CLUSTERS, CONVERGENCE, AND ECONOMIC PERFORMANCE

by

Mercedes Delgado *
Temple University

Michael E. Porter *
Harvard University

and

Scott Stern *
MIT and NBER

CES 10-34 October, 2010

Abstract

This paper evaluates the role of regional cluster composition in the economic performance of industries, clusters and regions. On the one hand, diminishing returns to specialization in a location can result in a convergence effect: the growth rate of an industry within a region may be declining in the level of activity of that industry. At the same time, positive spillovers across complementary economic activities provide an impetus for agglomeration: the growth rate of an industry within a region may be increasing in the size and "strength" (i.e., relative presence) of related economic sectors. Building on Porter (1998, 2003), we develop a systematic empirical framework to identify the role of regional clusters – groups of closely related and complementary industries operating within a particular region – in regional economic performance. We exploit newly available data from the US Cluster Mapping Project to disentangle the impact of convergence at the region-industry level from agglomeration within clusters. We find that, after controlling for the impact of convergence at the narrowest unit of analysis, there is significant evidence for cluster-driven agglomeration. Industries participating in a strong cluster register higher employment growth as well as higher growth of wages, number of establishments, and patenting. Industry and cluster level growth also increases with the strength of related clusters in the region and with the strength of similar clusters in adjacent regions. Importantly, we find evidence that *new* industries emerge where there is a strong cluster environment. Our analysis also suggests that the presence of strong clusters in a region enhances growth opportunities in *other* industries and clusters. Overall, these findings highlight the important role of cluster-based agglomeration in regional economic performance.

* This paper has benefited greatly from the sustained research by Richard Bryden and Weifeng Weng in assembling the US Cluster Mapping Project dataset. Ben Jones, Jim Davis, Bill Kerr, Matthew Kahn, Jim Dana, Wayne Gray, Megan MacGarvie, Anita McGahan, Alex Bryson, Iain Cockburn, and Christian Ketels contributed helpful suggestions. We are particularly grateful for many insightful comments by participants in the NBER Productivity Lunch Series, the Research Seminar at IESE Business School, the NBER Summer Institute (IPE group), and the North American Regional Science Association International meetings. Mercedes Delgado gratefully acknowledges support from the Kauffman Foundation and the Innovation Policy group at the NBER. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

1. Introduction

A striking feature of the US economy is significant variation in regional economic performance. For example, using the Bureau of Economic Analysis Economic Areas (EAs) as the unit of analysis, Porter (2003) documents large cross-EA differences in employment and wage growth during the 1990s, even when one conditions on the initial level of EA employment and wages (see Figure 1). Numerous theories have been proposed to explain why some regions achieve significantly higher growth rates than others, with particular emphasis on the role of initial conditions, the potential for innovation and knowledge spillovers, and the composition of economic activity (among others, Porter, 1990; Glaeser et al., 1992; Barro and Sala-i-Martin, 1995; Venables, 1996; Henderson, 1997; Fujita, Krugman, and Venables, 1999). Policymakers and researchers have focused considerable attention on areas such as Silicon Valley which seem to have achieved strong economic performance through the presence of innovative clusters of related companies and industries (Porter, 1990, 1998; Saxenian, 1994; Swann, 1998; Bresnahan and Gambardella, 2004). However, there is a surprisingly small literature examining the empirical impact of cluster composition on regional economic performance (Glaeser et al., 1992; Henderson, et al, 1995; Feldman and Audretsch, 1999; Porter, 2003).

Any empirical investigations of regional performance must account for two central, yet potentially competing, economic forces: convergence and agglomeration. Convergence arises when the potential for growth is *declining* in the level of economic activity as a result of diminishing returns (Barro and Sala-i-Martin, 1992). While many studies of convergence focus on diminishing returns at the regional level, (Barro and Sala-i-Martin, 1995), convergence may also arise at narrower units such as the region-industry (Henderson et al., 1995; Dumais et al., 2002). In this case, the region-industry growth rate (in terms of employment or other performance dimensions) will be declining in the initial level of economic activity.

Agglomeration exerts a countervailing force on regional performance. In the presence of agglomeration economies, growth is *increasing* in the level of economic activity (Glaeser et al., 1992), which can increase inequality across regions over time (Dumais et al., 2002). Agglomeration arises from interdependencies across complementary economic activities that give rise to increasing returns. The literature has tended to contrast two potential types of agglomerating forces: localization (increasing returns to activities within a single industry) and urbanization (increasing returns to diversity at the overall regional level). Distinguishing the impact of any of these types of agglomeration effects has been hindered because of the influence of convergence on regional growth. If both convergence and agglomeration effects are present, regional economic performance growth will reflect a balancing of the two effects, making it difficult to identify either effect in isolation (Henderson et al., 1995).

This paper moves beyond this impasse by focusing on the impact of clusters – groups of related industries operating in a given location – on economic performance. Our key insight is that while convergence is likely to be most salient at the industry level (or at relatively narrow levels of industry aggregation), strong agglomeration forces operate across industries within a cluster (or across closely related clusters).

Our focus on clusters allows two other related contributions. First we are able to move beyond the traditional dichotomy between the localization of individual industries and the potential urbanization effects arising from the overall diversity of regional economic activity. Instead, building on Porter (1990, 1998, 2001), we examine the agglomeration forces arising among closely related and complementary industries. By sharing common technologies, knowledge, inputs and cluster-specific institutions, industries within a cluster benefit from complementarities. Second, our empirical methodology allows examination of the impact of agglomeration among related industries while simultaneously accounting for convergence within a given industry (or within a narrow unit of economic analysis). Our methodology exploits the fact that conditional convergence operating at narrower economic units (e.g., within a single industry) can coexist with agglomeration across related economic units (e.g., across industries within a cluster). Hence we can test our main hypothesis: after controlling for the impact of convergence, the growth rate of an industry within a region is increasing in the strength of the regional cluster environment within which that industry operates.

We investigate this idea utilizing a novel panel dataset developed by the US Cluster Mapping Project (CMP). This database, drawing on the County Business Patterns data, provides a systematic classification system for mapping clusters within the US economy. Clusters are defined as groups of industries with high levels of co-location in terms of employment. The CMP identifies 41 "traded" clusters incorporating 589 "traded" industries. Traded industries and clusters are those which concentrate in particular regions and sell products or services across regions and countries, in contrast to local industries serving primarily the local market whose employment is evenly distributed across regions.[1]

The database includes numerous attributes of cluster composition and economic performance at the region-cluster-industry level between 1990 and 2005, covering 177 mutually exclusive Economic Areas (EAs) in the contiguous United States. We explore several measures of the cluster environment surrounding each region-industry, including a measure based on the strength of each cluster, a measure of the strength within the region of related clusters, and a measure that captures the strength of similar clusters in neighboring regions. For example, motor vehicles and car bodies (SIC 3711) is one of 15 industries in the automotive cluster in which a region may have strength. The automotive cluster can also be linked to as many as 6 related clusters (such as metal manufacturing) that may be present in the region, and to automotive clusters in geographically adjacent regions (see Table A2).

Our empirical strategy simultaneously accommodates convergence and agglomeration. A major identification challenge is to account for bias from unobserved factors, such as the size of the region or policies that might be associated with certain types of regions or clusters, that may be correlated with a region's cluster composition and subsequent economic performance. While we explore several alternatives, our core specifications incorporate region and industry fixed effects in the context of a region-industry growth regression. As a result of the inclusion of these detailed fixed effects, our empirical findings will be based exclusively on variation arising from the relative initial size of a cluster within a given region. By accounting for the potential for spurious correlation between cluster size and the overall growth rate of a given region and/or industry, we are able to isolate the dynamic impact of clusters on economic performance.

Our findings provide strong support for the simultaneous yet distinct influences of convergence and cluster-based agglomeration. We find that the rate of growth of employment at the region-industry level is declining in the initial level of employment at the region-industry level, consistent with the impact of convergence at the narrowest economic unit. At the same time, the growth of region-industry employment is increasing in the size and strength of the regional cluster to which that industry belongs, the size and strength of related clusters, and the size and strength of common clusters in neighboring regions. We also find that a strong regional cluster facilitates the *creation* of new industries within that cluster.

---

[1] See Porter (2003), and the discussion in Section 4.

The same overall pattern holds at the region-cluster level. While employment growth at the region-cluster level is declining in the initial (relative) size of that regional cluster, the cluster grows faster in the presence of strong related and neighboring clusters. The magnitude and statistical significance of these findings are similar to those in the region-industry model. Finally, the employment growth of a region (outside of its strong clusters) increases with the presence of strong clusters in the region, suggesting that strong clusters enhance opportunities for job creation in *other* activities in the region.

While the primary focus of this paper is on employment growth (a particularly salient performance dimension, both in terms of theory and policy), we also find that clusters have a positive impact on other dimensions of economic performance. Specifically, we find a positive impact of the cluster strength on the growth rate of average wages (productivity related) and the number of establishments (investment related). We also find a positive influence on the growth rate of patenting, a measure of innovation. In a related paper, Delgado, Porter and Stern (2010) find evidence for the positive impact of clusters on entrepreneurship. Hence our findings suggest that the positive impact of clusters on employment growth does not come at the expense of wages, investment, or innovation, but enhances them. Clusters are positively associated with multiple dimensions of regional economic performance.

Our findings generalize and extend the striking conclusions of more qualitative studies of cluster dynamics (Porter, 1990, 1998; Swann, 1992; Bresnahan and Gambardella, 2004; Bönte, 2004; Cortright, 2006). Prior studies (such as Feldman and Audretsch (1999)) have usefully demonstrated the impact of *science-based* related industries on region-industry innovation performance, but our analysis suggests that the impact of complementarities across related industries is far more pervasive. Rather than being confined to particular types of industries or operating through particular channels (such as the university-industry linkages), our results suggest that the effect of spillovers across related economic activity is a fundamental driver of growth and job creation across a broad range of industries and regions.

These findings suggest a number of policy implications, many of which diverge from the received wisdom among some practitioners. First, effective regional policy should harness complementarities across related economic activity rather than prioritize high-wage or high-tech clusters where there is little pre-existing strength within the region. Hence policy makers should pursue policies that leverage a region's cluster strengths (Porter, 2003; Cortright, 2006; Ketels and Memedovic, 2008; Rodríguez-Clare, 2007). Second, regional economic performance depends crucially on the composition of economic activity rather than the vagaries of political boundaries. The spillovers arising from related economic activity typically span multiple jurisdictions (and even states). Policies aimed at shifting the location of activity within narrow areas will be much less effective than those which operate to harness complementarities across jurisdictions.

The remainder of the paper is organized as follows. We begin by describing the role of clusters in intra-regional and inter-regional spillovers and develop the main hypotheses. Section 3 presents the empirical framework. Section 4 explains the data, and Section 5 discusses the main empirical findings. A final section concludes.

2. Clusters and Economic Geography

The agglomeration of related economic activity is a central feature of economic geography (Marshall, 1920; Porter, 1990; Krugman, 1991; Ciccone and Hall, 1996), but its prominence and role has been a puzzle. In a given location, limitations on resources can result in diminishing returns. This can lead to convergence in economic activity (employment, income, productivity) across regions over time (Barro and Sala-i-Martin 1991, 1995; Higgins et al., 2006). However, the striking geographic concentration of related economic activity, with copious examples ranging from textiles in northern Italy to financial services in New York City, reveals the powerful role of agglomeration. Starting with Marshall (1920), economists have highlighted at least three distinct drivers of agglomeration: input-output linkages, labor market pooling, and knowledge spillovers. Though conceptually distinct, each of these mechanisms is associated with increasing returns to geographically proximate economic activity.

As suggested by Glaeser et al. (1992), the relationship between the current structure of economic activity within a region and subsequent economic performance is subtle. Differing scopes of agglomeration forces may be at work. Agglomeration may arise from the specialization of a region in a particular industry where firms share common inputs or knowledge (so called localization economies). At the other extreme agglomeration may be the result of exploiting the overall diversity of industries in an entire regional economy (so called urbanization economies).[2]

[2] The empirical findings of the papers testing for localization and urbanization economies are mixed, and depend on the unit of analysis (region, industry or plant level) and the performance indicator. Some studies find evidence that

Empirical identification of these effects has been hampered because of the role of convergence (mean reversion) on regional growth patterns. As a result, the prior empirical literature has focused on identifying the balance of these two forces. Consider the relationship between the growth of economic activity and the initial level of economic activity within a region-industry. At the industry level of analysis, both convergence and agglomeration effects may be present. Region-industries may be subject to convergence effects (i.e., the coefficient on the initial *level* of economic activity is negative), either as the result of diminishing returns or a form of mean-reversion.[3] For example, the returns to economic activity can be diminishing in the level due to cost-based competition and congestion costs. A large presence of firms in an industry relative to the size of the region can intensify local competition for inputs, dampening incentives for entry and business expansion. For example, if the price of specialized (labor or capital) inputs is increasing in the number of local firms, there could be diminishing returns as a result of congestion costs (Sorenson and Audia, 2000; Swann, 1998).[4] An alternative interpretation of a negative relationship between the growth rate of employment and the initial level of employment is mean-reversion, where a region-industry with a relatively high level of economic activity at t_0 (compared to the average size of the industry in other similar regions) will likely have a lower, stochastically determined, growth rate between t_0 and t_1 (Barro and Sala-i-Martin, 1991; Quah, 1996).

