Clusters and Entrepreneurship[*]

Mercedes Delgado
Michael E. Porter
Scott Stern

December, 2009

Abstract

This paper evaluates the role of regional clusters on entrepreneurship in regional industries. We focus on the distinct influences of convergence and agglomeration on the rate of growth in the number of start-up firms and in employment by start-up firms. While reversion to the mean and diminishing returns to specialization within a location 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 channel by which location-based complementarities are realized. Using a novel panel dataset, there is significant evidence for the impact of clusters on entrepreneurship, after controlling for the impact of convergence at the region-industry level: industries located in regions with a large presence of related industries (i.e., strong clusters) experience higher growth in new business formation and start-up employment. Furthermore, strong clusters contribute to the level of employment in young start-ups in regional industries.

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

---

[*] Acknowledgements: William Kerr, Ben Jones, Christian Ketels, Jim Davis, Christian Asmussen, Roc Armenter and two referees contributed helpful suggestions.. The research in this paper was conducted while the authors were Special Sworn Status researchers of the US Census Bureau at the Boston Census Research Data Center (BRDC). Research results and conclusions expressed are our own and do not necessarily reflect the views of the Census Bureau. This paper has been screened to insure that no confidential data are revealed. Author contact information: Mercedes Delgado (Temple University, Fox School of Business; mdelgado@temple.edu), Michael E. Porter (HBS; mporter@hbs.edu); and Scott Stern (Northwestern University, MIT, NBER; s-stern2@northwestern.edu).

1. Introduction

In this paper, we examine the impact of agglomeration on the growth of entrepreneurship at the regional 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 firm creation are a striking feature of the US economy (Porter, 2003).[1] While a significant body of work explores why some regions experience more rapid growth than others (Porter, 1990, 1998; Saxenian, 1994; Glaeser, et al, 1992; Barro and Sala-i-Martin, 1995; Venables, 1996; Henderson, 1997; Fujita, Venables, and Krugman, 1999), there is increasing academic and policy interest in the particular role played by entrepreneurship. Startups seem to be an important driver of net regional employment growth (Davis et al, 1996; Haltiwanger, et al, 2009), and there is large regional variance in startup formation across regions (Armington and Acs, 2002). 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 1998; Saxenian, 1994; Feldman, 2001; Armington and Acs, 2002; Acs, et al, 2009; Glaeser and Kerr, 2009).

In an effort to explain region-industry growth two countervailing economic forces must be accounted for: 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, 1992).[2] Agglomeration exerts an opposite force on regional evolution. In the presence of agglomeration economies, the potential for growth is *increasing* in the level of economic activity (Glaeser, et al, 1992; Henderson, et al, 1995). From an empirical perspective, distinguishing the relative importance and 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

[1] For example, using Bureau of Economic Analysis Economic Areas (EAs) as the unit of analysis, Porter (2003) documents large cross-EA differences in employment growth during the 1990s, even when conditioning on the initial level of EA employment, and there are even larger cross-EA differences in the creation of new firms.

[2] While many studies of convergence focus on diminishing returns at the regional level, (Barro and Sala-i-Martin, 1995), convergence may also arise at more micro levels of analysis, such as the region-industry level (Henderson, et al, 1995; Dumais, et al, 2002, Bostic, et al, 1997; Cingano and Schivardi, 2004; Higgins et al, 2006). See Magrini (2004) for a recent review of the convergence literature.

growth will reflect a balancing of the two effects, making it infeasible to identify either effect in isolation (Henderson, et al, 1995).

We move beyond this traditional impasse by identifying the impact of industrial 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 (Delgado, Porter, and Stern, 2007). Building on Porter (1990, 1998, 2001), we focus specifically on the role of clusters –complementary industries related by technology, skills, demand and other linkages. Agglomeration is a consequence of the presence of complementary economic activity and clusters are a particularly important mechanism by which location-based complementarities are realized.

The main contribution of this paper is to examine 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 and risk of starting a business, enhancing opportunities for innovation, and enabling better access to a more diverse range of inputs and complementary products. (Saxenian 1994, Porter 1998, 2000, Feldman and Francis 2004, Glaeser and Kerr 2009). The co-location of companies, customers, suppliers, and other institutions increases the perception of innovation opportunities while amplifying the pressure to innovate (Porter, 2000). Since entrepreneurs are essential agents of change and innovation, a strong cluster environment should foster entrepreneurial activity.

The empirical analysis exploits the establishment-level Longitudinal Business Database (LBD) of the Census Bureau and the cluster definitions from the US Cluster Mapping Project (Porter 2001, 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. Using both databases we measure entrepreneurship and industrial composition at the region-cluster-industry level. We focus on a dataset that spans the years 1990-2005, includes 177 mutually exclusive Economic Areas (EAs) in the contiguous United States, and incorporates (up to) 588 "traded" industries spanning 41 clusters for each EA.[3]

Our empirical work focuses on early stage entrepreneurship, which we measure using two related indicators of start-up activity. We measure the number of new establishments by new firms in a region within a given traded industry (which we refer to as the level of *start-up establishments*), and by the employment in these new firms (which we refer to as the level of *start-up employment*). We then compute the growth rate in start-up establishments and start-up employment in each regional industry, and estimate the impact of cluster-driven agglomeration while accounting for the impact of convergence. Our core specifications incorporate detailed controls, including region and industry fixed effects. In other words, we estimate the impact of cluster composition on entrepreneurship, relying exclusively on variation arising from the relative size or strength of each cluster within a given region, and accounting for the overall growth in start-up activity of a given region and industry.

We find striking evidence for the simultaneous yet distinct influences of agglomeration and convergence on the growth rate of start-up establishments and start-up 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. At the same time, the start-up employment growth rate is increasing in measures of the strength of the cluster environment. We find similar findings for the growth rate 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 positive impact of clusters is both quantitatively and statistically significant, and robust to a variety of checks. The results provide support for the hypothesis that strong clusters facilitate growth in entrepreneurship over time.

While we primarily focus on the formation of start-up firms, for robustness we also look at total new establishments (including new establishments of existing firms).

[3] Traded industries are those that sell products and services across regional and national boundaries. See Porter (2003) and Section 4 for detailed explanation on traded industries and clusters.

We find that stronger clusters also facilitate the entry of new establishments of existing firms. While more research is needed on the locational and organizational decisions of multi-establishment (and multinational) firms (Holm, Malmberg, and Solvell, 2003; Mudambi and McCann, 2005; Alcacer, 2006), these firms may be opening establishments in new locations to seek complementary clusters and benefit from the competitive resources of each location (Enright, 2000; Bresnahan and Gambardella, 2004; Manning 2008). The contribution of these firms to generate entrepreneurial opportunities in a particular regional cluster is an open research question that we examine in related work.

Finally, we also examine the role of clusters in the performance of start-up firms. A strong cluster may raise the productivity of the participating firms, raising as well the bar for survival of new businesses (Sorenson and Audia, 2000). At the same time, a strong cluster in a region could facilitate the growth of start-up firms by providing better access to the necessary specialized inputs to commercialize their products and services. Thus, we expect clusters will enhance the performance for the most productive start-ups. To test this, we study the level of employment in young start-up survivors in a region-industry. After controlling for the level of start-up activity in the base period, we find that the cluster environment contributes to improve the level of employment in young start-up survivors, suggesting that clusters facilitate the survival and potentially the growth of successful 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 types of mechanisms are associated with entry of new businesses in agglomerated areas. On the one hand, 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, specialized institutions and the structure of regional business and social networks (Porter 1990, 1998, 2000; Malecki, 1990; Saxenian 1994; Markusen 1996). While most empirical studies of agglomeration focus on variables such as the aggregate rate of employment growth, an emerging theoretical and empirical literature emphasizes the role of the formation and the dynamics of new businesses in regional economic growth (Davis et al, 1996; Acs and Armington, 2006; Haltiwanger, et al, 2009). Relative to business expansions by incumbent firms, entrepreneurs may be more likely to be able 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 large churning of firms in most countries and sectors, especially amongst young and small businesses (Davis et at, 1996; Dunne et al, 1988, 2005; Barteisman, et al, 2005). Start-up firms have greater exit rates than new establishments of existing firms since they lack experience in the industry and in the location (Dunne et al, 2005; Kerr and Nanda, 2009). Importantly, those start-up firms that survive tend to have greater growth potential than incumbent firms (Barteisman, et al, 2005).[4] As such, the presence of a strong cluster environment that reduces barriers to entry and to firm growth and enables regional comparative advantage will be a central driver of entrepreneurial growth.

The precise mechanisms by which the structure of 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 these 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 (relative to Route 128) and a more decentralized organization of production. An extensive literature also highlights the broader relationship between entrepreneurship and the regional innovation system (e.g., Audretsch, 1995; Feldman 2001; Shane 2001; Armington and Acs, 2002; Acs, et al, 2009). Recently, Glaeser and Kerr (2009) carefully test for the impact of specific Marshallian economies of

[4] Barteisman, et al (2005) find that start-ups are of significantly lower size in the US than in Europe, but they seem to enjoy lower barriers to firm growth reaching rapidly a higher average size.

