METHOD FOR IDENTIFYING LOCAL AND DOMESTIC INDUSTRIAL CLUSTERS USING INTERREGIONAL COMMODITY TRADE DATA

Rainer vom Hofe
School of Planning
University of Cincinnati
Cincinnati OH[1]

Saurav Dev Bhatta
Urban Planning and Policy Program
University of Illinois at Chicago
Chicago IL

ABSTRACT

Defining geographic as well as industrial boundaries in industrial cluster analysis is a challenging task. The most widely used methods for identifying industrial clusters utilize data from input-output tables. Given that inter-industry linkages are not confined to political boundaries, the questions discussed in the relevant literature on input-output based approaches focus on whether to use a regional or a national input-output framework. But neither framework can capture non-local transactions, which may be of major importance for smaller study regions.

In the presented research, we expand the regional Chicago input-output framework to an interregional framework which accounts for both regional and non-local (i.e., commodity imports and exports by industry) inter-industry transactions. Applying factor analytical techniques to the data in this inter-regional framework, we are able to derive two sets of industrial clusters which we refer to as local and domestic clusters. In addition, we also identify the key sectors for each cluster based on indicators for industry backward and forward linkages. A comparison of these local and domestic clusters shows that while there are some similarities between these two types of clusters, there are also some significant differences between them. Thus, we demonstrate that it is important to include non-local inter-industry transactions as well in applied industrial cluster analysis

[1] Many thanks to Ke Chen, Assistant Professor of Economics, Finance, Geography and Urban Studies at the East Tennessee State University, for her assistance with the factor analysis.

INTRODUCTION

In recent years, industrial cluster analysis has received increasing attention from regional economists, economic development practitioners and policy makers alike. After decades of *first-wave* industrial attraction efforts (i.e., smokestack chasing) followed by years of *second-wave* industrial retention and expansion strategies in the 1980s, the beginning of the 1990s has seen the emergence of what might be called *third-wave* economic development strategies (Herbers 1990). These third-wave approaches are characterized by a focal shift from firm and business-level assistance to strategies that stimulate the formation of networks and clusters of industries, workers, social and political institutions, and demand markets (Fosler, 1992). Furthermore, these strategies are industrial-policy oriented, rather than project-oriented, and recognize the importance of community and business collaborations in strengthening entire and globally competitive industrial sectors (Bradshaw and Blakely, 1999).

As earlier work by Isard (1959), Czamanski (1974), and Ó hUallacháin (1984) indicates, the concept of cluster-based economic development is not a new economic development theory. Its roots can actually be traced to Marshall's (1920) ideas of agglomeration economies. But this concept gained currency only in the 1990s when it emerged as the new economic development panacea (Martin and Sunley, 2003). And much of the popularity boost of industrial clusters is attributable to Porter's (1990) widely-cited *Diamond of Advantage* in which he identifies four determinants that foster cluster development and where governments can play a proactive role.[2] But despite all the theoretical and conceptual understanding of the working mechanisms of industrial clusters, there is still no consensus in the relevant literature on the use of quantitative methods for proper cluster identification.

One important strain of cluster identification methodologies has evolved around using input-output (I-O) tables. The basic idea here is to identify inter-industry linkages based on vertical purchase-sales relationships, or group industries according to similarities in buying/selling patterns. Though widely applied, I-O based approaches for identifying regional clusters are open to a number of criticisms. One major shortcoming of these methods is that since I-O tables are aspatial by design, clusters identified using I-O based methods do not contain any information on how close firms and businesses are located to each other (Latham 1976). As suggested by Ó hUallacháin (1984), this criticism can be partially addressed by arguing that clusters identified using I-O tables do represent spatial agglomerations of industries when the region under consideration is relatively small. But even when the aspatial nature of I-O tables is not an issue, the identified regional clusters have yet another limitation, namely their inability to account for nonregional industry trade (Feser and Luger 2002). As a result, local clusters that cross regional boundaries might not necessarily be identified using these methods, a particularly relevant limitation in these times of increasing globalization.