---

supports the role of industry diversity in region-industry employment growth (Glaeser et al., 1992; Combes, 2000). In contrast, other studies find support for localization economies in region-industry productivity growth (Cingano and Schivardi, 2004) and in plant-level productivity (Henderson, 2003). Notably, Frenken et al. (2007) focus on *related* (within broad sectors) and *un-related* diversity (across sectors); and find that related diversity enhances regional employment growth. For further analysis on the respective influence of regional specialization and diversity see Glaeser et al. (1992) and the review of Rosenthal and Strange's (2004), among others.

[3] We draw on the convergence concept used by the cross-sectional growth literature to study economic activity across countries, regions and regional industries (Barro and Sala-i-Martin, 1991; Henderson et al., 1995; Bostic et al., 1997; Higgins et al, 2006; Geppert and Stephan, 2008; see also the review by Magrini (2004)). Other studies examine whether the distribution of industrial activity across regions is stable or diverging over time (Dumais et al., 2002). This paper focuses on (conditional) convergence to a steady state, meaning that growth is declining in the level of economic activity conditioning on differences across economies in their underlying fundamentals.

[4] Indeed, a number of prior studies find that convergence effects at the region-industry level may be sufficiently large to compensate the localization economies that take place within industries (Henderson et al., 1995; Dumais et al., 2002; Cingano and Schivardi, 2004; Klepper, 2007). The extent of agglomeration and convergence forces may vary by firm type (new entrants versus established; see e.g., Dumais et al., 2002), by industry type (mature versus new high-tech; see Henderson et al., 1995; Duranton and Puga, 2001); and across regions (depending on the degree of labor and capital mobility and other factors affecting the diffusion of advanced technologies; see e.g., Benhabib and Spiegel, 1994). Additionally, the convergence forces resulting from crowding out of demand may be lower for firms, industries and regions that are more traded-oriented (i.e., with substantial demand outside the region). For most of the analysis, we do not explore the sources of these differences, but simply control for industry, cluster and region heterogeneity using fixed effects.

However, region-industries may also be subject to agglomeration effects. There may be externalities across firms within individual industries in learning, innovation, and spawning entrepreneurs (Audretsch, 1995; Gompers et al., 2005; Glaeser and Kerr, 2009; Delgado, Porter and Stern, 2010).

Our first hypothesis concerns the relationship between convergence and agglomeration within a narrow economic unit. The empirical relationship between regional industry specialization and the growth of employment in that industry will be ambiguous, and will depend on the precise nature of competition (cost-based or innovation-based) and the pattern of strategic interaction among firms.

While the relationship between the growth rate of employment within a region-industry and its initial level may be ambiguous, it is possible to examine the potential for agglomeration as a distinct driver by considering the role of the cluster environment that surrounds a particular region-industry. The presence of complementary economic activity – e.g., specialized suppliers, a large or advanced local customer base, producers of complementary products and services, specialized institutions – increases the pool of available inputs in a location while giving rise to externalities of various sorts. This, in turn, enhances the incentives and resources available for entrepreneurship, innovation, and firm growth (Porter, 1990, 1998; Feldman and Audretsch, 1999; Delgado, Porter and Stern, 2010). Whereas prior work has emphasized individual channels through which particularly types of complementarities are realized (e.g., examining the relationship between scientific knowledge base in a region and industry innovation performance (Feldman and Audretsch, 1999), we focus on the broader role of complementarities by considering the overall impact of clusters on the performance of industries within a cluster.

Our second hypothesis concerns the role of clusters in regional performance. After controlling for the convergence effect, the growth of employment in an industry within a region will be increasing in the strength of the regional cluster environment within which that industry operates.

While convergence effects may prevail at the narrower industry level, an industry participating in a strong cluster should grow employment faster than the same industry in a region with limited presence in the cluster. A strong cluster will enable greater agglomeration economies, including larger pools of skilled employees, specialized suppliers, related industries, sophisticated buyers, and intense local competition. Proximity of related economic activity can also reduce transaction costs, enhance knowledge transfers and the flow of information, and induce the growth of specialized local institutions such as educational programs, trade groups, and quality or certification organizations that reinforce the complementarities across related industries. Thus, a strong cluster environment surrounding a particular region-industry should enhance growth at the region-industry level through increasing efficiency, driving productivity and job creation, and increasing returns to expansion, investment, and innovation (see, e.g., Porter, 1990, 1998, 2003; Saxenian, 1994; Feldman and Audretsch, 1999; Bresnahan and Gambardella, 2004; Bönte, 2004; Delgado, 2005; Cortright, 2006). In our empirical work, we examine several different facets of the regional cluster environment, including the strength of the region in other related industries that constitute the cluster and the strength in related clusters.

Entrepreneurship and new business formation by established firms are also particularly important channels for cluster-driven agglomeration (Porter, 1998; Saxenian, 1994; Feldman, 2001; Glaeser and Kerr, 2009; Feser et al., 2008; Wennberg and Lindqvist, 2008; Delgado, Porter, and Stern, 2010). Clusters facilitate new business formation and growth by lowering the costs of entry, enhancing opportunities for innovation, and allowing firms to leverage local resources to expand businesses more rapidly.

As stated in hypothesis 2, we expect that the industry growth rate will increase with the strength of the region in other complementary industries that constitute the cluster, as well as its specialization in related clusters. To illustrate this consider the following two Economic Areas in North Carolina, Raleigh-Durham-Cary and Greenville, both highly specialized in the pharmaceutical preparations industry (SIC 2834). The Raleigh EA is highly specialized in other complementary industries within the biopharmaceutical cluster and in related clusters such as medical devices and education and knowledge creation. The Greenville EA, in contrast, has a weak regional cluster environment. As expected, we find that pharmaceutical industry employment growth over 1990-2005 was significantly higher in Raleigh-Durham-Cary compared to Greenville (53% vs. -2%).[5]

Finally, there can be spillovers between a regional industry and the presence of strong clusters in nearby regions. While regional studies have often focused on regional units in isolation from other regions, economic geography theory suggests that neighboring regions can play an important role in shaping opportunities for growth. (Fujita, Krugman and Venables,

[5] See Cortright and Mayer (2002) for an informative discussion of US biotechnology clusters.

1999; Neary, 2001; Fujita and Thisse, 2002; Baldwin et al., 2003; Henderson, 2005). Indeed, some studies find that economies of agglomeration attenuate with distance (see, e.g., Rosenthal and Strange 2003, Henderson, 2003), and others find relevant spatial interactions among cities (Dobkins and Ioannides, 2001).

In this paper, we consider the impact of the presence of similar clusters in neighboring regions on industry and cluster growth in a region. The presence of a strong cluster in a neighboring region can be a source of locational competition, particularly for constrained inputs and demand. However, industries and clusters that are co-located in nearby regions may benefit from inter-regional spillovers, which lower the costs and risk of entrepreneurship and business expansion (e.g., by providing access to suppliers and customers, by allowing firms to leverage proximate inputs, technology, institutions, etc). There is likely to be asymmetry in the type and the extent of inter-regional spillovers among neighboring clusters depending on the depth and breadth of clusters in a region. For example, inter-regional spillovers may be lower for large versus small regions, and lower for leading national clusters versus smaller clusters.[6] *Our final hypothesis is that the impact of cluster strength in neighboring regions will have an ambiguous effect on the growth of employment in regional industries and clusters, depending on the relative salience of inter-regional spillovers versus locational competition.*

While our discussion has focused on employment growth, agglomeration effects should also affect other facets of economic performance, such as wages, innovation (e.g., patenting rates), and entrepreneurship (e.g., new business starts). In our empirical analysis, we examine these additional dimensions of performance to clarify the role of clusters in overall regional economic performance. In particular, we explore whether a strong cluster environment results in better overall region-industry performance, or whether greater employment growth comes at the expense of lower wages, lower entrepreneurship, or reduced innovation. However, since the theoretical relationship between clusters, productivity (wages), innovation and employment is subtle (e.g., see Bostic, Gans and Stern, 1997; Cingano and Schivardi, 2004), a full and simultaneous assessment of the role of clusters on multiple dimensions of economic performance is left to a companion paper.

[6] The co-location of a similar cluster in nearby regions may be driven by several mechanisms, such as input-output linkages; human capital composition; and the contribution of a leading cluster to the development of nearby clusters. We abstract from identifying the mechanism that generates the inter-regional spillovers, instead focusing on the impact of neighboring regions' cluster composition on the economic performance of individual industries and clusters.

In general, our framework suggests that the impact of convergence will be more salient for narrower economic units, and cluster composition may be a key driver of regional agglomeration. While convergence effects in industry employment growth may occur at the industry level, the most powerful agglomeration forces arise at a more aggregated level: the cluster, among related clusters, and among similar clusters in nearby regions. We can also use our framework to evaluate a number of related, ancillary hypotheses. We can test whether the employment growth rate of a cluster is increasing in the strength of related clusters in the region (i.e., clusters which are likely complementary to each other) or the strength of similar clusters in geographically adjacent regions. This type of analysis is valuable to test whether our core insight concerning the impact of complementary economic activity is robust across different levels of data aggregation. Additionally, we can examine whether the presence of a strong cluster environment spurs regional growth more generally. Specifically, we evaluate how the employment growth rate of a region (outside the strong clusters in that region) depends on the presence of strong clusters within a given region.[7]

3. Model

Industry Growth Model. To examine region-industry growth we draw on studies of conditional convergence (e.g., Barro and Sala-i-Martin 1991, 1995; Combes, 2000; Dumais et al., 2002; Higgins et al., 2006) that examine economic growth as function of the level of economic activity and observable attributes of the region. While convergence forces may prevail at the region-industry level, we argue that the important agglomeration forces are due to the presence of clusters of related industries (Porter, 1990, 1998; Swann, 1998; Feldman and Audretsch, 1999; Paci and Usai, 2000; among others).

We test these ideas utilizing a dataset that examines region-industry growth between 1990 and 2005 for 177 regions and up to 589 traded industries, totaling 55083 region-industries in which we observe positive employment in 1990. To separate convergence and agglomeration forces in regional industry growth, we distinguish between the level of regional specialization in a particular industry and the strength of the cluster environment around that region-industry. Once we control for the average growth of the industry and the region, and after conditioning out

[7] Porter (2003) suggests that regional prosperity may be driven by the relative performance of the strong clusters in the region. The ability of a region to perform well in whichever clusters with meaningful position seems more important for regional economic performance than the region's efforts to specialize in nationally high-wage clusters.

the effect of the (relative) size of the industry, we are able to disentangle the effect of the cluster environment on region-industry growth. Following the conditional convergence literature, our core econometric specification is:

$$\ln\left(\frac{\text{Employ}_{\text{icr},2005}}{\text{Employ}_{\text{icr},1990}}\right) = \alpha_0 + \delta \ln(\text{Industry Spec}_{\text{icr},1990})$$
$$+ \beta_1 \ln(\text{Cluster Spec}^{\text{outside i}}_{\text{icr},1990}) + \beta_2 \ln(\text{Related Clusters Spec}^{\text{outside c}}_{\text{cr},1990}) \qquad (1)$$
$$+ \beta_3 \ln(\text{Cluster Spec in Neighbors}_{\text{cr},1990}) + \alpha_{\text{i}} + \alpha_{\text{r}} + \varepsilon_{\text{icr}}.$$

Our primary dependent variable is employment growth of the (4-digit SIC) industry *i* in cluster *c* in region (EA) *r* over the period 1990-2005; and the explanatory variables are specified at 1990.[8] To capture the balance of the convergence and agglomeration forces at the industry level, we specify a model that includes regional specialization in the industry (Industry Spec). We also include several measures of the relative strength of the cluster environment surrounding that region-industry (see Section 4 for a precise definition of these measures): Cluster Spec (a measure of cluster strength in the set of closely related industries constituting the cluster), Related Clusters Spec (a measure of the strength of related clusters in the region), and Cluster Spec in Neighbors (a measure of cluster strength in adjacent geographic regions). We include industry (α_i) and region fixed effects (α_r) to control for other differences across regions and industries that affect employment growth. Thus, this specification examines the impact of the level of industry specialization and the strength of the cluster environment, fully controlling for differences in the average growth of a region or the average growth of a particular industry. Our analysis thus accounts for unobserved factors (such as the size of the industry and region, regional policies, etc) that might be correlated both with our explanatory variables and region-industry employment growth. In other words, we isolate the impact of clusters on economic performance by conditioning out the potential sources of spurious correlation between initial cluster size and the overall growth rate of a given region and/or industry.