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 (and mostly independent) literature within regional and international business studies examine how the location decision and the benefits of agglomeration depend on attributes of the firm as well as on attributes of the industry and the location (Saxenian 1994; Henderson 2003; Rosenthal and Strange 2003; Alcacer 2006; McCann and Mudambi 2005; among others). There are key interactions between the internal organization of the firm (start-ups, multi-location, multinational, small or large, young or old, corporate organization, etc.) and the agglomeration benefits from a geographical location. One dimension that has received special attention is the role of small firms in extracting and generating economies of agglomeration. For example, Henderson (2003) finds that the extent of localization economies is larger for single-unit plants (vs. multi-unit plants), which tend to be more dependent on the external environment. Related work suggests that the presence of smaller (and younger) firms spurs additional new business creation and regional employment growth (Glaeser et al. 1992, Rosenthal and Strange, 2009; Glaeser, Kerr and Ponzetto 2009, and Feberman, 2007, among others).

In contrast, some regional and cluster studies highlight the particular importance of so-called "anchor" firms (including multi-establishment and multinationals) that induce spin-offs and attract firms from related industries (Agrawal and Cockburn 2002, Enright 2000, Feldman, Francis, and Bercovitz 2005, Klepper 2007, Greenstone, et al, 2008; Manning, 2008). Further research is needed to understand the location decision of multi-location firms and their contribution to the entrepreneurial activities 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. Drawing on the cluster theory we suggest that to capture important agglomeration forces the focus should be the presence of complementary and related industries instead of industry diversity per se (Porter 1990, 1998; Feldman and Audretsch, 1999; Delgado, Porter and Stern, 2007).[5] Thus, we use Porter's (1998, 2003) empirical cluster framework to explore agglomeration across set of industries related by technology, skills, demand, or other linkages, abstracting from the individual underlying mechanisms that induce agglomeration benefits.

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. This relationship will be subject to a convergence effect, which can be interpreted in terms of mean-reversion or diminishing marginal returns to entrepreneurial opportunities. Mean reversion simply 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; Henderson, et al, 1995). At the same time, it is possible that the returns to entrepreneurial activity are 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 particular region-industry at a point in time may result in diminished near-term opportunities for entrepreneurship in that same 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 level of region-industry start-up activity.*

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

---

[5] The literature on agglomeration economies highlights numerous channels that may facilitate positive externalities (see the review by Rosenthal and Strange, 2004), including increasing returns to industry specialization and to diversity at the regional level (Glaeser et al, 1992; Henderson, et al, 1995). The cluster theory challenges this conceptualization of industry specialization and regional diversity by focusing on the role of clusters of related and complementary industries.

the one hand, the returns to entrepreneurship are lower in the face of intensive competition, and so the incentives for start-up entry in a particular location will be lower in the presence of a higher level of local price-based competition (Porter, 1980; Bresnahan and Reiss, 1991). 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 and diversity 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 particular 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 ultimate empirical relationship between industry specialization and the growth rate of entrepreneurship 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.*[6]

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 (Porter 1990, 1998, 2003; Feldman and Audretsch, 1999; Delgado, Porter and Stern, 2007). A strong cluster environment enhances growth at the region-industry level by raising the returns to business expansion, capital investment, and

[6] In the empirical analysis we do not test how the nature of competition affects start-up activity. Instead, we test whether positive externalities (e.g., complementarities with the established firms) or congestion forces (e.g., price-based competition effects) prevail at the region-industry level.

innovation, thereby increasing job creation and productivity (see e.g., Porter 1990, 1998, 2003; Saxenian 1994; Swann 1998; Feldman and Audretsch, 1999; Delgado, Porter and Stern, 2007; Bresnahan and Gambardella, 2004; Bonte, 2004; Delgado, 2005; Cortright, 2006). More specifically, clusters of related and complementary industries facilitate new business formation and the growth of successful start-ups by lowering the costs and risks of entry (e.g., by providing low-cost access to specialized capital 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 demanding local consumers, thus providing increased opportunities for entrepreneurial entry into emerging and 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 economic performance (Porter 1998; Saxenian 1994; Swann 1998; Feldman 2001; Feldman and Francis 2004; Feser, Renski, and Goldstein, 2008; Glaeser and Kerr, 2009; 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 of 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 and industries that are co-located in nearby regions benefit from inter-regional spillovers in employment growth. 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 entrepreneurship. Hence entrepreneurs may move to a neighboring region to open a business when that neighboring region has a strong cluster environment, reducing the potential for entrepreneurship growth for locations with weak cluster environments adjacent to strong cluster environments. *Therefore, our hypothesis is that the impact of the strength of neighboring regions' cluster environment on the growth rate of entrepreneurship is ambiguous, and will depend on the relative salience of inter-regional spillovers versus locational competition.*

While this paper focuses on new business formation, we recognize that the link between entrepreneurship and regional growth depends also on the survival and growth of these new enterprises. Thus, we tentatively examine the role of clusters in the survival of start-ups in regional industries. As mentioned earlier, a strong cluster in a region could reduce the barriers to the growth of start-up firms by providing better access to the necessary inputs to develop and commercialize their products and services. At the same time, a strong cluster may raise the productivity of the participating firms, raising as well the productivity bar for survival of new businesses. *Thus, our hypothesis is that a strong cluster environment will enhance the potential for growth for the most productive start-ups, while also raising the quality and productivity bar for survival*.

3. Econometric Model

To test our hypotheses, we need an empirical framework that allows us 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 (with payroll) 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 particularly 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 adapt the conditional convergence framework (Barro and Sala-i-Martin, 1991; Henderson, et al, 1995) and evaluate how the *growth* in start-up activity at the region-industry level is impacted by the level of start-up activity, industry specialization, the strength of the cluster environment, and region and industry fixed effects. Our core econometric specification is therefore:

$$\ln\left(\frac{\text{Start-up Activity}_{\text{i,c,r,2002-05}}}{\text{Start-up Activity}_{\text{i,c,r,1991-94}}}\right) = \alpha_0 + \delta \ln(\text{Start-up Activity}_{\text{i,c,r,1991-94}}) + \beta_1 \ln(\text{ Industry Spec}_{\text{i,c,r,1990}}) +$$
$$\beta_2 \ln(\text{Cluster Spec})_{\text{i,c,r,1990}}^{\text{outside i}} + \beta_3 \ln(\text{Linked Clusters Spec}_{\text{c,r,1990}}^{\text{outside c}}) +$$
$$\beta_4 \ln(\text{Cluster Spec in Neighbors}_{\text{c,r,1990}}) + \alpha_{\text{i}} + \alpha_{\text{r}} + \varepsilon_{\text{i,c,r,t}}. \qquad (1)$$

The dependent variable is the growth in start-up activity of traded industry *i* in cluster *c* at region (EA) *r*, where the base period is the mean level of start-up activity during the years 1991-1994, and the end period is the mean level of start-up activity during the year 2002-2005 (Section 4 includes a detailed explanation of this measure and its construction). The explanatory variables include the level of start-up activity in the region-industry, industry specialization and measures of the strength of related economic activity: cluster specialization, the strength of linked (related) clusters, and the strength of similar clusters in neighboring regions. These measures capture the relative scale and strength of different types of economic activity potentially impinging on start-up activity at the region-industry level. Our main 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 ($\beta_2 > 0$ and $\beta_3 > 0$), 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 the start-up growth rate through the inclusion of industry (α_i) and EA 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 measures of cluster specialization and the start-up growth rate in a particular region-industry. Thus, our core identification structure estimates the impact of cluster composition on entrepreneurship, relying exclusively on variation arising from the relative size or strength of that cluster within a given region, accounting for the overall growth of a given region and industry. Finally, to account for correlation across industries within each regional cluster, the standard errors are clustered by region-cluster.[7]

4. Data

To estimate equation (1), our dataset includes measures of start-up activity at the region-industry level (at two points in time), as well measures of industry and cluster specialization during a baseline period. 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 precise variable definitions, it is useful to provide an overview of these two data sources.

The LBD provides annual observations of the universe of US establishments with payroll from 1976 onward. For each establishment, the LBD includes the date of entry, physical location, industry code, and number of employees of that establishment. Importantly, the LBD offers both an establishment-level identifier and a firm-level identifier, so it is possible to distinguish 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.[8] Our approach aggregates this 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.[9]

Our approach combines the LBD with a classification system for cluster definitions drawn from the US Cluster Mapping Project (Porter, 2001, 2003). While the

---

[7] Additionally, since nearby regions tend to specialize in the same type of clusters, there might be spatial dependence of the performance and unobserved attributes of a region and its neighbors. For instance spatial dependence in performance exists if the growth of neighboring industries and clusters influences own-industry growth. Similarly, unobserved attributes of the neighboring regions, such as human capital composition, may induce spatial dependence in the error terms. We take into account this potential spatial dependence directly by including the cluster specialization of adjacent regions in our main specifications.