The primary goal of this paper is to present a revised I-O-based methodological framework for identifying meaningful industrial clusters and apply it to the Chicago Metropolitan Statistical Area (MSA). The contribution of this research to the existing body of cluster literature is that our approach uses an extended regional I-O table that incorporates commodity trade between the region and the rest of the nation. This modification allows

[2] The four determinants included in Porter's *Diamond* are: firm strategy and rivalry, factor and demand conditions, and a network of related and supporting industries.

us to identify two types of clusters, which we refer to as local and domestic clusters. Following Czamanski and Ablas (1979) and Feser and Bergman (2000), we define an industrial cluster as a group of industries that are linked by significant inter-industry trade, similarities in buying/selling patterns, or similarities in markets for resources and outputs. A local industrial cluster is defined as a cluster confined to the region under consideration while a domestic cluster could be part of a supra-regional cluster with a strong presence in the local economy.

An important aspect of this research is the recognition that clusters do not represent purely local or purely national phenomena. As the focus of many cluster studies is on sub-national regions, such as metropolitan, state, or multi-state regions, non-regional trade activities are often very important for the study area. Therefore, clusters cannot be properly studied in isolation from the surrounding national economy. Focusing on selected clusters for the Chicago Metropolitan region, we will show that a comparison of local and domestic clusters can help us to understand the existing strengths of the local economy as well as provide insights for devising potential cluster-based economic development strategies.

The next section provides a brief overview of industrial cluster concepts and cluster identification methodologies based on input-out tables. The relevance of this research is outlined in section III, followed by a description of the data and study region in section IV. The next two sections present the methodology and findings of this study. And the final section presents some concluding remarks.

IMPORTANT CONCEPTS AND METHODOLIGIES IN INDUSTRIAL CLUSTER ANALYSIS

In his widely cited book *Principles of Economics*, Marshall (1920) described three agglomeration forces which explain the existence of spatial conglomerations of firms. These economies of localization are *knowledge spillovers* among firms, *labor market pooling*, and cost advantages produced by the *sharing of industry-specific non-traded inputs*. Subsequently, Hoover (1948) extended Marshall's concept of external *localization economies* by identifying two more agglomeration forces, namely, the existence of large, diverse markets (external *economies of urbanization*) and the availability of large and specialized factors of production increases (*internal returns to scale*). These theories of agglomeration economies have played a key role in shaping the definitions of industrial clusters.

Cluster concepts that focus on regional industrial specialization or the similarity among industries can trace their roots to Marshall's and Hoover's theories of agglomeration economies. Hanson (2000), for instance, emphasizes the role of localization economies in the formation of clusters and states that "the existence of localized externalities implies that firms prefer to be near large agglomerations of other firms in their own industry or related industries" (p. 4). Similarly, Hill and Brennan (2000) define a competitive industrial cluster as a "concentration of competitive firms or establishments in the same industry" (p. 67). Based on similar concepts, Munnich et al. (1998), Rex (1999), Botham et al. (2001), and Peters (2004) measure regional specialization and identify industry agglomerations through location quotients computed using industry employment data. However, as pointed out by Doeringer and Terkla (1995) and Rosenfeld (1997), location quotients are by no means sufficient to identify industrial clusters as they fail to account for the role of inter-industry trade linkages in the spatial grouping of industries.

Isard et al.'s (1959) pioneering work on industrial complex analysis is the first approach to grouping firms into clusters that do not necessarily belong to the same or similar industries. What Isard et al. loosely referred to as "combinations of industrial activities for a region" (p. 1) became the foundation for cluster analysis based on production value chain linkages derived from input-output tables. Intensive research on the use of input-output tables for identifying industrial clusters was done in the 1970's and 1980's by scholars like Czamanski (1974, 1976), Roepke (1974), and Ó hUallacháin (1984). Input-output tables continue to be widely used in industrial cluster analysis to this day. However, scholars like Feser and Bergman (1999, 2000) have raised serious concerns regarding the appropriateness of sub-national input-output tables because of the exclusion of non-local buying / selling patterns (i.e., domestic exports / imports). Furthermore, despite the long tradition of using input-output tables for cluster analysis, there is still a lack of consensus on two important issues: the choice of an appropriate method for identifying clusters and the question of whether or not these identified clusters really capture industries that are located next to each other.