The key coefficients to test our main hypotheses are δ and $\boldsymbol{\beta}$. In line with our first hypothesis, our prediction concerning the coefficient on the initial (relative) size of an industry in

---

[8] To illustrate the unit of observation, consider the pharmaceutical preparations industry (SIC-2834) in the biopharmaceutical cluster in the Raleigh-Durham-Cary (NC) region. For this region-industry observation we look at the specialization of the region in the industry; the specialization of the region in biopharmaceuticals (excluding industry SIC-2834); the presence in the region of other related clusters (such as medical devices, chemical products, and education and knowledge creation); and the surrounding regions' specialization in biopharmaceuticals (including the focal industry).

the region (Industry Spec) is ambiguous, depending on the relative salience of convergence and agglomeration forces at the region-industry level. For the role of the cluster environment, our framework offers sharper predictions. After controlling for the convergence effect, we expect industries co-located with industries in the same cluster, or in regions with strong related clusters, to perform better than industries in regions lacking cluster strength ($\beta_1 > 0$ and $\beta_2 > 0$).

Importantly, we test whether this relationship is driven by industry complementarities within clusters or simply results from random aggregation of industries. Specifically, we implement a Monte Carlo falsification test in which, for each round of the Monte Carlo, we construct "random" clusters by randomly assigning the 589 traded industries into 41 clusters (see Section 5 for a more detailed description of our procedure). By constructing "random" clusters, we are able to evaluate whether our results are simply an artifact arising from the inclusion of *any* sets of industries or whether they depend on clusters constructed based on complementarities.

To account for correlation of the error terms across industries within a regional cluster, the standard errors are clustered by region-cluster. Finally, since nearby regions tend to specialize in like clusters, there might be unobserved spatial autocorrelation (Anselin, 1988). We account for this in two ways, by including in our main specifications the strength of similar clusters in neighboring regions and region fixed effects, and by directly testing for spatial correlation using spatial econometric techniques.[9]

Cluster Growth Model. To examine the relationship between a cluster's employment growth and the regional cluster environment, we specify the following model:

$$\begin{aligned} \ln\left(\frac{\text{Employ}_{\text{cr},2005}}{\text{Employ}_{\text{cr},1990}}\right) &= \alpha_0 + \delta \ln(\text{Cluster Spec}_{\text{cr},1990}) \\ &+ \beta_1 \ln(\text{Related Clusters Spec}^{\text{Outside c}}_{\text{cr},1990}) + \beta_2 \ln(\text{Cluster Spec in Neighbors}_{\text{cr},1990}) \\ &+ \alpha_c + \alpha_r + \varepsilon_{\text{cr}}. \end{aligned} \qquad (2)$$

The dependent variable is the employment growth of cluster *c* in region *r* over the period 1990-2005, with the explanatory variables specified in 1990. We expect that the region-cluster growth

[9] Following Anselin (1988) and the extensions and implementation developed by LeSage (1999, p. 171), we test for spatial autocorrelation using a first-order spatial autoregressive (FAR) model. We first implement a fixed effects OLS specification (e.g., equation (1)) and then estimate the following FAR model: $\hat{\varepsilon}_{\text{icr}} = \rho \text{W} \hat{\varepsilon}_{\text{icr}} + \mu_{\text{icr}}$, where $\hat{\varepsilon}$ are the residuals from the OLS estimation, and W is an NxN matrix (where N is the total number of region-industries) with elements equal to 1 for adjacent regions and 0 otherwise. Under the null hypothesis of no spatial autocorrelation if ρ=0. LeSage and collaborators generously provide MATLAB software for the FAR test, available at http://www.spatial-econometrics.com.

rate will be ambiguous in the initial size of the cluster (Cluster Spec), depending on the relative impact of convergence and agglomeration. Conditioning on cluster specialization, we also test for agglomeration forces involving related clusters (Related Clusters Spec) and similar clusters in neighboring regions (Cluster Spec in Neighbors).

Region Growth Model. Finally, we specify a region-level model to test whether the set of strong traded clusters in a region (traded clusters highly over-represented in the region) contribute to the employment growth of other clusters in the region. To test this relationship, we specify the following model:

$$\ln\left(\frac{\text{Employ}_{r,2005}}{\text{Employ}_{r,1990}}\right)^{\substack{\text{Outside}\\ \text{strong clusters}}} = \alpha_0 + \delta \ln(\text{Reg Employ}^{\substack{\text{Outside}\\ \text{strong clusters}}}_{r,1990}) + \beta \text{Reg Cluster Strength}_{r,1990} \quad (3)$$
$$+ \lambda \text{National Employ Growth}^{\text{Strong clusters}}_{r,1990-05} + \alpha_{\text{Census Region}} + \varepsilon_r.$$

The dependent variable is the growth in regional employment *outside its strong clusters*, over the period 1990-2005. Convergence forces that could prevail at the regional level are captured by including the level of regional employment outside its strong clusters. Cluster agglomeration forces are measured by the presence of the strong clusters in the region (Reg Cluster Strength). We expect that the set of strong clusters (and their linkages) in the region will contribute to growth of other activities in the region ($\beta > 0$). The model controls for national shocks affecting the region's strong clusters (National Employment Growth) because these shocks might be correlated to subsequent technological and demand shocks in the region, affecting regional performance growth and cluster composition. Finally, the model also controls for broader regional differences by including six broad Census-region dummies (e.g., Northeast, South Atlantic, etc.).

4. Data

Data from the County Business Patterns (CBP) dataset is coded with cluster definitions drawn from the US Cluster Mapping Project (CMP; see Porter, 2001, 2003). Before turning to the precise variable definitions, it is useful to provide an overview of the data sources. The CPB dataset is a publicly available database that provides annual county-level measures of private sector non-agricultural employment and establishments at the level of four-digit SIC codes (which we refer to as industries). This data is aggregated 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 defined by the Bureau of Economic Analysis) as the geographic unit. The analysis focuses on the 1990-2005 period.[10]

The cluster definitions and relationships among clusters are drawn from the US Cluster Mapping Project (Porter, 2001, 2003). Directly measuring complementarity in economic activity in a consistent and unbiased manner represents a considerable challenge.[11] The CMP used an indirect methodology for grouping four-digit (and some three-digit) SIC codes into clusters and related clusters.[12] The methodology first distinguishes three types of industries with different patterns of competition and locational drivers: traded, local, and natural resource-dependent. Local industries are those that serve primarily the local markets (e.g., utilities) whose employment is roughly evenly distributed across regions. Resource-dependent industries are those whose location is tied to local resource availability (e.g., logging). Traded industries are those that tend to be geographically concentrated and produce goods and services that are sold across regions and countries.[13]

Traded industries are the focus of our analysis. Using the pairwise correlations of industry employment across locations, the CMP assigns each of the 589 traded industries to one of 41 mutually exclusive traded clusters (referred to as "narrow cluster" definitions). This empirical approach is based on revealed co-location patterns, and captures any type of externalities present across industries (e.g. technology, skills, demand, or others) rather than

---

[10] There are 179 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).

[11] 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 et al., 2005). Finally, Ellison, et al, (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), Forni and Paba (2002) and Alcacer and Chung (2010).

[12] In order to use EA-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 a significant way in 2003). By construction, recent NAICS-based data can be translated (with some noise) into the older SIC system. The use of NAICS or SIC definitions has no meaningful impact on our core empirical results.

[13] Traded industries account for 30% of US employment and over 87% of US patents. See Porter (2003) for additional detail on the methodology to define these 3 industry groups.

focusing on any one type.[14] Examples of clusters include automotive, apparel, biopharmaceuticals, and information technology. Within a cluster such as information technology, for example, there are 9 4-digit SIC code industries, including electronic computers (SIC 3571) and software (SIC 7372), reflecting the fact that location of employment in computer hardware and software are highly correlated (Appendix A provides a comprehensive list of the 41 traded clusters and some key attributes).

The CMP also develops "broad" cluster definitions in which a given industry may be associated with multiple clusters (as inferred through the locational correlation of employment patterns). While the narrow cluster definitions form our key measures of complementary industries, we use the broader clusters to identify related clusters and to develop a measure of the strength of related clusters surrounding a given region-cluster.

Other Sources of Data. In addition to employment growth in a given EA-industry, we also examine the growth in the number of establishments, average wages and patenting. The establishment data is drawn from the CBP dataset. Wage indicators are obtained from the Longitudinal Business Database (LBD) of the Census Bureau, which provides annual observations of the universe of US establishments with payroll.[15] The EA-industry patent data is drawn from the United States Patent and Trademark Office (USPTO). This dataset includes detailed information about all USPTO utility patents, including inventor location and technology classification. Constructing patenting measures is complicated, both because USPTO patents are assigned to patent classes but are not directly matched to SIC codes, and because the multiple inventors of a given patent may be located in different regions. We utilize a patent-SIC code concordance algorithm developed by Silverman (1999), in which USPTO patents are assigned,

---

[14] Porter (2003) describes the methodology and classification system in detail. The primary classifications are based on empirical patterns of employment co-location among industries. Using this method, clusters may contain service and manufacturing industries as well as industries from different parts of the SIC system. It is possible that industries with high co-location may have little economic relationship. Two adjustments are made in the CMP to the cluster definitions to eliminate spurious correlations. First, the SIC industry definitions and list of included products and services are used to reveal logical links. Second, the National Input-Output accounts are used to identify meaningful cross-industry flows.

[15] For detailed information on the LBD data see Jarmin and Miranda (2002) and Delgado, Porter and Stern (2010). For the wage analysis, we omit establishments with zero employment and with very low wages (below half of the minimum wage) or very high wages (above $2 million USD) to avoid part time employment and potential measurement error.

on a fractional basis, to four-digit SIC codes in a consistent (albeit somewhat noisy) manner.[16] Each patent is then assigned, on a fractional basis, among the locations of the inventors.

Sample Description and Dependent Variable Definitions

Our main measure of economic performance is employment growth in a given EA-industry over the period 1990-2005. We also examine other dimensions of economic performance, such as the growth in the number of establishments in the EA-industry, where an increase in the growth rate would reflect an increase in entry of new establishments, the rate of survival, or both; the growth in average wages; and the growth in patenting (see Table 1 and Appendix B, Table B2).

The most straightforward approach to evaluating growth, taking $\ln(Employ_{i,r,2005}/Employ_{i,r,1990})$, must account for the fact that there are many EA-industries in which there is a zero level of employment. We need to either exclude those observations or impose a positive lower bound on the level of employment.[17] Since the empirical analysis focuses on explaining the growth of existing regional industries, the core sample uses EA-industries that have a positive level of employment in 1990 (55083 observations). To include EA-industries where we observe subsequent zero employment in 2005 (6162 observations), we set employment equal to one for these observations.

We also examine the *creation* of new EA-industries, by using the sample of EA-industries with zero employment in 1990 (48430 observations). The dependent variable is a dummy variable equal to one if there is a positive level of employment as of 2005 (Any Positive Employment_{2005} variable; with a mean of 0.29 and a standard deviation of 0.45). These analyses include probit specifications of the probability of having industries new to the region as of 2005 as a function of the initial cluster environment.

---

[16] The Silverman algorithm cleverly exploits the simultaneous assignment of patent classes and industry codes in the Canadian patent system. It (a) links each USPTO patent to Canadian patent classes (using each patent's IPC classes), (b) links each Canadian patent class to the Canadian SIC scheme according to the historical propensity for patents from each IPC class to be associated with particular Canadian SIC codes, and (c) links the Canadian SIC and US SIC classification systems according to a separate concordance developed by Industry Canada and the US Department of Commerce. For a detailed exposition of this procedure, see Silverman (1999).

[17] Since we are using narrow regional units and individual industries (mainly four-digit SIC), there are numerous regional industries with zero employment. Most zeros concentrate in small EAs and small national industries and clusters (e.g., Tobacco and Footwear related industries).