[8] For detailed information on the LBD data see Jarmin and Miranda (2002). Other papers that examine start-up formation based on this detailed Census Bureau data include Armington and Acs (2002), Glaeser and Kerr (2009) and Kerr and Nanda (2009).

[9] There are 179 BEA-defined EAs covering the entirety of the United States. To minimize concerns about differences in transportation costs and the definition of neighboring regions, we exclude the Alaska and Hawaii EAs. The boundaries of EAs are drawn to reflect meaningful economic regions, ensure comprehensive regional coverage and have been highly stable over time (Johnson and Kort, 2004). EA include both rural and urban areas, facilitating the mapping of clusters that span urban and proximate rural areas (Porter, et al., 2004).

measurement of complementary economic activity in a consistent and unbiased manner is a considerable challenge,[10] the US Cluster Mapping Project (USCMP) develops a methodology for grouping four-digit (and some three-digit) SIC codes into cluster and linked cluster groupings.[11] The methodology first distinguishes between three "types" of industries with very 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 underlying hypotheses, we focus exclusively on the traded industries, where the relationship between start-up activity and cluster-driven agglomeration is likely to be most salient. These traded 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.[12] Porter (2001, 2003) assigns each traded industry into one of 41 mutually exclusive traded clusters (referred to as "narrow clusters"), where the set of industries included in each cluster primarily reflects pairwise correlations of industry employment across locations (Appendix A provides a comprehensive list of the 41 traded clusters, and key attributes).[13] Examples of clusters include automotive, apparel, biopharmaceuticals, and

[10] A small literature considers alternative classification schemes. Ellison and Glaeser (1997) study the coagglomeration of manufacturing industries, creating an index reflecting "excess" concentration. Feldman and Audretsch (1999) group those manufacturing 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 and/or market proximity (Scherer, 1982; Jaffe, Trajtemberg and Henderson, 1993; Bloom, Schankerman and Van Reenen, 2005). Finally, Ellison, Glaeser and Kerr (2007) test various mechanisms inducing co-agglomeration of industries, and conclude that input-output linkages are the most relevant factor followed by labor pooling. This reasoning is consistent with the methodology developed in Porter (2001, 2003). *See also* Feser and Bergman (2000) and Forni and Paba (2002).

[11] For consistency with Porter's cluster definitions, we focus on industries that are included in the publicly available County Business Patterns data (i.e., private sector non-agricultural production, non-household, and non-railroad employment). In order to use industry data back to 1990, the analysis employs SIC system rather than the more refined NAICS systems, which was introduced in 1997 (and modified in 2002). By construction, recent NAICS-based data can be translated (with some noise) into the older SIC system.

[12] Traded industries account for over 87% of domestic US patents and 30% of total US employment (Porter, 2003). In contrast, local industries do not agglomerate (are *not* localized) and focus on local demand.

[13] While the co-location of industries in a region does not guarantee interaction or spillovers (Boschma 2005, Torre 2008), consistent co-location across many regions strongly suggests that such interactions are present. However, it is possible that in a few cases industries with high co-location of employment across regions may have little economic relationship. Thus, in the USCMP two adjustments are made to the cluster definitions to eliminate spurious correlations. First, the four-digit SIC industry definitions and list of products and services are used to reveal the presence of logical externalities. Second, the National Input-Output accounts are used to look for meaningful cross-industry flows (see Porter, 2003 pp. 563). While industries that have meaningful economic interactions tend to co-agglomerate in space (see Ellison and

information technology. Within a cluster such as information technology, 9 individual 4-digit SIC code industries are incorporated, including electronic computers (SIC 3571) and software (SIC 7372). These cluster definitions form our key measures of complementary economic activity.

Variable Definitions and Sample Description

Entrepreneurship. Our two main measures of entrepreneurial activity are start-up employment and start-up establishments by new firms within a given EA-industry. By focusing on start-up firms (those opening their first establishment), our measure offers a reasonably proxy for the level of entrepreneurial activity in a given industry and location at a given point in time. Specifically, *Start-up Employment* is defined as total level of employment in new firms during their first year of operation (with payroll); and *Start-up Establishment* is the count of these new firms. Consistent with prior work (Armington and Acs, 2002; Glaeser and Kerr 2009), we computer 4-year averages for these annual start-up activities.[14] Using a multi-year span (and including a Census-year in the base and terminal period) both allows for a more informative signal of the true level of entrepreneurial activity and also significantly reduces the number of EA-industries in which we observe zero entrants during a given period.[15]

One of the main goals 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($Start\text{-}up\ Activity_{i,r,2002\text{-}05}$ / $Start\text{-}up\ Activity_{i,r,1991\text{-}94}$)-- must account for the fact that there are many EA-industries in which there is a zero level

---

Glaeser 1997), we recognize that there are non-geographical dimensions of proximity that could also facilitate the interactions between industries and their firms. For example, Boschma (2005) and Torre (2008) suggest that institutional, organizational, and temporary geographical proximity may be as important as geographical proximity in facilitating knowledge transfers among firms. Our cluster definitions can only indirectly capture these dimensions to the extent that they complement geographical proximity.

[14] When we aggregate the data at the region-industry level, we exclude establishments with missing industry information, but include these observations to compute region-level and US-level totals. Additionally, we drop establishments with zero employment and with very low wages (below half of the minimum wage) or very high wages (above $2 million USD). Our key findings are robust to including these wage outlier observations.

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

of employment (i.e., non existing regional industry) or, relatedly, a zero level of start-up activity in the study period. In either of these cases, we are required to either exclude those observations with a zero value in the starting or ending period, or impose a positive lower bound on the level of startup activity.

To convey our main results in the most concise way, we focus the bulk of our analysis on a sample where we include EA-industries that have a non-zero level of employment during 1990, and then focus on the growth rate in start-up activity among those EA-industries where there is a pre-existing level of economic activity (thus making a growth rate analysis meaningful). The resulting sample consists of 53,213 EA-industries. While excluding EA-industries with zero employment is meaningful, our core findings are robust to alternative treatments of this data issue that we discuss below.

To include in the analysis EA-industries where we observe zero start-up activity in either the baseline or terminal periods, we set a minimum level of start-up activity of 1 employee and 0.01 establishments.[16] We also demonstrate that our results are robust to the (un-scaled) subsample which conditions on a positive level of EA-industry start-up activity in both the baseline and terminal period (sample of 11,981 EA-industries).[17]

Interestingly, despite the fact that our regional clusters include both regions and industries units that are quite narrow, a very high share of EA-clusters experience at least a minimal level of entrepreneurial activity – for example, 85% of all EA-clusters have at least one start-up establishment during the 1990-2005 period. This suggests that the large number of EA-industries with zero start-ups is not due to the lack of start-up activity within-clusters, but likely due to the small scale of some industries and EAs.

Finally, we further account for the large number of zeros by examining the impact of the industry and cluster environment on the level of entrepreneurial activity, using all potential EA-industry pairs (i.e., 588 industries by 177 EAs). These analyses include probit specifications that directly evaluate whether "missing" EA-industry start-up

[16] In other words, we scale the start-up activity indicators by adding the minimum annual start-up employment and start-up establishments in the sample, which is a standard procedure to scale variables (see e.g., Glaeser and Kerr, 2009). Since the average start-up establishment in a EA-industry is less than one, we scale the number of start-up establishments by adding the smallest number in the sample (0.01).

[17] While more than 60% of the 53,213 EA-industries experience some start-up activity over the 1990-2005 period, only 11,981 experience start-up activity in both the base and terminal periods. This high skewness of start-up activity in regional industries has been documented in other studies that use narrow regional and/or industry units (see e.g., Glaeser and Kerr, 2009; Rosenthal and Strange, 2003).

activity is related to the overall cluster environment, and count models that explicitly accounts for the skewed (count data) distribution of start-up activity in our dataset (see Table 7).

There are large differences across EAs and clusters in the level of start-up activity. At the cluster level, the average (1990-2005) annual start-up establishments (as % of establishments in the cluster) varies from a maximum of 4.4% in Business Services to a minimum of 0.59% in Power Generation and Transmission (See Table A1). At the EA level, the annual start-up establishment rate during 1990-2005 (as % of traded establishments in the EA) is 2.60% on average (with a standard deviation of 0.79). The top-EAs by the rate of start-up establishments include Las Vegas-Paradise-Pahrump (NV) and Austin-Round Rock (TX), and the bottom-two EAs are Grand Forks (ND-MN) and Mason City (IA) (See Table A3).