One appropriate way of identifying industrial clusters using input-output tables is by means of principal components factor analysis (PCA). The goal of PCA is to reduce the number of correlated variables in the dataset to a smaller number of meaningful dimensions or factors. As a data reduction method, PCA reduces the number of industries into a smaller number of industrial clusters using the maximum common variance criteria between industries and clusters (Tinsley and Tinsley 1987). Although straightforward conceptually, there are at least three different ways of identifying clusters using this tool. The first approach involves applying PCA to the inter-industry transaction matrix to group industries based on similarities in their buying patterns (i.e., R-mode analysis). The second applies PCA to the transposed transaction table to account for similarities in selling patterns (i.e., Q-mode analysis). And the third involves applying it to a symmetric matrix of correlation coefficients derived from correlating the normalized transaction and transposed transaction matrices with each other. This last approach, which matches industry pairs where one industry's buying pattern is similar to the other industry's selling pattern, accounts to some extent for vertical value chain linkages. But regardless of which of the three methods is given preference, all I-O based PCA approaches have two important limitations. First, they cannot actually measure the extent to which the identified industries co-locate in geographic proximity. And second, grouping industries into clusters according to similarities in buying and selling patterns does not provide us with the information on which key industries policy makers should focus to get the largest return for their investment.

The issue of geographic proximity in cluster definition has repeatedly been raised in the relevant literature. For instance, Czamanski and Ablas (1979) explicitly distinguish between industrial clusters and industrial complexes based on whether or not the industry groups are spatial conglomerations. According to them, a cluster is a set of industries connected by flows of goods and services, while a complex has the additional characteristic of industrial concentration in a well defined location. Roepke et al. (1974) define an industrial cluster as: "…. a base group of industries that have similar patterns of transactions, and it also includes other industries which are major suppliers or markets for those within the group" (p. 15). Accordingly, they perform an R-mode and a Q-mode analysis on a highly aggregated (44x44) matrix of inter-industry linkages for the Province of Ontario and identify thirteen groups of industries. Using the Streit Index to evaluate the intensity of inter-industry flows within identified industry groups, Roepke et al. conclude that they have identified industrial complexes. However, Czamanski and Ablas (1979) argue that Roepke et al. identified industrial clusters rather than complexes as only *aspatial* information from a very highly aggregated input-output table was used.

While acknowledging the above distinction between clusters and complexes, Ó hUallacháin (1984) stresses that the distinction might not be as important in the regional context. More specifically, he argues that when regional input-output tables are used for regional cluster analysis, "users of regional input-output tables view functional clusters and spatial complexes as identical phenomena because the data are confined to a single region" (p. 422). As for the cluster identification approach, Ó hUallacháin (1984) uses PCA to group industry sectors according to their complementary relationships, i.e., similarities in their buying / selling patterns, and demonstrates on a (49x49) input-output table of Washington State that Chenery and Watanabe's (1958) backward and forward linkage indices, are sufficient to identify industrial complexes.

Czamanski (1974, 1976) takes a different approach to addressing concerns raised by the aspatial nature of I-O tables in the identification of industrial clusters. In his 1974 study, he focuses on detecting vertical relationships in a highly disaggregated (172x172) input-output table for the US economy by applying PCA to a symmetric intercorrelation matrix containing the largest correlation coefficients, the outcomes of correlating the normalized transaction table with the normalized transposed transaction table. Altogether, Czamanski identifies sixteen industrial clusters which he describes as *purely aspatial*. Then in his 1976 study, he uses population and employment data for 191 Standard Metropolitan Statistical Areas (SMSA) to assess whether or not the identified clusters fulfill the spatial 'geographic proximity' characteristic of industrial complexes. Regressing first employment by industry *i* for each region *k* (E_{ik}) on the region's population (P_k) and then regressing the error terms of the first regression (ε_{ik}) on employment in industry j in region k (E_{jk}), Czamanski again derives sixteen groups of industries. Spatial industrial clusters (or industrial complexes) are then identified by analyzing the similarity between these two independently derived industry groups.