Finally, we implement numerous sensitivity analyses involving alternative approaches to deal with the zeros. We use the sample of regional industries that have positive employment in both the end and the terminal period (i.e., "non-zeros" sample of 48921 observations); and we also vary the base and terminal periods (e.g. 1990-1996 and 1997-2005 versus 1990-2005). Our main findings remain essentially unchanged.

Explanatory Variables: Industry and Cluster-Level Models

To examine the impact of the industry strength and different aspects of cluster strength on the growth of regional industries, we need measures of industry and cluster specialization as well as measures of 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). Employment-based industry specialization in the base year (1990) is measured by the share of regional employment in the regional industry as compared to the share of US employment in the national industry: $\text{INDUSTRY SPEC}_{\text{Employ,i,r,90}} = \dfrac{\text{employ}_{\text{i,r}} / \text{employ}_{\text{r}}}{\text{employ}_{\text{i,US}} / \text{employ}_{\text{US}}}$, where *r* and *i* indicate the region (EA) and the industry, respectively. This indicator captures the degree to which the industry is "over-represented" (in terms of employment) in the EA. In our sample, the industry specialization of regions has a mean of 2.34 and a standard deviation of 7.21 (Table 1). As noted earlier, we include EA and industry fixed effects in the regional industry model (equation 1), and so, the independent variation in our main empirical specifications is driven exclusively by variation in employment in the EA-industry.

Cluster Specialization. We use 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 regional cluster (*outside the industry*) as compared to the share of US employment in the national cluster (*outside the industry*): $\text{CLUSTER SPEC}_{\text{Employ,icr,90}} = \dfrac{\text{employ}_{\text{c,r}}^{\text{outside i}} / \text{employ}_{\text{r}}}{\text{employ}_{\text{c,US}}^{\text{outside i}} / \text{employ}_{\text{US}}}$. In our sample, the average employment-based Cluster Spec is 1.31 (and the standard deviation is 2.20; Table 1). Analogously, in the EA-cluster level models, cluster specialization is measured using an employment-based location quotient, but including all the industries that constitute a given cluster: $\text{CLUSTER SPEC}_{\text{Employ,cr,90}} = \dfrac{\text{employ}_{\text{c,r}} / \text{employ}_{\text{r}}}{\text{employ}_{\text{c,US}} / \text{employ}_{\text{US}}}$.

Since cluster specialization is measured relative to the overall size of the region, a region may exhibit specialization within a particular cluster even though that region only holds 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 pockets of automotive cluster strength in smaller regions, such as Lexington-Fayette-Frankfort-Richmond (KY) and Louisville-Elizabethtown-Scottsburg (KY-IN) (see Table A3 and Figure A1.1 in the Appendix). Thus, our analysis includes both large and small regional clusters.

Finally, it is useful to note that, in the EA-industry models the inclusion of region and industry fixed effects, means that the independent variation utilized in the regressions comes exclusively from the employment within a given cluster (outside the industry). Similarly, in the EA-cluster models (with region and cluster fixed effects), the variation in cluster specialization comes from differences across clusters in the initial level of employment.

Strong Clusters in a Region. To identify the set of strong clusters in a given EA in 1990 we use the magnitude of cluster specialization, based on the distribution of Cluster Spec_{cr} across EAs for each of the 41 clusters. We then define as strong clusters those in EAs with the top 20% of specialization (i.e., above the 80^{th} percentile value of Cluster Spec_{cr}; see Table A3). Even after controlling for cluster differences, there are small regions with very low employment that manage to hit the location quotient threshold. To address these cases, the high cluster specialization criterion is supplemented with a minimum share of national cluster employment and number of establishments.[18] The strong clusters in an EA are those that satisfy these three criteria (REG STRONG CLUSTERS). On average, EAs have 6.91 strong clusters (Table 1).[19]

Strength of Related Clusters. The measures of the strength of related clusters are developed using the "broad" cluster definitions in Porter (2003). We identify as clusters related to a given cluster *c* those broad clusters that have at least one of cluster *c*'s narrow industries in

---

[18] The threshold values are selected by cluster, using the employment and establishment values that correspond to the 20^{th} percentile.

[19] Using these criteria, all the EAs prove to have at least one strong cluster. Some regions have numerous strong clusters (e.g., San Diego-Carlsbad-San Marcos, CA) while others have only one or two (e.g., Lewiston, ID-WA).

common. For example, in the case of automotive, related clusters include production technology and metal manufacturing, among others (see Table A2). [20] Having identified the set of clusters related to a given cluster (C^*), we then measure the degree of overlap between each pair of clusters (c, j) using the average proportion of narrow industries that are shared in each direction: $\omega_{c,j} = \text{Avg}\left(\frac{\text{shared industries}_{c,j}}{\text{total industries}_c}, \frac{\text{shared industries}_{j,c}}{\text{total industries}_j}\right)$. [21] The strength of a region in clusters related to cluster c is then defined by a weighted sum of the location quotients associated with each (narrowly defined) related cluster:

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

For instance, based on this weighting which emphasizes the degree of overlap between clusters, our measure of the strength of related clusters for industries within the automotive cluster weights the presence of the metal manufacturing cluster more heavily than the presence of the furniture cluster (Table A2).

Strength in Neighboring Clusters. Specialization in a particular cluster tends to be spatially correlated across neighboring regions – for example, 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) (see Figure A1.1).[22] We develop a measure of the presence of similar clusters in neighboring regions (CLUSTER SPEC in NEIGHBORS variable) to explore the role of neighboring clusters in employment growth in a region-industry. We compute the (average) specialization of adjacent regions in the cluster (including the focal

---

[20] The related clusters selected using this method will be the most relevant ones for each given cluster *c*. This concept of related clusters is conservative since we count industry linkages in only one direction (i.e., we do not look at how the broad set of industries in cluster c are shared with other narrow clusters). Clusters related with many other traded clusters include analytical instruments and communications equipment, among others; and clusters with few connections to other clusters include tobacco and footwear.

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

[22] Service-oriented clusters, such as Financial Services, also tend to co-locate in nearby regions (see Fig. A1.2). More generally, Table B1 shows that the specialization of a region in a given cluster is significantly correlated to the average specialization of neighbors in the same cluster (correlation coefficient of 0.50).

industry). In other words, the strength of neighboring clusters is measured by the average (employment-based) LQ of a cluster in adjacent regions.

Finally, as mentioned earlier, in the EA-industry (EA-cluster) models we include a complete set of EA and industry (cluster) fixed effects in our main specifications to control for unobserved factors that may be correlated with industry and cluster specialization.

Explanatory Variables: Region-level Model

In the region growth model specified in equation 3, the measure of cluster-driven agglomeration is regional cluster strength (REG CLUSTER STRENGTH), defined as the share of regional traded employment accounted for the set of strong clusters in the region:

$$\text{REG CLUSTER STRENGTH}_{\text{Employ}} = \frac{\sum\limits_{c \in \text{Strong Clusters}_r} \text{employ}_{c,r}}{\text{Traded regional employ}_r}.$$

As mentioned earlier, the strong clusters in a region are those with high cluster specialization (relative to like clusters in other regions). Thus, the regional cluster strength variable captures benefits of having an array of clusters highly over-represented in the region. Regional cluster strength will be higher if there are a few strong clusters that account for much of the regional traded employment (e.g., automotive related clusters in Detroit-Warren-Flint, MI), or if there are numerous strong clusters (e.g., South Bend-Mishawaka, IN-MI).

Since a region with strength in clusters that are rapidly growing nationally (e.g., some areas of high-tech and services) is likely to experience distinct technological and demand shocks, we include the average national employment growth outside the region of a region's strong clusters as a control (NATIONAL EMPLOY GROWTH of STRONG CLUSTERS, Table 1).

Finally, in the region-level model we cannot include EA fixed effects since we are using a cross section. Instead, we use six broad Census regions (West Pacific, West Mountain, Midwest, South Central, Northeast, and South Atlantic) to control for differences in growth rates and cluster structure across broader regions of the US.[23]

[23] For example, the Northeast area is specialized in clusters that have been growing nationally, such as business services, education and knowledge creation, and medical devices. In contrast, strong clusters in the South Atlantic region include textiles, apparel, and furniture.

5. Results

We now turn to the key empirical findings. Table 2 provides a simple analysis of how the growth rate of EA-industry employment varies with the initial level of industry specialization and cluster specialization (i.e., initial industry and cluster strength in an EA). The 55083 EA-industries are divided into four categories based on whether they have low or high (below or above the median) employment-based industry and cluster specialization for their industry in 1990. For each of these four groups of EA-industries, we compute the average annualized percentage employment change between 1990 and 2005. We find that there is a significant decrease in the average employment growth rate when one moves from an EA-industry with *low* initial industry specialization to one with *high* industry specialization, consistent with the convergence effect (e.g., for EA-industries with low initial cluster specialization, the annual growth rate decreases from 13% to 0%). Regardless of the initial level of industry specialization, however, there is a significant increase in the growth of employment when one moves from an EA-industry with a *low* cluster specialization to one with *high* cluster specialization (e.g., for EA-industries with low initial industry specialization, the annual growth rate increases from 13% to 20%). In other words, regional industries that are located in relatively strong clusters experience much higher growth rates in employment.

A more systematic regression analysis is shown in Table 3. We begin in (3-1) with a simple model relating EA-industry employment growth to the level of industry specialization. The estimated coefficient is negative, suggesting that convergence dominates the impact of agglomeration at the narrowest unit of analysis. In (3-2), we introduce a second variable, Cluster Spec. While the coefficient on Industry Spec continues to be negative (and roughly of the same magnitude), the coefficient on Cluster Spec is positive.[24]

There could be alternative factors, such as region or industry effects, driving our findings. Therefore, in (3-3), we introduce a comprehensive set of industry and EA fixed effects. This is our core specification, relying exclusively on variation in the (relative) strength of the cluster environment, conditioning on the overall growth rate for the region and industry. The results are

[24] In our sample of EA-industries with positive employment in 1990, there are observations with zero employment in the cluster (outside the industry), in related clusters or in neighboring clusters. To be able to use these observations, we replace ln(variable) with the minimum value of ln(variable) in the sample. To control for unobserved attributes of these types of observations we also add across all models (Tables 3-to-5; and models 6-3 to-6-6) a dummy equal to 1 if the particular variable was not corrected. For example, model (3-3) includes a dummy equal to 1 if an EA-industry has any employment in 1990 in the cluster (outside the industry). All our findings across Tables 3-6 only change trivially when dropping these dummies.

essentially unchanged: there is a large negative impact of industry specialization, and a large positive impact of cluster specialization. For example, using the coefficients from (3-3), the annual rate of convergence is 3.5%.[25] At the same time, a one standard-deviation increase in cluster specialization (2.20) is associated with a 2.9% increase in the expected *annual* employment growth rate (relative to the average annual growth rate of -2%). In other words, the suggestive evidence from Table 2 is not simply the result of spurious correlation or an artifact of particular high-growth industries or regions; instead, our findings reflect a systematic and precise relationship between the cluster environment and the potential for growth at the region-industry level.

Table 3 also includes a number of robustness checks. We re-estimate the specification (3-3) using a sample of *all* EA-industries, including both traded and local clusters (3-4). The estimates are essentially unchanged, suggesting that cluster-driven agglomerations benefit both local and traded industries. We then turn in (3-5) to a Monte Carlo falsification test in which we construct "random" clusters by randomly assigning the 589 traded industries into 41 traded clusters (without replacement). We repeat this process 2000 times to create 2000 simulated cluster definitions, and select the 200 cluster definitions that have the closest distribution to our original cluster definitions (where closest is defined in terms of the distribution of size of region-clusters). For each of the 200 random cluster definitions, we estimate the exact same specification as (3-3); (3-5) reports the average coefficient from this exercise (with the standard errors based on the empirical distribution from the simulation exercise). While the impact of Industry Spec continues to be negative (and of the same magnitude), the coefficient associated with the *random* cluster specialization measures are essentially zero and statistically insignificant. In other words, cluster-driven agglomeration depends critically on grouping industries into meaningful groups of related activities (consistent with Porter (2003)), and does *not* simply reflect random industry groupings.