While we primarily focus on start-up establishments, for robustness we also look at total new establishments (including both start-up firms and new establishments of existing firms). We expect the strength of the cluster environment in a location will foster the entry of both types of new establishments. To test this, we compute two related entry indicators. *Entry employment* is defined as the level of employment in all new establishments within a given EA-industry; and *entry establishment* is the count of these new establishments. Δ *Entry Employment* is then defined as ln($Entry\ Employment_{i,r,02\text{-}05}$ / $Entry\ Employment_{i,r,91\text{-}94}$), and Δ *Entry Establishments* is ln($Entry\ Establishment_{i,r,02\text{-}05}$ / $Entry\ Establishment_{i,r,91\text{-}94}$).

Finally, to test for the role of clusters in the performance of start-up firms, we examine the level of employment in young (up to five years old) start-ups in regional industries. Specifically, $Employment\ in\ Start\text{-}up\ Survivors_{i,r,2004\text{-}05}$ is defined as the average annual level of employment (over 2004-2005) in start-up firms borne during 2001-2003.

Industry 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 a body of prior work which uses location quotients (LQ) as a primary measure of regional specialization (Glaeser, et al. 1992, Feldman and Audretsch 1999, Porter 2003, among others). 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: $\text{INDUSTRY SPEC}_{\text{Employ},i,r,90} = \dfrac{\text{employ}_{i,r} / \text{employ}_{r}}{\text{employ}_{i,US} / \text{employ}_{US}}$, where *r* and *i* indicate the region (EA) and the industry, respectively. This indicator captures to what extent the industry is "over-represented" (in terms of employment) in the EA. Note that the specialization indicators are based on employment in the start-up employment models, and based on establishments in the start-up establishment models. In the data, the employment-based industry specialization of regions has a mean of 2.01 and a standard deviation of 6.42 (the establishment-based industry specialization has a mean of 1.78 and standard deviation of 3.20; Table 1). As mentioned earlier, we include region and industry fixed effects, and so the independent variation in our main empirical specifications is driven exclusively by variation in employment in the region-industry.

Cluster Specialization. We utilize an analogous procedure to develop a measure for cluster specialization. For a particular EA-industry the specialization of the EA in cluster *c* is measured by the share of regional employment in the cluster (outside the industry) as compared to the share of US total employment in the national cluster (outside the industry): $\text{CLUSTER SPEC}_{\text{Employ},icr,90} = \dfrac{\text{employ}_{c,r}^{\text{outside } i} / \text{employ}_{r}}{\text{employ}_{c,US}^{\text{outside } i} / \text{employ}_{US}}$. The average (employment-based) cluster specialization is 1.21 (and the standard deviation is 2; Table 1). Since this measure of specialization is relative to the overall size of the region, a region may exhibit specialization within a particular cluster even though that region only maintains a small share of the overall national employment of that cluster. While it is not surprising that leading regions in the automotive cluster include Detroit-Warren-Flint (MI) and Cleveland-Akron-Elyria (OH), there are also pockets of automotive cluster strength in smaller regions, such as Lexington-Fayette-Frankfort-Richmond (KY) and Louisville-Elizabethtown-Scottsburg (KY-IN) (Figure A1). It is useful to note that, with the inclusion of region and industry fixed effects and a measure of industry specialization, the independent variation that is utilized in the regression comes exclusively from the employment within a given cluster (outside the industry).

Table 2 illustrates key attributes of the top regional clusters based on cluster specialization in 1990.[19] These top clusters tend to have a higher level of start-up activity, a larger average size of establishments as well as a greater mix of older incumbents than other regional clusters. Interestengly, on average over 20% of all the establishments in a regional cluster belong to firms that have establishments in more than one geographical market (EA), and this figure increases to over 30% in the top regional clusters. This fact suggests that clusters are not isolated geographical units and they may establish linkages with other locations, in part facilitated by the presence of these regionally diversified firms.[20]

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 defined in Porter (2003). Specifically, while the narrow cluster definitions used for the earlier measure come from a classification scheme in which each industry is assigned to a unique cluster, Porter (2003) also develops a broad cluster definition in which each industry may be associated (measured by locational correlation of employment) with multiple clusters. To develop a measure based on linkages to cluster *c*, we include 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 (see Table A2).[21] 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:

[19] We define the set of top EA-clusters by selecting the top-10 EAs by Custer Specialization for every cluster. The high cluster specialization criterion is complemented with a minimum threshold for the share of national cluster employment (above values that correspond to the 20th percentile) to limit the concern about small regions with very low employment and very large location quotients.

[20] While systematic quantitative analysis is still missing, the view is that in response to global competition clusters are becoming more specialized on certain activities within the value chain and increasing their participation in national and global value chains (see e.g., Porter, 1998b, Bresnahan and Gambardella, 2004; Ketels and Memedovic, 2008). For example, multi-plants and multinational firms can use their network of subsidiaries to coordinate across clusters (Dunning 1998; Rugman and Verbeke, 2003; Manning, 2008).

[21] Clusters with linkages with many other traded clusters include analytical instruments and communications equipment, among others; while clusters with few connections to other clusters include tobacco and footwear.

$\omega_{c,j}=\text{Avg}\left(\frac{\text{shared industries}_{c,j}}{\text{total industries}_c}, \frac{\text{shared industries}_{j,c}}{\text{total industries}_j}\right)$.[22] The strength of a region in clusters linked to cluster *c* is then defined by a weighted sum of the location quotients associated with each linked cluster:

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

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 (Table A2).

We also develop a measure of the presence of like clusters in neighboring regions. In part, we include this measure based on the simple empirical observation that specialization in a particular cluster tends to be spatially correlated across neighboring regions – the historical strength of the automotive cluster near Detroit is likely reinforced by cluster specialization in automotives in neighboring EAs such as Grand Rapids-Muskegon-Holland (MI), Toledo-Fremont (OH) and Fort Wayne-Huntington-Auburn (IN). To explore the role of neighboring clusters in start-up growth in a region-industry, we compute the (average) specialization of adjacent regions in the cluster (including the focal industry). In other words, the strength of neighboring clusters is measured by the average LQ of the adjacent regional clusters.

Finally, as mentioned earlier, we include a complete set of region and industry fixed effects in our main specifications, and so control for unobserved factors that may be correlated with these measures of industry and cluster specialization, including regional or industry-level demand shocks, regional policies, or national regulatory changes that affect all firms within certain industries.

[22] For example, automotive has 5 narrow industries (out of 15) in common with production technology, and production technology shares 7 narrow industries (out of 23) with automotives; the degree of overlap between these two clusters is then $\omega_{c,j} = .32$.

5. Results

We now turn to our key findings. The sample consists of a cross-section of region-industries in 177 EAs and 588 four-digit SIC code industries, grouped into 41 traded clusters. After conditioning on those EA-industries where we observe a positive level of employment in the base period, our core sample consists of 53,213 observations.

Our analysis begins in Table 3 where we compute the average start-up growth rates for region-industries based on their *initial* levels of start-up activity and cluster specialization. Specifically, we divide all region-industries into four categories based on whether they are above or below the median level of start-up employment (for their industry), and above or below the median level of cluster specialization (for their industry). There are striking differences in the start-up employment growth rate across these conditions. Whereas region-industries with a relatively high level of start-up employment and a low level of cluster specialization experience, on average, a 33% decline in the rate of start-up activity between 1991-1994 and 2002-2005, those region-industries with low levels of start-up employment and a high level of cluster specialization register a 36% growth rate on average during this period. There is a statistically and quantitatively significant increase in the average start-up employment growth rate when moving from high initial level of start-up employment to a low initial level of start-up employment, consistent with a mean-reversion process in start-up activity. However, regardless of the initial level of start-up employment, there is a statistically and quantitatively significant increase in the growth rate of start-up activity when one moves from a region-industry with a low level of cluster specialization to one with high level of cluster specialization.[23] In other words, those regional industries that are located in a relatively strong cluster experience much higher growth rates in entrepreneurial activity.

While the sharp contrasts in Table 3 are intriguing, it is of course possible that alternative factors, such as industry specialization, or industry and region effects, are driving these striking results. We therefore turn in Table 4 to a more systematic regression analysis. The dependent variable is the start-up growth rate between the

[23] We find the same conclusions when looking at start-up establishment growth; and using the no-zeros subsample.

baseline period (1991-1994) and a terminal period (2002-2005). In (4-1), we include only the level of start-up employment, the level of industry specialization and the level of cluster specialization. The results provide evidence for the two main findings of this paper. First, there is a large convergence effect – a doubling of the initial level of start-up employment is associated with a 31% 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 influence on the growth rate of entrepreneurship. Both industry and cluster specialization are associated with higher growth rates of start-up activity.

These results are reinforced in (4-2), where we incorporate both 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 rate of start-up activity. This latter finding is consistent with the hypothesis that, while a strong local cluster environment enhances incentives for entrepreneurship, 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, thus controlling for unobserved shocks such as an increase in demand or a change in the policy environment towards a particular industry. The main results concerning mean reversion and the impact of cluster specialization are robust. The only meaningful change in the estimates concerns the impact of neighboring clusters; not surprisingly, given that the sign on the coefficient on *Cluster Spec in Neighbors* is ambiguous from the perspective of theory, the estimated coefficient depends on whether we control for industry and region heterogeneity.

