	**Yes**	**Yes**	**Yes**
R-Squared	**.325**	**.328**	**.454**	**.455**

Notes: Bold and italic numbers refer to coefficients significant at 1% and 10% levels. Standard errors are clustered by EA-Cluster. The variables are based on employment in 6-1-6-2 and count of establishments in 6-3-6-4.

Table 7: EA-industry level of start-up activity (using the full sample)

	Probit (Marginal Effects)	Negative Binomial (Incidence-Rate Ratios)	
	ANY START-UP ACTIVITY (during 2002-05)	START-UP ESTABLISHMENTS (during 2002-05)	ANNUAL START-UP ESTABLISHMENTS (1997, 2002)
	1	2	3
ANY START-UP ACTIVITY $_{90-96}$	**.054**	**1.349**	**1.289**
	(.003)	**(.016)**	**(.019)**
Ln INDUSTRY SPEC $_{Estab, 90}$	**.011**	**1.286**	**1.360**
	(.000)	**(.005)**	**(.006)**
Ln CLUSTER SPEC $_{Estab, 90}$	**.008**	**1.338**	**1.385**
(Outside the industry)	**(.002)**	**(.011)**	**(.012)**
Ln LINKED CLUSTERS SPEC $_{Estab, 90}$	.004	**1.117**	**1.097**
	(.003)	**(.017)**	**(.019)**
Ln CLUSTER SPEC in NEIGHBORS $_{Estab, 90}$	**.025**	**1.107**	**1.044**
	(.003)	**(.014)**	**(.015)**
EA FEs	**Yes**	**Yes**	**Yes**
INDUSTRY FEs	**Yes**	**Yes**	**Yes**
YEAR FEs			**Yes**
R-Squared	.450		
Log-likelihood	-29471	-58404	-63412
Obs.	103368	103014	206028

Note: Bold numbers refer to coefficients significant at 1% level. In (7-1) the dependent variable is a dummy equal to one for EA-industries with *any* start-up activity during 2002-05; the coefficients are marginal effects from the probit model. In (7-2) we examine the count of start-up establishments during 2002-05 using a Negative Binomial (NB) model, and the coefficients are the incidence-rate ratios. In (7-3) we examine the *annual* count of start-up establishments in 1997and 2002 using a NB model. To avoid convergence problems, in the NB models we combine the 10% smallest EAs and national industries by employment size in a single EA and a single industry dummy, including in the models a total of 160 EA dummies and 529 industry dummies. Alternatively, we drop the 10% smallest EAs and industries and the same findings hold.

Table 8: EA-industry level of employment in start-up survivors (N=53213)

| | Ln EMPLOYMENT in START-UP SURVIVORS$_{2004\text{-}2005}$ | | |
| | | | No-zeros N=11981 |
	1	2	3
Ln START-UP EMPLOYMENT$_{91\text{-}94}$	**.357**	**.352**	**.158**
	(.010)	**(.010)**	**(.011)**
Ln INDUSTRY SPEC$_{Employ, 90}$	**.116**	**.112**	**.238**
	(.003)	**(.003)**	**(.014)**
Ln CLUSTER SPEC$_{Employ, 90}$	**.024**	**.014**	**.096**
(Outside the industry)	**(.003)**	**(.004)**	**(.019)**
Ln LINKED CLUSTERS SPEC$_{Employ, 90}$		**.078**	**.123**
		(.008)	**(.027)**
Ln CLUSTER SPEC in NEIGHBORS$_{Employ, 90}$		*.016*	.036
		(.008)	(.026)
EA FEs	**Yes**	**Yes**	**Yes**
INDUSTRY FEs	**Yes**	**Yes**	**Yes**
R-Squared	**.338**	**.339**	**.431**

Notes: Bold and Bold-Italic numbers refer to coefficients significant at 1% and 5%. Robust standard errors clustered by EA-Cluster. The dependent variable is the (log of) average employment in 2004-05 of start-up establishments borne during 2001-2003 in a region-industry. We add 1 to the dependent variable before taking logs. The explanatory variables are in logs