Finally, we examine an alternative measure of complementary economic activity drawn solely from the SIC system itself. In (3-6), we define related industries as all four-digit SIC code (traded) industries within a two-digit SIC code (i.e., industries that are related based on their list of products/services; variable SIC2 Spec). This mechanical use of SIC codes captures some forms of relatedness, but fails to capture complementarities among service and manufacturing

[25] The annual convergence rate is 100*(ln(-0.395+1))/-15) following Barro and Sala i Martin (1991).

industries or among four-digit industries not included in the same two-digit SIC code. Despite these limitations, the relationship between SIC2 Spec and EA-industry growth rate is also positive and significant, consistent with our earlier findings.[26]

In (3-7) and (3-8) we examine the impact of strength in related clusters and the presence of strong clusters in neighboring regions. To do so, we introduce two new measures: Related Cluster Spec and Cluster Spec in Neighbors. The only difference between (3-7) and (3-8) is the exclusion or inclusion of industry and region fixed effects. Across both specifications, the estimated impact of industry specialization continues to be negative (consistent with a convergence effect), but we find separate and positive impacts for each of the three measures of the cluster environment surrounding a particular region-industry.[27] Moreover, the estimated impact is large: using the estimates from (3-8), a one standard-deviation increase in each of the aspects of the cluster environment is associated with a 3.7% (2.2%, 0.6% and 0.9%, respectively) increase in the expected *annual* employment growth rate.

Creation of New Industries. A significant limitation of the analysis to this point is that the sample excludes those region-industries with *zero* employment in the initial period. In other words, the analysis has examined the impact of the cluster environment on the growth rate of *existing* region-industries. We also examine the impact of the initial cluster environment on the creation of *new* EA-industries in Table 4. The sample is composed of EA-industries with zero employment as of 1990. The dependent variable in the probit equations in Table 4 is a dummy variable equal to one if there is a positive level of employment as of 2005. The results are striking. A higher level of Cluster Spec, Related Cluster Spec and Cluster Spec in Neighbors is associated with a higher probability of new industry creation. This result holds with the inclusion or exclusion of region and industry fixed effects, and the estimated magnitudes are

---

[26] The findings in Tables 3 are robust to a number of additional sensitivity checks, such as using the non-zero sample (i.e., EA-industries with positive employment in both 1990 and 2005), using alternative regional units (e.g., MSAs), dropping the 5% smallest and largest EAs, and dropping small regional clusters that may consist mostly of sales and distribution offices of clusters based elsewhere (i.e., for every cluster, we drop those regional clusters in the bottom 20% of the distribution of employment). As described in Section 3, we also test for spatial correlation running a first order spatial autoregressive (FAR) model on the residuals of our core fixed effects models (3-3 and 3-8), under the null hypothesis of no spatial autocorrelation if $\rho=0$. We cannot reject the null hypothesis. In both specifications, the R-squared of the FAR models are zero, the coefficients of ρ are around -.004, and the standard deviation of ρ is large (1.77).

[27] Importantly, the inter-regional spillover effects are robust to using alternative weighting schemes for adjacent regions (e.g., weighting more the neighboring clusters with higher levels of employment versus using the un-weighted average). Furthermore, the inter-regional spillovers are higher when neighbors specialize in the same cluster than when neighbors specialize in related clusters.

large. For example, in (4-3), a one-standard deviation increase in Cluster Spec is associated with an ~ 5 percentage point increase in the probability of new industry development in a region (relative to a mean of 0.3 for the dependent variable). Overall, these findings suggest that cluster strength facilitates the creation of related new industries in a region, hence playing an important role in the diversification of regional economies.

Other Performance Measures. Our analysis thus far has focused on employment growth. In Table 5 we evaluate additional performance dimensions to further clarify the role of clusters in regional performance. Here we examine the influence of industry and cluster specialization on the growth rate of average wages, the number of business establishments, and patenting in regional traded industries. For each of these dimensions of performance we estimate fixed effect models similar to (3-3) and (3-8), but with the dependent variable and explanatory variables constructed using measures specific to that performance dimension (see Table B2 for the descriptive statistics).

In the wage growth model (5-1), the explanatory variables are the 1990 level of average wages within the EA-industry and the 1990 level of average wages within the cluster to which that EA-industry belongs. The results are striking. The growth rate of wages is declining significantly in the initial level of EA-industry average wages, but is simultaneously increasing in the average wage of the cluster (excluding the EA-industry). In (5-2) we incorporate additional dimensions of the cluster environment and find that wage growth is also positively associated with the average wage of clusters in neighboring regions. These large and robust effects are present despite the inclusion of industry and EA fixed effects. In other words, these effects are not simply reflecting the fact that wage growth in certain types of industries and regions was more rapid than others. Instead, these results suggest a systematic and positive relationship between the growth rate in region-industry wages and the initial cluster environment.

We find a similar pattern for the growth in the number of establishments and in patenting in traded industries (models 5-3 to 5-6). In these growth models the explanatory variables are the initial level of industry specialization (based on number of establishments or patents) and various aspects of the cluster environment (based on number of establishments or patents) of a given EA-industry. We find that high initial levels of industry specialization within the EA-industry is associated with a reduction in the growth rate in the number of establishments and patents, while an initially stronger surrounding cluster environment is associated with an increase in the growth rate of these variables.[28]

These findings suggest that clusters matter not only for job creation but for wages, new business establishments and innovation, providing strong evidence of the positive role of clusters in regional performance. Note, however, that the theoretical arguments for the role of clusters in productivity and innovation growth may be different from the ones that explain employment growth (e.g., see Bostic, Gans and Stern, 1997; Cingano and Schivardi, 2004). Further research is needed on the complex relationship between clusters, productivity, innovation and employment. A simultaneous assessment of the role of clusters on multiple dimensions of economic performance will be an important subject for future research.

When Do Clusters Matter More? While our findings highlight the strong average impact of the local cluster environment on regional economic performance, prior theoretical work suggests that the impact of agglomeration forces can vary across types of clusters and regions. For example, while our cluster measures are constructed in terms of the *relative* size of the cluster (relative to the region and the national cluster), some of the externalities and complementarities harnessed by strong clusters may be subject to economies of scale and scope. Thus, the absolute size of the local region can matter.

We investigate the relationship between cluster driven agglomeration and region size in (6-1), introducing a set of interaction effects between industry and cluster specialization and a dummy variable (EA SIZE) which is equal to one for regions above the median employment size.

The main findings hold for both large and small regions, though there are important differences. Notably, the impact of cluster composition on performance is greater in larger regions (for both Cluster Spec and Related Cluster Spec). Also, and perhaps not surprisingly, the impact of neighboring clusters is significantly greater for industries and clusters in smaller regions.[29] A greater importance of spillovers with neighbors for smaller regions is consistent with the idea that larger EAs enhance the opportunities for cluster development in adjacent

[28] We need to be cautious about the patenting analysis because the patents generated in an EA-industry in a given year is a flow variable and may not necessarily capture the innovative activity of that specific year due to the delay between patent application and patent granted.
[29] The findings are robust to alternative samples and specifications.

regions, as some cluster activities are outsourced nearby and the large region provides access to demand and skills.

We investigate the degree of heterogeneity across regions and clusters more systematically by estimating a separate industry and cluster specialization coefficient for each of the 177 EAs and for each of the 41 clusters. Taking advantage of the fact that each industry (and cluster) is present in multiple locations, we are able to estimate a model similar to (3-3) but allow for cluster-specific coefficients. The results are presented in Figure 2A and Table 7. In Figure 2A, we simply plot the empirical distribution of these coefficients. The cluster-specific industry specialization coefficient is always negative, while the cluster-specific cluster specialization coefficient is almost always positive. Moreover, both of these distributions are single-peaked and relatively tight. This finding is reinforced when we examine the coefficients by clusters in detail in Table 7. Except for two cases (distribution services and entertainment), the industrial specialization coefficient is negative (and statistically significant) and the cluster specialization coefficient is positive (and statistically significant). Moreover, certain service-oriented clusters such as Business Services and Financial Services are estimated to have relatively high industry and cluster specialization coefficients (in absolute value), while the impact of cluster specialization is muted in some other clusters (such as Tobacco or Footwear, which are highly mature clusters).[30]

Finally, in Figure 2B, we plot the distribution of coefficients from a specification where we allow for EA-specific industry specialization and cluster specialization coefficients. As before, the results strongly suggest that our overall findings – convergence at the level of the industry, dynamic agglomeration at the level of the cluster – are not simply the result of a small number of "outlier" EAs but represent a robust and central tendency in the data.

Performance at the Cluster Level. So far we have focused on employment growth at the region-industry level. We have found that convergence forces dominate at the narrowest unit (i.e., at the EA-industry), and dynamic agglomeration forces arise among complementary economic activity. The same logic should also apply to employment growth at the region-cluster

[30] Industries with high (low) convergence effects tend to have high (low) within-cluster effect. The correlation between the industry spec and cluster spec coefficients is -0.17. Though the relationship is somewhat noisy, this negative correlation suggests that those clusters with particularly salient complementarities among industries tend to be subject to a high level of diminishing returns within industries.

level. Here, we examine this using a cross-sectional dataset of 6694 EA-clusters with positive employment in 1990, spanning 41 traded clusters and up to 177 EAs.

In Table 8, we specify cluster employment growth as a function of the strength of that cluster, related clusters, and common neighboring clusters, with (8-2) or without (8-1) EA and cluster fixed effects. The cluster growth models reveal convergence in employment at the EA-cluster level. Regional clusters with higher initial employment levels tend to experience lower employment growth (the annual convergence rate is 3.4% based on model 8-2). While cluster specialization in a region displays convergence, employment growth of a cluster is positively influenced by the presence of strong related clusters in the region and by the specialization of neighboring regions in the same cluster.[31] A one standard deviation increase in these variables is associated with a 2% increase in the annual EA-cluster employment growth (relative to the mean annual EA-cluster growth of 0.5%).[32]

Performance at the Region level. Finally, we test for the role of a region's cluster strength in the employment growth of *other* traded and local clusters in the region using the cross-section of 177 EAs. In Table 9, we first specify regional employment growth outside a region's strong clusters over 1990-2005 as a function of regional cluster strength in 1990 (9-1); we then implement our core specification (9-2). The results shown in (9-1) and (9-2) suggest that regions with high (relative) cluster strength (i.e., a high share of total traded employment in the region is accounted for by the strong clusters) are associated with higher employment growth outside the strong clusters. In other words, regional cluster strength matters for employment creation in other traded and local activities in the region.[33] For example, doubling the cluster strength in a region is associated with an increase of 1.7% in annual employment growth rate.

---

[31] These findings are robust to multiple sensitivity analyses. The results change only marginally when we condition on EA-clusters with positive employment in both 1990 and 2005; drop the 5% smallest and largest EAs; use MSAs instead of EAs; and exclude small regional clusters. As described in Section 3, we also test for spatial correlation running a FAR model on the residuals of model (8-2), under the null hypothesis of no spatial autocorrelation if $\rho=0$. We cannot reject the null hypothesis. The R-squared of the FAR model is zero, the coefficient of ρ is -.004, and the standard deviation of ρ is very large (.66).

[32] Similarly to the industry-level model findings, the cluster-level model finds greater inter-regional agglomeration benefits for smaller EAs (not reported).

[33] These findings also hold when we use alternative indicators of regional cluster strength such as the number of strong clusters in the region, the total level of employment in strong clusters, cluster strength weighting by the overlap among the strong clusters, or, alternatively, weighting each strong cluster by its degree of overlap with other traded clusters. The findings are also robust to dropping the 5% smallest and 5% largest EAs.

Overall, the findings suggest that the (relative and absolute) presence of strong traded clusters in a region will generate job opportunities for other (probably related) traded and local activities in the region. The ways and extent to which the traded clusters interact with local clusters will be the subject of related research.

6. Conclusions

In this paper we investigate the role of agglomeration in regional economic performance. Our core empirical finding is the co-existence of convergence within narrow economic units with agglomeration across complementary economic units captured by clusters. Industries located within a strong cluster are associated with higher employment growth, a finding which is robust across different clusters and regions. Industry employment growth is increasing in the strength of related clusters in the region as well as the strength of the cluster in geographically adjacent regions. Strong clusters also foster growth in wages, the number of establishments, and patenting. Importantly, these findings are conditioned on the average growth rate of individual industries and regions; our empirical methodology allows us to disentangle the impact of clusters from alternative drivers of regional economic performance.

We find that the cluster and related clusters surrounding a region-industry matters not only for the growth of existing industries but also for the *creation* of new industries in a region. In other words, new industries are born out of strong regional clusters. These findings suggest that clusters play a crucial role in the path of regional economic development (Porter 1990, 1998, 2003; Swann, 1992).

These findings, taken together, offer several important conclusions. First, the traditional distinction between industry specialization and regional diversity is misplaced. This dichotomy overlooks the powerful role played by complementary economic activity in shaping economic growth, and the central role of clusters as the manifestation of complementarity. Narrow regional specialization in an industry is likely to result in diminishing returns, and the presence of *unrelated* economic activity is unlikely to significantly enhance opportunities for growth but may increase congestion. However, the presence of complementary activity via clusters is a strong driver of growth through allowing firms ready access to key inputs, better interactions with customers, and facilitating experimentation and innovation.