Finally, in (4-5), we conduct a robustness check in which we drop all region-industries in which we observe zero start-up employment during either the base period or the terminal period (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 not only robust to a subsample that focuses on those region-industries with a positive level of start-up activity in both the base and terminal period, but the estimated coefficients on each of the parameters associated with the cluster environment increase in a meaningful way.

To illustrate the size of the effects, consider a one-standard deviation shift in each of the measures of industrial and cluster specialization using the coefficient estimates (and sample) from (4-5).[25] A one standard-deviation increase in industry specialization (4.26) is associated with a 11% increase in the annual start-up employment growth rate, while a one standard-deviation shift in cluster specialization (1.91) is associated with a 1.6% increase in the expected annual start-up employment growth rate.[26] In other words, after controlling for the impact of convergence, there is a quantitatively important impact of related economic activity on start-up activity.

The core findings persist when we shift attention towards measures of entrepreneurship and specialization based on the number of establishments (rather than the total employment within those new establishments). In Table 5, both the dependent and independent variables are now based on counts of establishments, and the structure of the specifications mirrors the logic of Table 4. We begin in (5-1) by including the level of start-up establishment activity, industry specialization and cluster specialization; we add the linked cluster and neighboring cluster measures and control for the total establishments in the region in (5-2); include region and industry fixed effects in (5-3) and (5-4); and condition on a sample that excludes all region-industries with zero start-up establishments in either the baseline or terminal period in (5-5). Our core results concerning the convergence effect and the impact of cluster specialization are robust. Interestingly, the only significant difference in the results concerns the estimated effect of industry specialization. Whereas the coefficient on industry specialization was positive in the start-up employment models (Table 4), the coefficient on industry specialization in

[25] We focus our analysis of the magnitudes on this latter specification since this subsample is not subject to the re-scaling of the dependent variable that we implement in (4-1)-to-(4-4).
[26] Alternatively, an increase in the level of industry specialization (cluster specialization) from the 25th to the 75th percentile value is associated with a 3.4% (0.7%) increase in the annual start-up employment growth rate.

the first two columns of Table 5 is negative (and significant).[27] While the coefficient becomes positive and significant when we include region and industry (or only industry) fixed effects, the heterogeneity of this parameter across specifications is consistent with the fact that the effect of industry specialization on entrepreneurial growth is ambiguous. In contrast, the coefficient on the impact of clusters (both using the narrow cluster definition as well as the impact of "linked" clusters) is positive across all specifications.

Finally, in Table 6, we consider an alternative measure of "new" economic activity by examining the growth in employment in "new" establishments (even if the firm may already exist in other locations) and counts of new establishments (including new establishments by already existing firms). While this measure combines "pure" entrepreneurship with more traditional types of business expansion, the opening up of new establishments (and employing a significant number of workers in those establishments) is a crucial channel by which entrepreneurial firms grow over time and contribute to aggregate economic performance. Each of the specifications in Table 6 includes region and industry fixed effects, and only vary in the number of measures of related economic activity that are included (i.e., whether the linked cluster and neighboring cluster variables are included) and whether the dependent and explanatory variables are measured based on employment (models 6-1 and 6-2) or based on counts of new establishments (models 6-3 and 6-4). The results are robust across all of the specifications – there is a mean-reversion effect in the data consistent with convergence, and a positive impact of cluster strength and scale on the growth rate of new establishments and of the employment in new establishments. The core findings hold when we examine only new establishments of existing firms (i.e., "new subsidiaries". Interestingly, these new establishments often belong to firms that operate in clusters in other locations (EAs) -- for example in 1990, more than 85% of these new subsidiaries belong to multi-location firms. While more research is needed on the locational and organizational decisions of multi-establishment firms, the findings suggests that firms may be opening establishments in new locations to seek complementary clusters and benefit from the competitive resources of each location.

[27] The coefficient becomes positive if we include the (log of) the number of establishments in the EA-industry; suggesting that the scale of the EA-industry matters for start-up establishment growth.

In Table 7 we examine the level of entrepreneurial activity using all potential region-industry pairs (i.e., 588 industries by 177 EAs) to further account for the large number of EA-industries with zero start-up activity. First, we implement a probit specification that evaluates whether the existence of start-up activity during 2002-2005 is related to the industry and the overall cluster environment while controlling for the existence of any start-up experience in an earlier period (*Positive Start-up Activity* during 1990-1996) and including region and industry fixed effects (7-1).[29] We report the marginal effects to facilitate the interpretation of the coefficients, and find that agglomeration forces that take place in the industry and in the set of related industries within the regional cluster and in neighboring clusters increase the probability of experiencing start-up activity in the future.[30]

In model 7-2 we use the same explanatory variables to examine the total count of start-up establishments during 2002-2005 using a fixed effects Negative Binomial model. The estimated incidence-rate ratios suggest that the strength of the cluster has more than a 33% boost on the subsequent count of start-up establishments; the local industry has also a large effect (29% boost), and the linked clusters and neighboring clusters matter but to a lower extent.[31] Finally, these findings are confirmed when we implement a panel specification to examine the annual count of establishments in 1997 and 2002 Census years using a Negative Binomial model with year, region and industry fixed effects (7-3 and 7-4).[32] Drawing on the dynamic count data model developed by Blundell et al., (1995), we include two alternative indicators of the pre-existing start-up activity in a region-industry, the existence of any start-up activity from 1990 to 1996 (7-3) and the average annual start-up activity during that period (7-4) to control for the unobservable heterogeneity of region-industries. Overall our findings suggest that the cluster environment facilitates the formation of new businesses in regional traded industries.

---

[29] In order to include the EA-industries with zero employment (i.e., zero industry specialization), we replace industry specialization with the minimum positive value of this variable.

[30] We also define the industry and cluster environment based on employment (versus count of establishments) and the same findings hold. Furthermore, the findings are robust to dropping the observations with zero employment in the base year (i.e., using our core sample of 53213 observations).

[31] We also examine the count of start-up employment and find robust evidence for the contribution of the industry and the cluster, while the effect of linked clusters and neighboring clusters gets noisier.

[32] As mentioned earlier, Census-years are preferred because they offer a better coverage of all new firms.

While the focus of this paper is the formation of new businesses, we also examine the role of clusters on the (medium-term) performance of start-up firms. In particular, in Table 8 we study the level of employment in young start-up survivors in a region-industry using our core econometric specification. We find that the cluster environment contributes to improve the level of employment in young start-up survivors, suggesting that clusters facilitate the survival and potentially the growth of successful start-ups. In related work, we will examine more carefully the role of clusters in the dynamics of new businesses and the attributes of successful start-ups.

Finally, 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 used a less nuanced measure of complementary activity, in which the elements of each "cluster" are defined as all traded four-digit SIC code industries in each two-digit SIC code.[33] We have included in the model a dummy indicator equal to one for region-industries with positive start-up activity in the base period to capture unobserved factors that may influence the entrepreneurial opportunities in a local industry. Regarding the sample, we have varied both the length and precise dates of both the base and terminal periods to test for sensitivity to other periods. Specifically, we consider the 1990-2001 period and define the base and terminal periods based on 5-year averages (versus 4-year averages). By using this shorter period, we also reduce some of the noise of translating the older SIC system into the new 2002 NAICS. Furthermore, we have considered a larger (but noisier) sample by including those establishments with very noisy wages (too low or too high).[34] Finally, we have examined whether the impact of clusters is particularly salient in certain types of regions (e.g., large versus small) and clusters (manufacturing-oriented versus service-oriented).[35]

[33] In other words, we estimate models (4-3) and (5-3) but using an indicator of the specialization of the region in other two-digit SIC traded industries (instead of the cluster specialization variable), and find that the presence in the region of other (two-digit SIC) related industries contributes to start-up activity growth.
[34] The new sample consists of 54,001 EA-industries with positive employment in the base period.
[35] Interestingly, we find that in larger regions the convergence effect is smaller and the cluster effects are larger. We also explore how the convergence and cluster effects vary across different types of clusters (manufacturing-oriented vs. service-oriented) and find that service-oriented clusters have the lowest convergence and largest cluster-driven agglomeration benefits in entrepreneurship growth.

6. Conclusion and Extensions

This paper finds striking evidence for the simultaneous yet distinct influences of agglomeration and convergence on the growth of 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 complementary economic activity -- most notably, the presence of strong local clusters -- accelerates the growth in start-up activities. We find that industries located within a strong cluster or that can access strong linked clusters are associated with higher start-up employment growth rates and higher entry of new firms.