The second limitation of PCA—the provision of inadequate information for identifying key sectors in the economy—can be addressed in a straightforward manner by using PCA results in conjunction with intersectoral relationship indicators derived from I-O tables. Perhaps, the simplest and most rudimentary of these linkage measures are the backward (BL) and forward (FL) linkage measures by Chenery and Watanabe (1958). Derived from the direct coefficient matrix, the Chenery and Watanabe backward and forward linkage indices measure only the direct effects of changes in the final demand for the different sectors. In other words, they only measure total intermediate purchases and total intermediate sales. It is, however, possible to account for the indirect effects as well by using partial multipliers from the

Leontief inverse matrix (i.e., the total requirements matrix) instead.

Hazari (1970), Hewings (1974), and Beyers (1976) were among the earliest to empirically identify key sectors by including both direct and indirect effects. Conceptually very similar, their approaches all focus on identifying key sectors by using two indices introduced by Rasmussen (1952), namely, the index of the power of dispersion and the index of the sensitivity of dispersion. But while Hazari (1970) and Hewings (1976) rely exclusively on the demand-driven Leontief inverse matrix to construct both indices, Beyers (1976) uses an inverse matrix derived from a supply-driven I-O model to derive the index of the sensitivity of dispersion. Beyers' inverse matrix is based on the sales coefficient matrix, where sectoral gross output relates to the primary inputs of production rather than to final demand as is done in the standard I-O model. We follow Beyer's approach in this paper.

Feser and Bergman (2000) bring to the fore another limitation of most I-O based cluster-identification approaches, namely their inability to account for the linkages of regional industries with the national economy. Defining clusters as: '…. specific constellations of linked firms' (p. 3), they argue that sub-national input-output tables are too restrictive due to the absence of non-local buying/selling patterns and use a (362x362) sector national-level input-output table to identify industrial clusters. The current study also attempts to address this limitation of I-O based approaches.

THE RATIONALE FOR PRESENTED RESEARCH

The literature reviewed above makes it clear that any sound methodology for identifying regional industrial clusters must account for the linkages among the industries, their similarities in purchase and sales patterns, and their spatial proximity to each other. In addition, and as indicated by Feser and Bergman (2000) and Feser and Luger (2002), the methodology must also recognize that while some clusters are local to a region, others can cross regional boundaries. Existing cluster identification methodologies, however, focus either on purely local clusters or identify national/supra-regional clusters that are aspatial in nature, and fail to identify supra-regional clusters that have a significant presence in the local economy. Our research attempts to address this shortcoming in the literature by proposing a methodology that identifies both local and domestic industrial clusters in the region. The main methodological innovation in this study is the integration of commodity imports and exports by industry into the regional I-O framework (in our case the I-O table for the Chicago MSA). We show that this extended I-O framework allows us to identify the two types of clusters mentioned above.

From a policy perspective, the rationale for studying both local and domestic clusters is that they provide different insights into the strengths and growth potential of the local economy. Local clusters, by grouping those industries in the region that currently exhibit significant links with each other, identify the sectors that are already in line with cluster-oriented development. They do not, however, help to identify other complementary industries that might be targeted by cluster-oriented development strategies in the future to enhance the growth of the local economy. Furthermore, since local clusters exclude industries that have especially strong non-regional trading relationships, some key local sectors could be ignored if policymakers focus exclusively on identified local clusters. Latham (1976) and Feser and Bergman (2000) suggest that we should, therefore, also examine clusters based on national I-O tables to develop economic development strategies that recognize complementary industries and existing gaps in supply-chains and product chains. More specifically, Feser and Bergmen (2000) propose using the national clusters as "templates for developing a strategic view of a regional manufacturing economy" (pp. 2). Using national clusters to identify complementary industries and gaps in chains as recommended by Latham (1976) and Feser and Bergman (2000), however, fails to account for the unique economic characteristics and competitiveness of the regional economy. Thus targeting industrial sectors on the basis of national clusters might not necessarily be an appropriate regional development strategy.