Second, prior studies have focused on individual channels through which complementarities might operate. Building on Feldman and Audretsch (1999), for example, numerous studies have emphasized the role of the local scientific knowledge base and the potential for knowledge spillovers in shaping opportunities for innovation and entrepreneurship. While our results are consistent with such findings, the impact of related economic activity on economic performance is far broader. The presence of clusters, which arise out of multiple types of complementarities, seems to be a primary driver of growth in employment (and in other performance dimensions) across essentially all regions and clusters.

Our findings have important implications for the theoretical understanding of the main drivers of agglomeration and the role of clusters in agglomeration. While most theoretical work on agglomeration emphasizes the potential for cost efficiencies, risk mitigation, or geographically localized knowledge spillovers, the pervasive impact of clusters and related clusters suggests that the underpinnings of agglomeration may be far broader. For example, clusters may not simply reduce the cost of production but the cost of exchange, by enhancing trading relationships and the transparency of local input and output markets. The impact of local knowledge spillovers likely does not simply accrue to a single firm in an isolated way; rather, related local discoveries may simultaneously enhance the knowledge base of multiple local firms. In addition, qualitative studies of clusters emphasize the central role of specialized local institutions (from training facilities to infrastructure investments) in allowing potential complementarities to be realized (Porter and Emmons, 2003; Sölvell et al., 2006). While the precise design, role, and operation of such institutions varies widely by circumstance, little theoretical or empirical research has examined the impact of these localized institutions in shaping regional economic performance.

Our findings also carry several important policy implications. One is that effective regional policy should prioritize complementarities across related economic activity rather than seek to attract any type of investment, offer incentives to benefit a small number of firms, or favor particular high-technology fields such as biotechnology or software if the region has little strength in those areas. Instead the focus should be on how to leverage a region's strong clusters (Porter, 2003; Cortright, 2006; Ketels and Memedovic, 2008; Rodríguez-Clare, 2007). New industries will grow out of the most successful existing clusters.

Our results suggest that regional economic performance depends crucially on the cluster composition across nearby regions rather than within narrow political boundaries. The benefits arising from clusters often span multiple jurisdictions (and even states). Policies that enhance complementarities across jurisdictions, such as supporting infrastructure and institutions that facilitate access to demand, skills or suppliers in neighboring clusters, are important tools for regional development.

Our analysis also suggests several intriguing directions for research. While the current analysis takes the initial state of the cluster environment as a given, and holds cluster definitions constant, cluster structure can evolve over time. For example, whereas electronic computers (SIC 3571) may have been the central industry within the information technology cluster on an historical basis, it is possible that software (SIC 7372) may be the core industry within that cluster going forward. However, few studies have examined how the co-location patterns of industries change over time, or the role of the historical composition of industries in a region in shaping new industry growth. Understanding the drivers of the evolution of clusters is a crucial direction for future research.

7. References

Alcacer, J and W. Chung, 2010, “Location Strategies for Agglomeration Economies, Harvard Business School Working Paper, No. 10-064.

Anselin,L., 1988, *Spatial Econometrics: Methods and Models*, Dorddrecht: Kluwer Academic Publishers.

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

Baldwin R., R. Forslid, P. Martin, G. Ottaviano and F. Robert-Nicoud, 2003, *Public Policies and Economic Geography*, Princeton: Princeton University Press.

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

Barro, R.J. and X. Sala-i-Martin, 1992, "Convergence," Journal of Political Economy, University of Chicago Press, 100(2), pp. 223-51.

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

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

Benhabib, J. and M. Spiegel, 1994, “The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data,” *Journal of Monetary Economics*, 34, pp. 143-173.

Bönte, 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.

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

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.

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

Combes, P.P., 2000, “Economic Structure and Local Growth: France, 1984-1993,” *Journal of Urban Economics*, 47, pp. 329-355.

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.

Cortright J.and H. Mayer, 2002, “Signs of Life: The Growth of Biotechnology Centers in the U.S.,” The Brookings Institution.

Delgado, M., 2005, “Do Clusters Really Matter for Companies’ Competitive Strategies? Evidence at the Country Level,” Manuscript.

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

Dobkins, L.H. and Y.M. Ioannides 2001, "Spatial Interactions Among U.S. Cities: 1900-1990," Regional Science and Urban Economics, 31, pp. 701-731.

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

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.

Feldman, M.P. and Audretsch, D., 1999, "Innovation in cities: Science-based diversity, specialization and localized competition," *European Economic Review*, 43, pp. 409-429.

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.

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

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(5), pp. 685-697.

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

Fujita, M. and Thisse J.F., 2002, *Economics of Agglomeration: Cities, Industrial Location and Regional Growth*. Cambridge: Cambridge University Press.

Geppert, K. and A. Stephan, 2008. "Regional disparities in the European Union: Convergence and agglomeration," *Papers in Regional Science*, 87(2), pp. 193-217.

Glaeser, E.L., H.D. Kallal, J.A. Scheinkman, and A. Shleifer, 1992, "Growth in Cities," *Journal of Political Economy* 100(6), pp. 1126-1152.

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

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.

Henderson, J.V., 2005,"Urbanization and Growth," at P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, P., North Holland.

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

Henderson, J.V., 1997, "Externalities and Industrial Development," *Journal of Urban Economics*, 42(3), pp. 449-470.

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

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

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.

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

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.

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

Neary, P., 2001, "Of Hype and Hyperbolas: Introducing the New Economic Geography," *Journal of Economic Literature*, 39(2), pp. 536–561.

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

Porter, M.E., 1998, 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., 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 W. Emmons, 2003, "Institutions for Collaboration: Overview," Harvard Business School Note 703-436.

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

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

Rodríguez-Clare, A., 2007, "Clusters and Comparative Advantage: Implications for Industrial Policy," *Journal of Development Economics*, 82, pp. 43-57.

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.

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

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

Silverman, B.S., 1999, "Technological resources and the direction of corporate diversification," *Management Science*, 45(8), pp. 1109-1124.

Sölvell, Ö., G. Lindqvist, and C. Ketels, 2006, "Cluster Initiatives in Developing and Transition Economies." Report, Ivory Tower AB, Stockholm, http://www.cluster-research.org/dldevtra.htm.

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), pp. 424-62.

Swann, P., 1992, *The Dynamics of Industrial Clusters*, Oxford.

Swann, P., 1998, "Clusters in the US Computing Industry," in Swann, 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.

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*, 34(3), pp. 221-241.

Table 1: Variables' Definitions and Descriptive Statistics

Variables	Definition	EA-Industry N=55083	EA-Cluster N=6694	Region (EA) N=177
		Mean (Std Dev)		
PERCENTAGE EMPLOYMENT $CHANGE_{1990\text{-}05}$	$(employ_{05}-employ_{90})/employ_{90}$	1.26 (7.92)		
EMPLOYMENT $GROWTH_{1990\text{-}05}$	$\ln(employ_{2005}/employ_{1990})$	-.32 (1.69)	.08 (1.03)	
REG EMPLOYMENT GROWTH (Outside Strong Clusters)	Regional (total) employment growth excluding strong clusters			.29 (.15)
$EMPLOYMENT_{1990}$	Employment level	576.45 (2545.35)	4750.64 (15044.92)	
INDUSTRY $SPEC_{1990}$	Industry employment-based Location Quotient $LQ_{i,r} = \frac{employ_{i,r}/employ_r}{employ_{i,US}/employ_{US}}$	2.34 (7.21)		
	Cluster employment-based LQ	1.31 (2.20)	1.29 (2.53)	
CLUSTER $SPEC_{1990}$				
LINKED CLUSTERS $SPEC_{1990}$	Related clusters' employment-based LQ (weighted by cluster overlap)	1.15 (1.04)	1.14 (1.34)	
CLUSTER SPEC in $NEIGHBORS_{1990}$	Neighboring clusters' average employment-based LQ	1.28 (1.28)	1.23 (1.40)	
SIC2 $SPEC_{1990}$	two-digit SIC code employment-based LQ	1.31 (1.95)		
REG STRONG $CLUSTERS_{1990}$	Number of strong clusters in a region (i.e., clusters with top 20% Cluster Spec by cluster)			6.91 (3.28)
REG EMPLOYMENT $_{1990}$ (Outside strong clusters)	Regional (total) employment (outside the strong clusters)			435015.3 (810172.5)
REG CLUSTER $STRENGTH_{1990}$	Share of regional traded employ in the region's strong clusters			.47 (.15)

Notes: These indicators are based on the CBP data. In the EA-industry models Cluster Specialization is measured outside the industry.

Table 2: Average EA-Industry Annualized Employment Growth over 1990-2005 by Level of Industry and Cluster Strength (N=55083)

		INDUSTRY $SPEC_{1990}$ Low	INDUSTRY $SPEC_{1990}$ High
CLUSTER $SPEC_{1990}$ (Outside the industry)	**Low**	**$\Delta EMPLOY_{i,r}$= 0.13** N= 16,732	**$\Delta EMPLOY_{i,r}$= 0.00** N=10,978
	High	**$\Delta EMPLOY_{i,r}$= 0.20** N= 10,964	**$\Delta EMPLOY_{i,r}$= 0.01** N= 16,409

Notes: ΔEMPLOY is the average EA-industry *annualized* employment percentage change ($1/15*(employ_{05}-employ_{90})/employ_{90}$). Low and High are based on the median of the variable for each industry. All the averages are significantly different from each other at 1% level.

Table 3: EA-Industry Employment Growth over 1990-2005 (N=55083)

	INDUSTRY EMPLOYMENT GROWTH							
	3-1	3-2	3-3	3-4 All Industries N=99047	3-5 Random Clusters	3-6 SIC2	3-7	3-8
Ln INDUSTRY SPEC	**-.397**	**-.437**	**-.395**	**-.414**	**-.355**	**-.388**	**-.442**	**-.405**
	(.006)	**(.006)**	**(.006)**	**(.003)**	**(.000)**	**(.006)**	**(.007)**	**(.006)**
Ln CLUSTER SPEC		**.187**	**.200**	**.216**	-.001		**.160**	**.149**
(Outside the industry)		**(.011)**	**(.008)**	**(.008)**	(.011)		**(.013)**	**(.009)**
Ln RELATED							***.042***	**.091**
CLUSTERS SPEC							***(.021)***	**(.013)**
Ln CLUSTER SPEC in							**.050**	**.104**
NEIGHBORS							**(.018)**	**(.012)**
Ln SIC2 SPEC						**.197**		
(Outside the industry)						**(.008)**		
EA Fixed Effects	**No**	**No**	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**Yes**
INDUSTRY Fixed Effects	**No**	**No**	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**Yes**
R-Squared	**.113**	**.125**	**.156**	**.159**		**.154**	**.126**	**.159**

Notes: Bold and Bold-Italic numbers refer to coefficients significant at 1% and 5% levels, respectively. Standard errors are clustered by EA-Cluster. Model (3-5) uses random cluster definition and reports the bootstrapped standard errors. In (3-6) traded industries are grouped by 2-digit SIC code and the standard errors are clustered by EA-SIC2. All the models (but 3-1) include a dummy equal to 1 if any employment in 1990 in the cluster (outside the industry); and models (3-7) and (3-8) also include dummies equal to 1 if any initial employment in the related clusters or in the neighboring clusters (not reported). Our findings only change trivially when dropping these dummies.

Table 4: Creation of New EA-industries over 1990-2005 (Probit Models; N=48430)

	ANY POSITIVE EMPLOYMENT$_{2005}$		
	1	2	3
Ln CLUSTER SPEC	**.016**	**.033**	**.026**
(Outside the industry)	**(.002)**	**(.002)**	**(.002)**
Ln RELATED CLUSTERS SPEC			**.015**
			(.004)
Ln CLUSTER SPEC in NEIGHBORS			**.021**
			(.003)
EA Fixed Effects	**No**	**Yes**	**Yes**
INDUSTRY Fixed Effects	**No**	**Yes**	**Yes**
(Pseudo) R-Squared	**.008**	**.304**	**.306**

Notes: Bold numbers refer to coefficients significant at 1% level. The sample conditions on EA-industries with zero employment in 1990. The reported coefficients are the marginal effects of the estimated probit model. Standard errors are clustered by EA-Cluster. All explanatory variables are in logs. In addition to the reported explanatory variables, all the models include a dummy equal to 1 if any employment in 1990 in the cluster (outside the industry); and model (4-3) also includes dummies equal to 1 if any employment in the related clusters or in the neighboring clusters. Our findings only change trivially when dropping these dummies.