These findings offer an important and novel 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, 1998), this paper moves beyond the traditional debate in which the presence of related economic activity simultaneously indicated the presence of complementarities as well as competition for inputs and customers, clouding the interpretation of any particular empirical finding. Instead, by first accounting for convergence and the potential for competition within each industry in a region, we are able to isolate the separate and quantitatively important impact of cluster-related complementarities on entrepreneurship. In other words, while at a (narrow) industry level firms 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. Cluster-driven entrepreneurship is, moreover, a dynamic process, as the new business creation at one point in time spurs further start-up activity on an ongoing basis.

While our analysis has focused at new business formation, we also examine young start-up survivors. We find that clusters contribute to the level of employment in young start-ups in regional industries, suggesting that a strong cluster environment enhances the potential for growth for the most productive start-ups. The impact of the cluster environment on 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 performance of start-up survivors, even after controlling for region and industry heterogeneity. There is large heterogeneity in the types of entrepreneurship that we could further explore. Start-up firms will differ in size, innovative-orientation, and their growth potential. Similarly, we could also exploit the attributes of the firms that participate in clusters (size, age, regional diversification). Further understanding the demography of entrepreneurship and clusters will help design more effective policies to promote entrepreneurship.

In the paper we focus exclusively on traded industries since they are exposed to greater competition and account for most innovations. In future work we will examine how the effect of cluster-driven agglomeration forces on entrepreneurship varies for local industries versus traded industries.

Finally, we find that clusters also matter for the formation of new subsidiaries of existing firms. These new subsidiaries often belong to firms that operate in clusters in other locations (i.e., "multi-cluster" firms), suggesting that firms may seek complementary clusters, benefiting from the competitive resources of each location. Similarly, the International Business literature emphasizes that in a global economy, local clusters integrate in national and global value chains often through the complex network of subsidiaries of MNCs (Dunning 1998; Porter 1998b; Enright, 2000; Manning, 2008). We conjecture based on our analysis of new establishments of existing firms that strong clusters may also facilitate the entry of foreign subsidiaries to a location. A related important questions raised by international business studies is what attributes of US clusters are more attractive for the location of (domestic and foreign) multi-establishment firms, and how the participation in multiple clusters affect the strategy, internal organization and performance of these companies. Furthermore, the contribution of these firms to generate entrepreneurial opportunities in a particular regional cluster is an open research question that we examine in related work.

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.

Acs, Z.J., E. Glaeser, R. Litan, L. Fleming, S. Goetz, W. Kerr, S. Klepper, S. Rosenthal, Ol. Sorenson, and W. Strange, 2008, "Entrepreneurship and Urban Success: Toward a Policy Consensus," Kaufman Foundation Report.

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, 1992, "Convergence," *Journal of Political Economy* 100(2), pp. 223-51.

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.

Bloom, N., M. Schankerman and J. Van Reenen, 2005, "Identifying Technology Spillovers and Product Market Rivalry," CEP Discussion Paper, 675.

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.

Bonte, W, 2004, "Innovation and Employment Growth in Industrial Clusters: Evidence from Aeronautical Firms in Germany," *International Journal of the Economics and Business* 11 (3), pp. 259-278.

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.

Ciccone, A. and R. Hall, 1996, "Productivity and the Density of Economic Activity," *American Economic Review* 86(1), pp. 54-70.

Cingano F. and F. Schivardi, 2004, "Identifying the Sources of Local Productivity Growth," *Journal of the European Economic Association* 2(4), pp. 720–742

Cortright, J., 2006, "Making Sense of Clusters: Regional Competitiveness and Economic Development," *The Brookings Institution Metropolitan Policy Program*, http://www.brookings.edu/reports/2006/03cities_cortright.aspx.

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

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.

Dumais, G., G. Ellison, E.L. Glaeser, 2002, "Geographic Concentration as a Dynamic Process," *Rev. Econ. and Stat.* 84 (2), pp. 193-204.

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, 1989, "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.

Duranton, G. and Puga, D., 2001, "Nursery Cities: Urban diversity, process innovation, and the life cycle of products," *American Economic Review* 91(5), pp. 1454-1477.

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.

Feldman, M.P., 1994, "The University and High-Tech Start-ups: The Case of Johns Hopkins University and Baltimore," *Economic Development Quarterly* 8, pp. 67-77.

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.

Forni, M and S. Paba, 2002, “Spillovers and the Growth of Local Industries,” *Journal of Industrial Economics* 50, pp. 151-171.

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

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

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

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).

Glaeser, E.L, W.R. Kerr, and G.A. M. Ponzetto, 2009, “Clusters of Entrepreneurship,” *Journal of Urban Economics* (forthcoming).

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.

Greenstein S., C. Forman and A. Goldfarb, 2008, “Understanding the Inputs into Innovation: Do Cities Substitute for Internal Firm Resources?” *Journal of Economics and Management Strategy* 17(2).

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,” Working Paper.

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.

Higgins, MJ., D. Levy, and AT Young, 2006, “Growth and Convergence across the United States: Evidence from County-Level Data,” *Review of Economics and Statistics* 88(4), pp. 671-681.

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.

Ketels, C. and O. Memedovic, "From Clusters to Cluster-Based Economic Development," *International Journal of Technological Learning, Innovation, and Development* 1, pp. 375-392.

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, "The Geography of Organizational Knowledge," manuscript.

Krugman, P., 1991, "Increasing Returns and Economic Geography," *Journal of Political Economy* 99(3), pp. 483–499.

Malecki, E.J., 1990, "New Firm Formation in the USA: Corporate Structure, Venture Capital, and Local Environment," *Entrepreneurship and Regional Development* 2, pp. 247-265.

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.

Magrini, S., 2004, Regional (Di)Convergence, in J. V. Henderson and J. F. Thisse (eds.) *Handbook of Regional and Urban Economics,* 4, Amsterdam: Elsevier North-Holland, pp. 2243-2292.

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

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.

Paci R. and S Usai, 2000, "The Role of Specialization and Diversity Externalities in the Agglomeration of Innovative Activity," Rivista Italiana degli Economisti 2(2), pp. 237-268.

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.

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.

Porter, M.E. and S. Stern, 2003, "The Impact of Location on Global Innovation: Findings from the National Innovative Capacity Index," in P. K. Cornelius (ed.), *The Global Competitiveness Report 2002-2003*, New York: Oxford University Press, pp. 227-253.

Porter, M.E., et al., 2004, "Competitiveness in Rural U.S. Regions: Learning and Research Agenda," Economic Development Administration, U.S. Department of Commerce.

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

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.

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

Rugman, A. M. and A. Verbeke, 2003, "Multinational Enterprises and Clusters: An Organizing Framework," *Management International Review*, 43(3), pp. 151–169.

Shane, S., 2001, "Technological Opportunities and New Firm Creation," *Management Science* 47, pp. 205–220.

Scherer, F. M., 1982, "Inter-industry Technology Flows and Productivity Growth," *The Review of Economics and Statistics* 64(4), pp. 627-634.

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.

Swann, G.M. P., 1998, "Clusters in the US Computing Industry," in 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, pp. 76-105.

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

Venables, A.J., 1996, "Equilibrium Locations of Vertically Linked Industries," *International Economic Review* 37(2), pp. 341-360.

Wennberg, K. and G. Lindqvist, 2008, "The Effect of Clusters on the Survival and Performance of New Firms," *Small Business Economics*.