Like national clusters, the domestic clusters identified by our methodology can reveal complementary industries, supply-chain gaps and product-chain gaps that are not identified by local clusters. But since domestic clusters focus on local industries that exist as members of supra-regional clusters, they also account for the regional characteristics of these industries. Hence the identification of domestic clusters can give additional insights that could be useful for devising appropriate regional development strategies that build upon the existing competitive advantage of local industries. For example, if a set of industries is found to belong to a domestic (but not local) cluster, then we know that these industries have strong links among themselves via their import-export relationship with the national economy. Thus one potential economic development strategy in this case would be to assist them in reorienting their inter-industry purchases patterns so that input imports are progressively replaced by local purchases.

Any cluster-based development strategy must recognize that different industrial clusters have very dissimilar requirements regarding capital, labor, infrastructure, and other local factors that affect the overall business environment in which they operate. At the same time, the above discussion indicates that the spatial attributes of a region's clusters (i.e., whether they are local, domestic or national) are equally important factors in designing a policy portfolio aimed at improving the competitiveness of local clusters. The comparison of Chicago's local and domestic clusters presented in this study should give further insights into the relevance of distinguishing between these two cluster types.

STUDY REGION AND DATA

The study area, hereafter referred to as the Chicago region, includes the following eight most urbanized counties from the Chicago-Naperville-Joliet Metropolitan Statistical Area (MSA) (Census 2006): Cook, DuPage, Kane, Lake, McHenry, and Will in Illinois, and Lake and Porter in Indiana. Combined, these eight counties account for 96.6 percent of the employment and 96.2 percent of all establishments in the MSA (Bureau of Economic Analysis 2003). The main data framework is the 2001 Chicago I-O table which has been derived using the commercially available IMPLAN® economic impact modeling system (MIG 2004). The I-O framework for the Chicago region contains all the information required for performing a principal components factor analysis of inter-industry transactions. In addition, the IMPLAN dataset also includes employment and value-added data for 2001. Supplementary data on employment and establishments are taken from the Quarterly Census of Employment and Wages (QCEW) of the Bureau of Labor Statistics (BLS 2006). The industry classification in the presented framework follows the North American Industry Classification System (NAICS). To achieve a sufficient level of industry detail, we distinguish among a total of 223 individual industries in the I-O framework: eight extraction sectors (i.e., agriculture and mining), 119 manufacturing industries (including construction), and 96 service industries.

For the purpose of the presented analysis, we rely on two data sets extracted from the IMPLAN database for the Chicago MSA: i) the intra-regional inter-industry flows as captured in the use and make matrices and ii) the inter-regional flows which map commodity exports and imports by industry. The Minnesota IMPLAN Group, Inc. (MIG) relies on information from various national, state, and county level data sources to construct the regional IMPLAN database.[4] For instance county-level data on employment, employee compensation, proprietary income, population, federal and state expenditures and selected wealth data are employed for the estimation of county-level databases. These data sources, together with national I-O matrices and national tables for regional purchase coefficients (PPC), are the main ingredients for building sub-national level I-O accounts.

IMPLAN I-O tables are built top-down using a non-survey approach.[5] Thus, the national level structural matrices (i.e., use and make matrices) form the basis for the regional level use and make matrices. Adjustments are made to the national structural matrices to reflect regional differences. State level data are appropriately scaled to ensure that they sum to the national totals. Similarly, county level data within each state are adjusted so that their sums match the state totals.

Regional trade flows (commodity imports and exports) are estimated based on regional purchase coefficients (RPC) which measure the portions of total commodity demand that are supplied by regional sources. The starting point for deriving the RPCs are empirical trade flow data taken from the U.S. Department of Health and Human Services' 1977 multi-regional I-O accounts (MRIOA), a 125 sector inter-regional I-O framework for all states and the District of Columbia. Using an econometric model, the RPCs are estimated based on regional employee compensation by industry, regional employment ratios, location quotients, and area ratios. It should be pointed out that, due to data limitations, IMPLAN distributes foreign imports proportionally among all regions within the U.S. We have, therefore, excluded the foreign trade flows from our analysis.

[4] A detailed technical discussion of the data methodology employed by Minnesota IMPLAN Group can be found in MIG (2004).