Table 5: EA-Industry Growth in Wages, Establishments and Patents (1990-2005)

	ΔWAGE		Δ ESTABLISHMENT		Δ PATENT	
	N=45843		N=48921		N=40266	
	1	2	3	4	5	6
Ln INDUSTRY WAGE	**-.798**	**-.799**				
	(.006)	**(.006)**				
Ln CLUSTER WAGE	**.076**	**.069**				
(Outside the industry)	**(.009)**	**(.010)**				
Ln RELATED CLUSTERS WAGE		.005				
		(.009)				
Ln CLUSTER WAGE in NEIGHBORS		**.077**				
		(.017)				
Ln INDUSTRY $SPEC_{Estab}$			**-.393**	**-.399**		
			(.005)	**(.005)**		
Ln CLUSTER $SPEC_{Estab}$			**.167**	**.132**		
(Outside the industry)			**(.007)**	**(.009)**		
Ln RELATED CLUSTERS $SPEC_{Estab}$				**.142**		
				(.011)		
Ln CLUSTER SPEC in $NEIGHBORS_{Estab}$				**.028**		
				(.010)		
Ln INDUSTRY $SPEC_{Patent}$					**-.780**	**-.787**
					(.009)	**(.009)**
Ln CLUSTER $SPEC_{Patent}$					**.136**	**.080**
(Outside the industry)					**(.013)**	**(.015)**
Ln RELATED CLUSTERS $SPEC_{Patent}$						**.159**
						(.025)
Ln CLUSTER SPEC in $NEIGHBORS_{Patent}$						**.154**
						(.021)
EA Fixed Effects	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
INDUSTRY Fixed Effects	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
R-Squared	**.464**	**.465**	**.215**	**.220**	**.498**	**.502**

Note: Bold numbers refer to coefficients significant at 1% level. Standard errors are clustered by EA-cluster. All models include EA-industry with positive economic activity (wages, establishments or patents) in both 1990 and 2005.

Table 6: EA-Industry Employment Growth by Region Size (N=55083)

	INDUSTRY EMPLOYMENT GROWTH	
Ln INDUSTRY SPEC	**-.352**	**-.365**
	(.010)	**(.010)**
Ln CLUSTER SPEC (Outside the industry)	**.166**	**.113**
	(.011)	**(.013)**
Ln RELATED CLUSTERS SPEC		**.051**
		(.018)
Ln CLUSTER SPEC in NEIGHBORS		**.151**
		(.019)
EA SIZE*Ln INDUSTRY SPEC	**-.068**	**-.065**
	(.012)	**(.012)**
EA SIZE*Ln CLUSTER SPEC	**.063**	**.065**
	(.013)	**(.015)**
EA SIZE*Ln RELATED CLUSTERS SPEC		**.078**
		(.024)
EA SIZE*Ln CLUSTER SPEC in NEIGHBORS		**-.082**
		(.024)
EA Fixed Effects	**Yes**	**Yes**
INDUSTRY Fixed Effects	**Yes**	**Yes**
R-Squared	**.157**	**.161**

Notes: Bold numbers refer to coefficients significant at 1%. EA SIZE is a dummy variable equal to 1 for large EAs.

Table 7: Estimated Coefficients of Industry and Cluster Specialization by Cluster: EA-Industry Employment Growth (model 3-3)

Cluster	**Type**	**Industry Spec 1**	**Cluster Spec 2**
Business Services	Service	**-0.753**	**0.361**
Financial Services	Service	**-0.598**	**0.307**
Metal Manufacturing	Other	**-0.383**	**0.304**
Analytical Instruments	High-tech	**-0.562**	**0.291**
Production Technology	Other	**-0.368**	**0.291**
Processed Food	Other	**-0.230**	**0.284**
Heavy Construction Services	Service	**-0.348**	**0.254**
Plastics	Other	**-0.384**	**0.253**
Communications Equipment	High-tech	**-0.557**	**0.252**
Automotive	Other	**-0.270**	**0.237**
Hospitality and Tourism	Service	**-0.497**	**0.226**
Medical Devices	High-tech	**-0.363**	**0.221**
Apparel	Other	**-0.665**	**0.218**
Agricultural Products	Other	**-0.260**	**0.208**
Oil and Gas Products and Services	Service	**-0.162**	**0.206**
Heavy Machinery	Other	**-0.346**	**0.199**
Aerospace Vehicles and Defense	High-tech	**-0.294**	**0.188**
Fishing and Fishing Products	Other	***-0.147***	**0.185**
Information Technology	High-tech	**-0.544**	**0.183**
Publishing and Printing	Other	**-0.501**	**0.174**
Textiles	Other	**-0.409**	**0.170**
Building Fixtures, Equipment and Services	Other	**-0.486**	**0.170**
Prefabricated Enclosures	Other	**-0.174**	**0.160**
Biopharmaceuticals	High-tech	**-0.393**	**0.160**
Motor Driven Products	Other	**-0.287**	**0.158**
Sporting, Recreational and Children's Goods	Other	**-0.640**	**0.157**
Chemical Products	High-tech	**-0.273**	**0.154**
Transportation and Logistics	Service	**-0.355**	**0.153**
Jewelry and Precious Metals	Other	**-0.353**	**0.144**
Forest Products	Other	**-0.222**	**0.119**
Furniture	Other	**-0.409**	**0.112**
Leather and Related Products	Other	**-0.504**	**0.108**
Education and Knowledge Creation	Service	**-0.422**	**0.107**
Construction Materials	Other	**-0.328**	**0.106**
Lighting and Electrical Equipment	Other	**-0.285**	***0.105***
Tobacco	Other	**-0.352**	***0.104***
Aerospace Engines	High-tech	**-0.213**	***0.102***
Footwear	Other	**-0.582**	***0.077***
Power Generation and Transmission	Service	**-0.483**	***0.076***
Entertainment	Service	**-0.540**	0.057
Distribution Services	Service	**-0.403**	-0.013
Avg		-0.399	0.179
Avg High-tech		-0.400	0.194
Avg Service-oriented		-0.456	0.173
Avg Other (Traditional Manufacturing)		-0.373	0.176

Note: Bold and Bold-Italic numbers refer to coefficients significant at 1% and 5% levels, respectively. These coefficients are obtained by estimating model 3-3 (Table 3) allowing the Industry Spec and Cluster Spec coefficients to vary by cluster.

Table 8: EA-Cluster Employment Growth over 1990-2005 (N=6694)

	CLUSTER EMPLOY GROWTH	
	1	2
Ln CLUSTER SPEC	**-.348**	**-.403**
	(.017)	**(.016)**
Ln RELATED CLUSTERS SPEC	**.075**	**.095**
	(.022)	**(.019)**
Ln CLUSTER SPEC in NEIGHBORS	**.150**	**.129**
	(.020)	**(.019)**
EA Fixed Effects	**No**	**Yes**
CLUSTER Fixed Effects	**No**	**Yes**
R-Squared	**.147**	**.367**

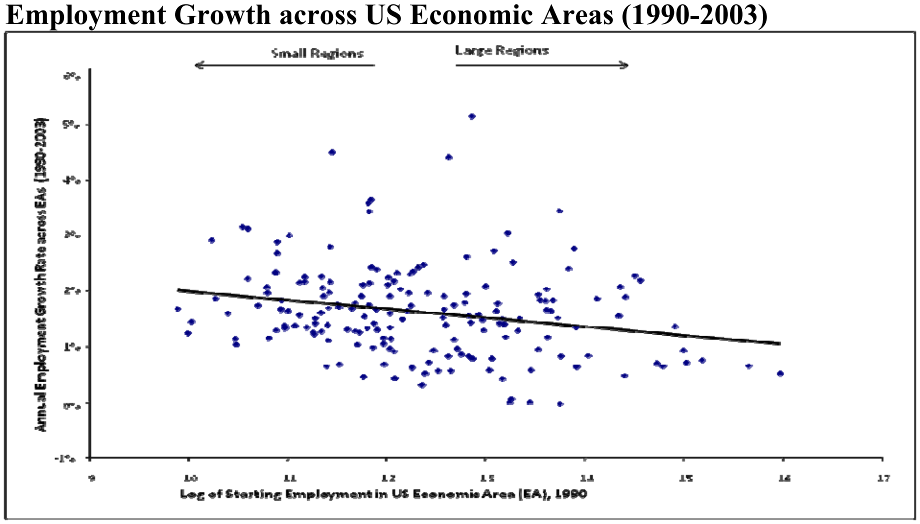
Note: Bold numbers refer to coefficients significant at 1% level. Standard errors clustered by EA. All the variables are in logs.

Table 9: Regional Employment Growth Outside the Strong Clusters (N=177)

	REG EMPLOYMENT GROWTH (Outside Strong Clusters)	
	1	2
REG CLUSTER STRENGTH	**.192**	**.251**
	(.081)	**(.065)**
Ln REG EMPLOYMENT		**-.017**
(Outside strong clusters)		**(.007)**
NATIONAL EMPLOY GROWTH of STRONG CLUSTERS		.056
(Outside the EA)		(.048)
CENSUS REGION Fixed Effects		**Yes**
R-Squared	**.038**	**.414**

Note: Bold numbers refer to coefficients significant at 5%. Robust standard errors are in parentheses.

Figure 1: Employment Growth across US Economic Areas (1990-2003)



Note: Adapted from Porter (2003). Based on County Business Patterns dataset.

Figure 2a: EA-Industry Employment Growth: Estimated Convergence and Cluster Effect by Cluster

Distribution of Coef. of Industry Spec by cluster (Mean=-.40 Std. Dev=.14; 41 clusters)	Distribution of Coef. of Cluster Spec by cluster (Mean=.18 Std. Dev=.08; 41 clusters)

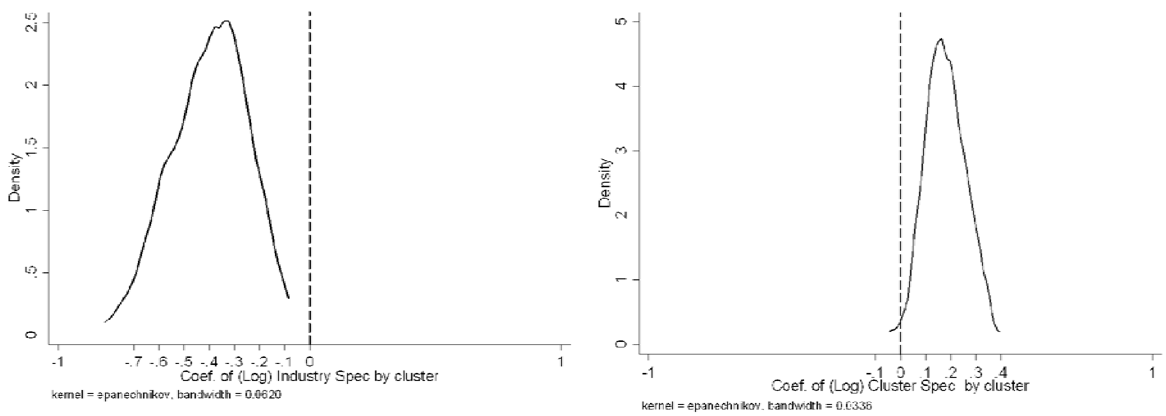
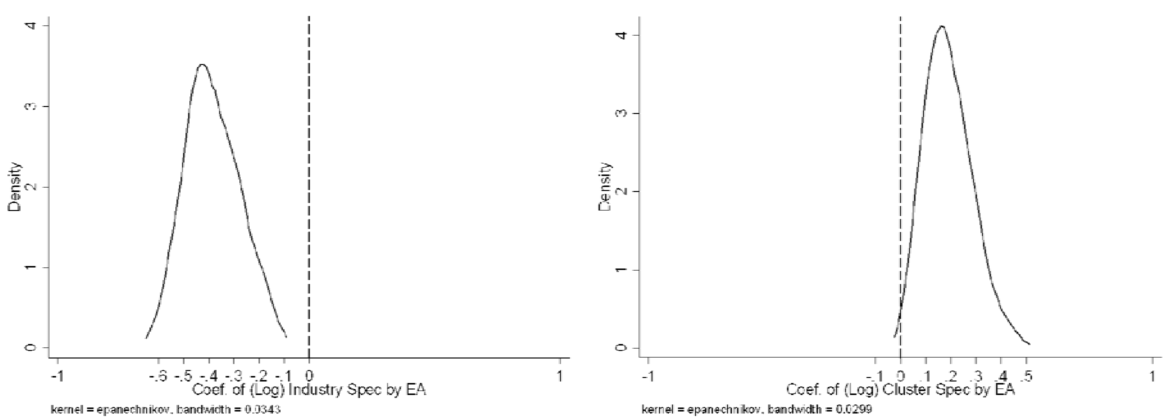


Figure 2b: EA-Industry Employment Growth: Estimated Convergence and Cluster Effects by EA

Distribution of Coef. of Industry Spec by EA (Mean= -.38 Std. Dev=.10; 177 EAs)	Distribution of Coef. of Cluster Spec by EA (Mean=.19 Std. Dev=.09; 177 EAs)



Note: The graphs plot the Kernel density of the estimated coefficients. These coefficients are obtained by estimating model 3-3 (Table 3) allowing the Industry Spec and Cluster Spec coefficients to vary by cluster (Figure 2a) or by EA (Figure 2b).