Table 1: Region-Industry Descriptive Statistics

	EA-Industries with positive employment in 1990*		
			"No-zeros" Positive start-up activity in 1991-94 & 2002-05
		N=53213	N=11981
Variables	**Definition**		
$\text{ESTABLISHMENTS}_{90}$	EA-industry establishments	16.16 (71.23)	56.00 (142.01)
$\text{STARTUP ESTABLISHMENT}_{91\text{-}94}$	EA-industry (4-year average) annual startup establishments (91-to-94)	.47 (2.59)	1.87 (5.21)
$\text{ENTRY ESTABLISHMENT}_{91\text{-}94}$	EA-industry (4-year average) annual entry establishments (91-to-94)	0.82 (5.28)	2.54 (9.39)
EMPLOYMENT_{90}	EA-industry employment	554.67 (2464.7)	1544.09 (4778.32)
$\text{STARTUP EMPLOYMENT}_{91\text{-}94}$	EA-industry (4-year average) annual startup employment	6.37 (86.42)	20.99 (179.05)
$\text{ENTRY EMPLOYMENT}_{91\text{-}94}$	EA-industry (4-year average) annual entry employment	15.39 (146.70)	43.75 (262.82)
EMPLOYMENT IN STARTUP $\text{SURVIVORS}_{04\text{-}05}$	EA-industry (2-year average) annual employment in start-up firms borne during 2001-2003	13.61 (120.10)	51.55 (241.85)
Δ STARTUP ESTABLISHMENT	Growth rate in startup establishments $\log\left(\frac{\text{startup estab}_{i,r,02-05}}{\text{startup estab}_{i,r,91-94}}\right)$	.10 (1.78)	.04 (.91)
Δ STARTUP EMPLOYMENT	Growth rate in startup employment	.16 (1.02)	.47 (1.73)
Δ ENTRY ESTABLISHMENT	Growth rate in entry establishments	.07 (1.86)	.02 (.94)
Δ ENTRY EMPLOYMENT	Growth rate in entry employment	.14 (1.25)	.36 (1.80)
$\text{INDUSTRY SPEC}_{\text{Estab, 90}}$	Industry establishment-based Location Quotient. $LQ_{i,r} = \frac{\text{estab}_{i,r}/\text{estab}_r}{\text{estab}_{i,US}/\text{estab}_{US}}$	1.78 (3.20)	1.40 (2.57)
$\text{CLUSTER SPEC}_{\text{Estab, 90}}$	Cluster establishment-based LQ (outside the industry)	1.06 (1.14)	1.13 (1.25)
$\text{LINKED CLUSTERS SPEC}_{\text{Estab, 90}}$	Linked clusters' establishment-based LQ (weighted by cluster overlap)	.97 (.46)	.98 (.35)
CLUSTER SPEC in $\text{NEIGHBOR}_{\text{Estab, 90}}$	Neighboring clusters' average establishment-based LQ	1.03 (.80)	1.02 (.81)
$\text{INDUSTRY SPEC}_{\text{Employ, 90}}$	Industry employment-based LQ	2.01 (6.42)	1.47 (4.26)
$\text{CLUSTER SPEC}_{\text{Employ, 90}}$	Cluster employment-based LQ (outside the industry)	1.17 (1.95)	1.18 (1.91)
$\text{LINKED CLUSTERS SPEC}_{\text{Employ, 90}}$	Linked clusters' employment-based LQ (weighted by cluster overlap)	1.04 (.92)	1.01 (.70)
CLUSTER SPEC in $\text{NEIGHBOR}_{\text{Employ, 90}}$	Neighboring clusters' average employment-based LQ	1.13 (1.19)	1.06 (1.06)

Note: The core sample includes EA-industries with positive employment in 1990, resulting in 53213 observations. For this sample, the start-up and entry indicators are scaled by adding 1 employee and 0.01 establishments (before taking logs). The "No-zeros" subsample includes EA-industries with positive startup activity in both the base and terminal periods (1991-1994 and 2002-2005). For the Entry variables (i.e., those based on *all* the new establishments) the "No-zeros" sample consists of 16016 observations.

Table 2: Attributes of EA-Clusters in 1990 (Mean and Std. Deviation)

	Start-up Estab$_{91\text{-}94}$ (Annual, level)	**Multi-EA Estab** (rate)	**Avg. Size Estab**	**Young Estab** Age<5 (rate)	**Old Estab** Age>10 (rate)
Top-10 EA-Clusters$_{1990}$ (by Cluster) N=410	**8.8** **(29.16)**	**.313** **(.248)**	**247.761** **(394.20)**	**.249** **(.162)**	**.520** **(.210)**
Other EA-Clusters$_{1990}$ N=6220	**3.11** **(12.97)**	**.193** **(.217)**	**55.661** **(93.301)**	**.304** **(.214)**	**.426** **(.240)**
t-test of diff. of means	**5.95**	**10.684**	**28.273**	**-5.095**	**7.717**

Note: We define the set of top EA-clusters by selecting the top-10 EAs by Custer Specialization for every cluster (E.g., for automotive the top one local cluster would be in Detroit EA). The high cluster specialization criterion is complemented with a minimum threshold for the share of national cluster employment (above the 20th percentile values) to limit the concern about small EAs with very low employment and very high cluster specialization. *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 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)

		STARTUP EMPLOYMENT$_{1991\text{-}94}$	
		Low	**High**
CLUSTER SPEC$_{Employ, 90}$ (Outside the industry)	**Low**	**ΔSTARTUP EMPLOYMENT= .25** N= 20,507	**ΔSTARTUP EMPLOYMENT= -.33** N=6,265
	High	**ΔSTARTUP EMPLOYMENT= .36** N= 17,474	**ΔSTARTUP EMPLOYMENT= -.09** N=8,967

Notes: Low versus High is based on the median of the variable for each industry.
All the averages are significantly different from each other at 1%.

Table 4: EA-Industry Growth in Start-up Employment (N=53213)

	STARTUP EMPLOYMENT GROWTH				
					No zeros N=11,981
	1	2	3	4	5
Ln STARTUP EMPLOY$_{91-94}$	**-.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 NEIGHBORS$_{Employ, 90}$		**-.018**		**.031**	**.076**
		(.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, Bold-Italic numbers refer to coefficients significant at 1% and 5% 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)

	STARTUP ESTABLISHMENT GROWTH				
					No zeros N=11,981
	1	2	3	4	5
Ln STARTUP 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, Bold-Italic numbers refer to coefficients significant at 1% and 5% levels. 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 $ENTRY_{91\text{-}94}$	**-.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 Bold-Italic numbers refer to coefficients significant at 1% and 5% levels. Robust standard errors clustered by EA-Cluster. The explanatory variables are based on employment (establishments) in the entry employment (establishment) growth models.

Table 7: EA-Industry Level of Start-up Activity (Using the Full Sample)

	Probit (Marginal Effects)	**Negative Binomial** (Incidence-Rate Ratios)		
	POSITIVE START-UP ACTIVITY (during 2002-05)	START-UP ESTABLISHMENTS (during 2002-05)	ANNUAL START-UP ESTABLISHMENTS (1997, 2002)	
	1	2	3	4
POSITIVE START-UP	**.054**	**1.349**	**1.289**	
$ACTIVITY_{1990\text{-}96}$	**(.003)**	**(.016)**	**(.019)**	
AVG ANNUAL START-UP				**1.007**
$ESTABLISHMENTS_{1990\text{-}96}$				**(.000)**
Ln INDUSTRY $SPEC_{Est\ 90}$	**.011**	**1.286**	**1.360**	**1.362**
	(.000)	**(.005)**	**(.006)**	**(.006)**
Ln CLUSTER $SPEC_{Est\ 90}$	**.008**	**1.338**	**1.385**	**1.367**
(Outside the industry)	**(.002)**	**(.011)**	**(.012)**	**(.012)**
Ln LINKED CLUSTERS $SPEC_{Est\ 90}$	.004	**1.117**	**1.097**	**1.100**
	(.003)	**(.017)**	**(.019)**	**(.019)**
Ln CLUSTER SPEC in	**.025**	**1.107**	**1.044**	**1.046**
$NEIGHBORS_{Est\ 90}$	**(.003)**	**(.014)**	**(.015)**	**(.015)**
EA FEs	**Yes**	**Yes**	**Yes**	**Yes**
INDUSTRY FEs	**Yes**	**Yes**	**Yes**	**Yes**
YEAR FEs			**Yes**	**Yes**
R-Squared	.450			
Log-likelihood	-29470.9	-58403.9	-63412.1	-63307.1
Obs.	103368	103014	206028	206028

Note: In (6-1) the dependent variable is a dummy equal to one for EA-industries with positive start-up activity during 2002-2005; the coefficients are marginal effects from the probit model. In (6-2) we examine the count of start-up establishments during 2002-2005 using a Negative Binomial model, and the coefficients are the incidence-rate ratios. In (6-3) and (6-4) we examine the annual count of start-up establishments in two Census years (1997, 2002) using a Negative Binomial model. All models omit data for which all observations corresponding to a given fixed effect have zero observed outcome. Note that to avoid convergence problems, in the Negative Binomial models we combine the 10% smallest EAs by employment size in a single region fixed effect, including in the models a total of 160 EA dummies (vs. 177). Similarly, we include 529 industry fixed effects in the model (vs. 588) by combining the 10% smallest national industries. 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
			11,981
	1	2	3
Ln STARTUP EMPLOY$_{91\text{-}94}$	**.338**	**.334**	**.172**
	(.011)	**(.011)**	**(.011)**
Ln INDUSTRY SPEC$_{Employ,\ 90}$	**.100**	**.097**	**.227**
	(.003)	**(.003)**	**(.014)**
Ln CLUSTER SPEC$_{Employ,\ 90}$	**.020**	**.012**	**.075**
(Outside the industry)	**(.003)**	**(.003)**	**(.019)**
Ln LINKED CLUSTERS SPEC$_{Employ,\ 90}$		**.053**	**.095**
		(.008)	**(.028)**
Ln CLUSTER SPEC in NEIGHBORS$_{Employ,\ 90}$		***.016***	.039
		(.007)	(.026)
EA FEs	**Yes**	**Yes**	**Yes**
INDUSTRY FEs	**Yes**	**Yes**	**Yes**
R-Squared	**.283**	**.284**	**.375**

Notes: Bold, Bold-Italic, and Italic numbers refer to coefficients significant at 1%, 5% and 10% levels. Robust standard errors clustered by EA-Cluster. The dependent variable is the (log of) average annual employment (over 2004-2005) of start-up establishments borne in 2001-2003. We add 1 to the dependent variable before taking logs. The explanatory variables are in logs