Appendix A

Table A1: Attributes of Traded Clusters in the US Economy, 2005.

Name (41 traded clusters)	Type	Narrow Clusters 2005 Employ (1000)	2005 Patents (1000)	#Industries total	svc.	Broad Clusters #Industries
Aerospace Engines	High-tech	82.669	0.084	2	0	23
Aerospace Vehicles & Defense	High-tech	319.800	0.393	6	0	28
Analytical Instruments	High-tech	578.593	9.654	10	0	27
Biopharmaceuticals	High-tech	278.582	2.201	4	0	26
Chemical Products	High-tech	359.703	2.853	21	0	40
Communications Equipment	High-tech	245.582	9.911	9	0	35
Information Technology	High-tech	644.532	12.390	9	3	27
Medical Devices	High-tech	375.063	2.765	8	0	29
Business Services	Service	4748.123	0.177	21	21	26
Distribution Services	Service	1803.523	0.052	19	19	23
Education & Knowledge Creation	Service	2779.839	0.367	10	9	38
Entertainment	Service	1174.900	1.482	13	9	20
Financial Services	Service	3212.496	0.016	21	21	32
Heavy Construction Services	Service	1752.938	0.767	19	6	30
Hospitality & Tourism	Service	2671.877	0.239	22	19	34
Oil & Gas Products & Services	Service	415.732	0.487	12	6	24
Power Generation & Transmission	Service	267.663	0.503	6	1	16
Transportation & Logistics	Service	1580.522	0.194	17	16	29
Agricultural Products	Other Mfg.	279.265	0.100	20	6	46
Apparel	Other Mfg.	289.520	0.166	27	0	34
Automotive	Other Mfg.	1164.530	4.033	15	0	32
Building Fixtures, Equip. & Services	Other Mfg.	665.425	0.849	25	2	57
Construction Materials	Other Mfg.	174.402	0.470	11	0	23
Fishing & Fishing Products	Other Mfg.	44.402	0.068	3	0	4
Footwear	Other Mfg.	22.926	0.066	5	0	7
Forest Products	Other Mfg.	361.999	0.521	8	0	21
Furniture	Other Mfg.	280.577	0.153	10	0	26
Heavy Machinery	Other Mfg.	354.144	1.177	10	2	26
Jewelry & Precious Metals	Other Mfg.	101.919	0.024	7	1	11
Leather & Related Products	Other Mfg.	120.546	0.394	13	0	22
Lighting & Electrical Equipment	Other Mfg.	229.775	1.070	10	0	24
Metal Manufacturing	Other Mfg.	1131.574	2.082	44	0	67
Motor Driven Products	Other Mfg.	322.544	1.466	12	0	28
Plastics	Other Mfg.	740.110	2.988	9	0	22
Prefabricated Enclosures	Other Mfg.	267.696	0.312	12	0	20
Processed Food	Other Mfg.	1321.019	0.560	43	2	49
Production Technology	Other Mfg.	535.772	3.688	23	0	46
Publishing & Printing	Other Mfg.	720.938	1.436	26	3	36
Sporting, Recreational & Children's Goods	Other Mfg.	86.492	0.312	3	0	8
Textiles	Other Mfg.	257.760	0.409	20	0	25
Tobacco	Other Mfg.	32.773	0.042	4	0	4
Totals		32798.242	66.922	589	146	1145
Average				14.4	3.6	27.9

Notes: Based on County Business Patterns (CBP) and CMP datasets. There are 589 traded four-digit SIC code industries (146 of them are service (svc.) industries). Service-oriented clusters are those with more than 35% of employment in service industries. High-tech clusters are manufacturing clusters with high patenting.

Table A2: Automotive Cluster: Constituent Industries and Related Clusters

4-digit SIC Industries		Clusters Related 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 related clusters	.32	.25	.08	.08	.03	.03

Source: Porter's (2003) cluster definitions. These 15 industries constitute the narrow cluster definition. 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).

Table A3: Location of Strong Automotive Clusters in 1990

Top 20% cluster specialization (LQ) & Top 10% share of US cluster employment (SHR)			Top 20% Cluster Specialization (LQ)		
	LQ	SHR %		LQ	SHR %
Detroit-Warren-Flint, MI	9.5	25.3	La Crosse, WI-MN	4.1	0.3
Toledo-Fremont, OH	6.5	2.7	Traverse City, MI	3.9	0.3
Fort Wayne-Huntington-Auburn, IN	6.4	2.1	Jonesboro, AR	3.8	0.3
Dayton-Springfield-Greenville, OH	5.1	2.9	Cape Girardeau-Jackson, MO-IL	3.3	0.3
Grand Rapids-Muskegon-Holland, MI	4.5	3.0	Lincoln, NE	3.2	0.4
Indianapolis-Anderson-Columbus, IN	3.6	4.3	Huntsville-Decatur, AL	3.1	1.0
South Bend-Mishawaka, IN-MI	3.6	1.3	Peoria-Canton, IL	3.0	1.0
Cleveland-Akron-Elyria, OH	3.2	6.1	Asheville-Brevard, NC	2.8	0.5
Nashville-Murfreesboro-Columbia, TN	3.0	2.4	Kearney, NE	2.6	0.3
Milwaukee-Racine-Waukesha, WI	2.2	2.2	Madison-Baraboo, WI	2.6	1.0
Columbus-Marion-Chillicothe, OH	2.1	1.8	Mason City, IA	2.5	0.1
			Alpena, MI	2.4	0.1
			Erie, PA	2.4	0.4
			Louisville-Scottsburg, KY-IN	2.2	1.1
			Fayetteville-Springdale-Rogers, AR-MO	2.2	0.2
			Buffalo-Niagara-Cattaraugus, NY	2.0	1.2
			Joplin, MO	2.0	0.2
			Oklahoma City-Shawnee, OK	1.9	1.0
			Evansville, IN-KY	1.8	0.5
			Knoxville-Sevierville-La Follette, TN	1.8	0.6
			Lexington-Fayette-Richmond, KY	1.8	0.7
			Waterloo-Cedar Falls, IA	1.8	0.1

Note: Based on County Business Patterns dataset. The strong Automotive clusters are the top 20% EAs by employment Cluster Spec (33 EAs out of 173 EAs with any employment in automotive cluster).

Figure A1: Location of Strong Regional Clusters in 1990 (Top 20% of EAs by employment Cluster Specialization)

- 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 considered strong clusters)

Fig. A1.1: Strong Automotive Clusters in 1990 (See Table A3. Based on CBP dataset)

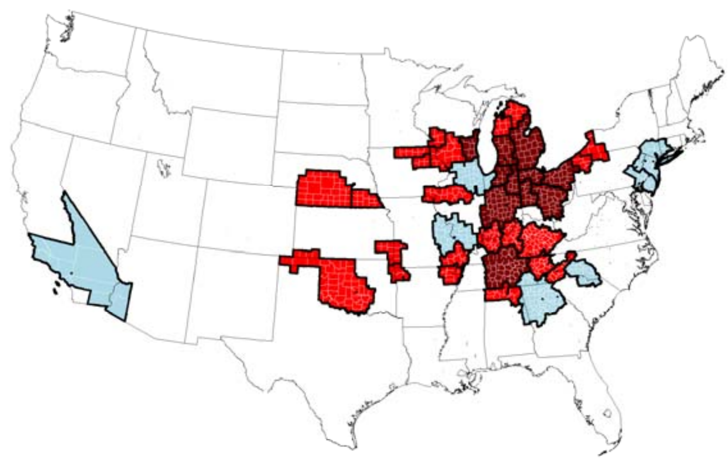
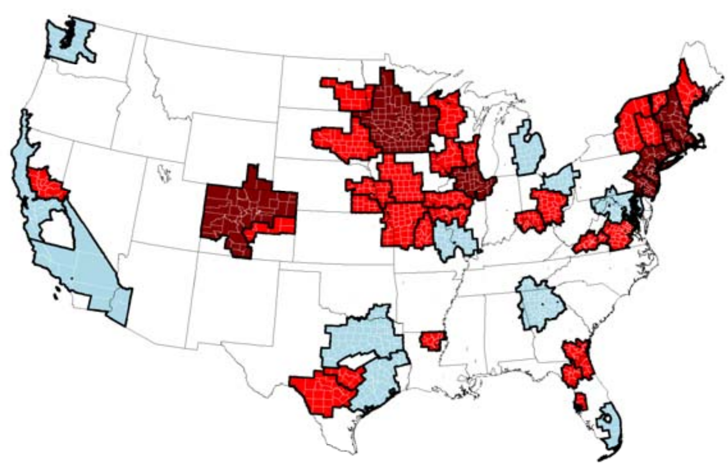


Fig. A1.2: Strong Financial Services Clusters in 1990 (Based on CBP dataset)



Appendix B

Table B1: Cluster-Level Specialization Variables: Correlation Table (N=6694)

		V_1	V_2	V_3
CLUSTER SPEC $_{Employ}$	V_1	1.0		
RELATED CLUSTERS SPEC $_{Employ}$	V_2	.28	1.0	
CLUSTER SPEC in NEIGHBOR $_{Employ}$	V_3	.50	.26	1.0

Note: All correlations are significant at 1% level. All the variables are in log.

Table B2: Variables' Definitions and Descriptive Statistics for Table 5 (EA-industry growth)

Variables	Definition	Mean (Std Dev)	Source
WAGE GROWTH$_{1990\text{-}05}$	Industry wage growth	.12 (.49)	LBD
ESTABLISHMENT GROWTH$_{1990\text{-}05}$	Industry establishment growth	.17 (.87)	CBP
PATENT GROWTH$_{1990\text{-}05}$	Industry patent growth	.19 (1.30)	USPTO
INDUSTRY WAGE$_{1990}$	Industry wage	34032.79 (25977.68)	LBD
CLUSTER WAGE $_{1990}$ (Outside the industry)	Industry cluster wage	35069.27 (18070.13)	LBD
RELATED CLUSTERS WAGE$_{1990}$	Wages in related clusters	54035.93 (28498.58)	LBD
CLUSTER WAGE in NEIGHBORS$_{1990}$	Neighboring clusters' average wage	33509.65 (9707.44)	LBD
INDUSTRY ESTABLISHMENT$_{1990}$	Industry count of establishments	19.31 (83.29)	CBP
INDUSTRY SPEC$_{Estab,\ 1990}$	Industry establishment-LQ	1.63 (2.93)	CBP
CLUSTER SPEC$_{Estab,\ 1990}$	Cluster establishment-LQ	1.06 (1.10)	CBP
RELATED CLUSTERS SPEC$_{Estab,\ 1990}$	Related clusters' establishment-LQ (weighted by cluster overlap)	.96 (.44)	CBP
CLUSTER SPEC in NEIGHBORS$_{Estab,1990}$	Neighboring clusters' average establishment-LQ	1.02 (.76)	CBP
INDUSTRY PATENT$_{1990}$	Patents granted	.90 (5.08)	USPTO
INDUSTRY SPEC$_{Patent,\ 1990}$	Industry patent-LQ	1.27 (2.32)	USPTO
CLUSTER SPEC$_{Patent,\ 1990}$	Cluster patent-LQ	1.13 (1.29)	USPTO
RELATED CLUSTERS SPEC$_{Patent,\ 1990}$	Related clusters' patent-LQ (weighted by cluster overlap)	1.04 (.48)	USPTO
CLUSTER SPEC in NEIGHBORS$_{Patent,\ 1990}$	Neighboring clusters' average patent-LQ	1.11 (.63)	USPTO

Notes: Sources: Longitudinal Business Database (LBD) of the Census Bureau; County Business Patterns (CBP); and United States Patent and Trademark Office (USPTO) datasets. Wages in 2002 $US. The number of EA-industries are 45843 (wage indicators), 48921(establishment indicators), and 40266 (patent indicators).