Appendix A: Traded Clusters Attributes

Table A1: Entry by Traded Clusters (Annual averages over 1990-2005)

Name (41 traded clusters)	Type	# Narrow Industries		Startup Estab	New subsidiary Estab	Startup Employ (1,000)	New subsidiary Employ (1,000)	Startup Estab (% Estab)
		total	svc.					
Aerospace Engines	High-tech	2	0	7.6	5.1	0.229	1.820	1.4
Aerospace Vehicles & Defense	High-tech	6	0	25.2	10.6	0.861	5.449	2.0
Analytical Instruments	High-tech	10	0	164.2	58.3	5.404	7.041	1.7
Biopharmaceuticals	High-tech	4	0	44.0	25.0	2.476	2.220	2.1
Chemical Products	High-tech	21	0	90.4	95.7	2.070	3.915	1.2
Communications Equipment	High-tech	9	0	60.1	40.8	2.781	4.282	1.8
Information Technology	High-tech	9	3	410.0	403.5	7.431	14.127	3.2
Medical Devices	High-tech	8	0	89.9	39.2	2.005	3.591	1.8
Distribution Services	Service	19	19	2452.6	1741.2	19.584	40.042	2.5
Education & Knowledge Creation	Service	10	9	1042.6	490.6	16.399	16.107	2.9
Business Services	Service	21	21	10539.3	4457.4	90.254	116.404	4.4
Entertainment	Service	13	9	1597.9	207.9	23.572	12.545	4.0
Financial Services	Service	21	21	2241.0	8692.7	35.963	137.821	1.9
Heavy Construction Services	Service	19	6	2811.4	521.4	22.056	16.493	2.5
Hospitality & Tourism	Service	22	19	1805.8	924.0	29.968	38.447	2.6
Oil & Gas Products & Services	Service	12	6	377.1	277.3	4.312	8.032	2.4
Power Generation & Transmission	Service	6	1	16.6	98.5	0.632	3.472	0.6
Transportation & Logistics	Service	17	16	1186.2	1544.3	14.752	38.298	2.3
Agricultural Products	Other	20	6	527.6	53.4	3.504	2.055	3.1
Apparel	Other	27	0	454.2	46.5	8.727	5.323	3.9
Automotive	Other	15	0	304.8	87.9	8.836	10.849	1.8
Building Fixtures, Equipment & Services	Other	25	2	578.0	68.5	5.855	3.188	2.6
Construction Materials	Other	11	0	125.3	33.5	1.733	1.762	2.3
Fishing & Fishing Products	Other	3	0	57.0	8.2	0.496	0.732	3.0
Footwear	Other	5	0	15.6	3.7	0.460	0.380	2.6
Forest Products	Other	8	0	114.3	38.7	2.529	3.477	2.4
Furniture	Other	10	0	165.5	28.9	2.394	2.538	2.4
Heavy Machinery	Other	10	2	188.0	106.3	2.793	3.439	1.4
Jewelry & Precious Metals	Other	7	1	261.5	15.7	1.114	0.477	2.7
Leather & Related Products	Other	13	0	114.3	12.7	1.342	0.765	2.7
Lighting & Electrical Equipment	Other	10	0	66.7	24.7	2.009	1.939	1.4
Metal Manufacturing	Other	44	0	386.6	122.8	10.683	7.935	1.6
Motor Driven Products	Other	12	0	46.3	21.8	1.691	2.698	1.7
Plastics	Other	9	0	202.3	112.9	4.831	7.834	1.6
Prefabricated Enclosures	Other	12	0	55.0	20.8	1.670	1.946	2.0
Processed Food	Other	43	2	325.6	237.7	8.956	11.205	1.3
Production Technology	Other	23	0	192.5	67.2	4.038	4.315	1.6
Publishing & Printing	Other	26	3	887.1	330.1	9.717	11.995	2.4
Sporting, Recreational & Children's Goods	Other	3	0	95.9	10.5	1.299	1.345	3.5
Textiles	Other	20	0	120.2	27.9	3.773	3.472	2.5
Tobacco	Other	4	0					

Notes: Data sources: Longitudinal Business Database and US Cluster Mapping Project (Porter, 2003).
There are 589 traded four-digit SIC code industries (146 of them are service (svc.) industries).
Service clusters are those with more than 35% of employment in service industries.
High-tech clusters are manufacturing clusters with high R&D and patenting.
The last column reports startup establishments (as % of establishments in the cluster).
"New subsidiary" refers to new establishments of existing firms.

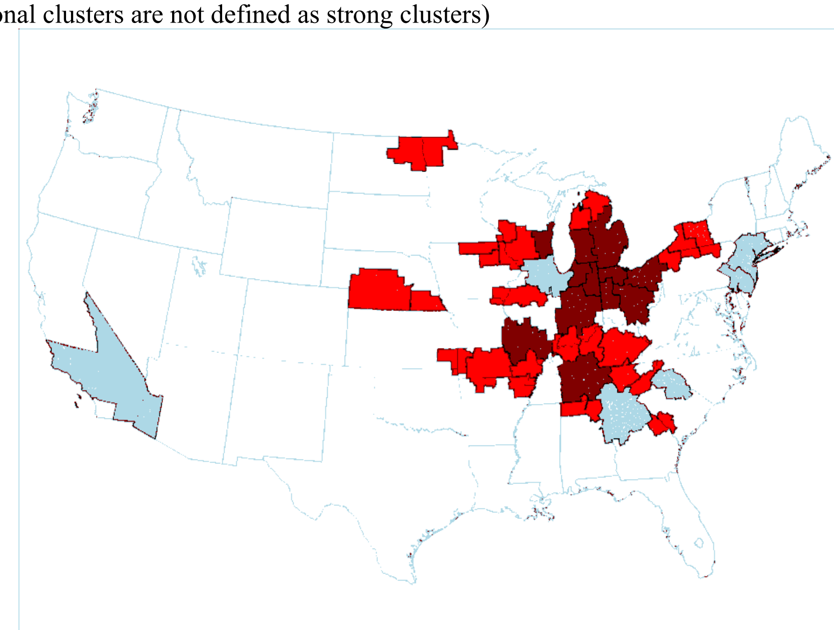
Table A2: Automotive (narrow) Cluster: Breadth of Industries and Linked Clusters

4-digit SIC Industries		Clusters Linked to Automotive (Shared industries) Production Technology	Metal Manufacturing	Heavy Machinery	Motor Driven Products	Aerospace Engines	Furniture
2396	Automotive and apparel trimmings						
3052	Rubber and plastics hose and belting						X
3061	Mechanical rubber goods					X	
3210	Flat glass						
3230	Products of purchased glass						
3322	Malleable iron foundries		X				
3465	Automotive stampings	X					
3519	Internal combustion engines, n.e.c.			X	X		
3544	Special dies, tools, jigs and fixtures	X	X				
3549	Metalworking machinery, n.e.c.	X	X				
3592	Carburetors, pistons, rings, valves		X				
3711	Motor vehicles and car bodies	X					
3714	Motor vehicle parts and accessories		X				
3799	Transportation equipment, n.e.c.						
3824	Fluid meters and counting devices	X					
	Cluster Overlap ($\omega_{c,j}$) with linked clusters	.32	.25	.08	.08	.03	.03

Porter's (2003) cluster definitions. The Automotive cluster has more than 30% overlap with the Production Technology cluster (by average of the percent of narrow industries shared in each direction)

Figure A1: Location of strong regional Automotive clusters (1997) (Top 20% of EAs by employment Location Quotient)

- EAs with high cluster specialization
- EAs with high cluster specialization and high share of US cluster employment (top 10% of EAs)
- EAs with high share of US cluster employment but without high cluster specialization (these regional clusters are not defined as strong clusters)



Source: Delgado, Porter and Stern (2007). Calculations based on US Cluster Mapping Project.

Table A3: Annual Start-up Activities in EAs (average 1990-2005): Top and Worse Performers

Top-5 EAs by start-up employment (% traded employment)	Top-5 EAs by start-up establishments (% traded establishments)
Miami-Fort Lauderdale-Miami Beach, FL	Las Vegas-Paradise-Pahrump, NV
Houston-Baytown-Huntsville, TX	Austin-Round Rock, TX
San Antonio, TX	Sarasota-Bradenton-Venice, FL
Pensacola-Ferry Pass-Brent, FL	Bend-Prineville, OR
Los Angeles-Long Beach-Riverside, CA	Miami-Fort Lauderdale-Miami Beach, FL

Bottom-5 EAs by start-up employment (% traded employment)	Bottom -5 EAs by start-up establishments (% traded establishments)
Mason City, IA	Grand Forks, ND-MN
Waterloo-Cedar Falls, IA	Mason City, IA
Kearney, NE	Aberdeen, SD
Cedar Rapids, IA	Waterloo-Cedar Falls, IA
Joplin, MO	Scotts Bluff, NE

Source: Authors' calculations based on LBD data. Start-up activities in the region are based on traded industries.