[5] Ideally, we would like to use I-O tables based on data from current surveys. But building survey-based regional I-O tables is an undertaking where the amount of time and money spent does not justify the gains in accuracy.

METHODOLOGY

As indicated earlier, this paper seeks to identify two types of industrial clusters—local and domestic—based on inter-industry linkages. In both cases, the cluster identification methodologies used are similar to the principal component analysis (PCA) approaches in Czamanski (1974), Roepke et al. (1974), and Feser and Bergman (2000). Both utilize information from I-O tables for the local economy to group industries into clusters based on inter-industry purchase and sales linkages, similarities in inter-industry trade patterns, or similarities in markets for resources and outputs. And in each case, the identified clusters can be considered spatial industry agglomerations since the geographical area under consideration is relatively small. But while identification of local clusters requires data only on commodity flows across industries within the region, identification of domestic clusters requires information on inter-regional flows as well. Thus the incorporation of inter-regional inter-industry flows in the identification of domestic clusters distinguishes our methodology from others found in the literature.

Identification of local industrial clusters

The first step in identifying local industrial clusters is the construction of an I-O inter-industry transaction table for the region under consideration. We derive the inter-industry transaction table from the commodity flow tables—the make $(\mathbf{M})$ and use $(\mathbf{U})$ matrices—that are part of the I-O framework for the region. Each element m_{ij} of $\mathbf{M}$ represents the value of commodity j made by industry i, and each element u_{ij} of $\mathbf{U}$ shows the value of commodity i used by industry j. The local inter-industry transaction table $(\mathbf{T})$ is then constructed as follows:

$$\mathbf{T} = \mathbf{M}\mathbf{C}^{-1}\mathbf{U} \qquad (1)$$

where $\mathbf{C}$ is a diagonal matrix whose diagonal elements represent the total values of the different commodities produced by the industries in the region for the regional market. Each element T_{ij} of $\mathbf{T}$ denotes the quantity (in dollar terms) of industry i's output purchased by industry j for intermediate use.

The next step involves performing PCAs based on $\mathbf{T}$ to identify industries that exhibit similarities in input purchase patterns or similarities in selling patterns.[6] If we denote the total intermediate sales of industry i by $T_{i\bullet}$, we can define a local purchase coefficient $a_{ij} = \frac{T_{ij}}{T_{j\bullet}}$, and a local sales coefficient $b_{ij} = \frac{T_{ij}}{T_{i\bullet}}$. Thus each column $\mathbf{a}_j$ of the matrix $\mathbf{A} = [a_{ij}]$ shows the proportions of total inputs bought by industry j from different industries while each row $\mathbf{b}_j$ of $\mathbf{B} = [b_{ij}]$ represents the proportions of industry $j's$ sales going to different industries.

The PCA of $\mathbf{A}$ with varimax rotation gives industry clusters that are based on similarities in input purchase patterns among the industries. It is performed by treating each industry column $\mathbf{a}_j$ as a variable and deriving a smaller number of latent factors which explain observed correlations between these industries using the maximum common variance criteria between variables and a factor. The strength of the relationship between any industry and a particular factor is given by the corresponding factor loading (Feser and Bergman 2000).[7] In this paper, industries

[6] These are the R-mode and Q-mode analyses referred to in Section 2.

[7] Since the literature provides numerous examples of how to apply principal components factor analysis to

with loadings of at least 0.35 or greater on a given factor are considered members of a cluster. And while an industry can belong to more than one cluster, the industry with the highest factor loading in a particular cluster is considered the primary industry of that cluster. Industry clusters based on similarities in selling patterns are identified by performing a PCA of $\mathbf{B}$.

It is clear that the clusters identified using the above PCAs do not consider the sales and purchase linkages between industry pairs, but focus on similarities in sales and purchase patterns. One way to account for these value-chain linkages is by performing a PCA on a new matrix that captures the correlations in I-O structures between pairs of industries. The similarities in the I-O structures between industry pairs can be described by the following correlations based on $\mathbf{A}$ and $\mathbf{B}$:

$Corr(\mathbf{a}_{i,}\mathbf{a}_j)$: correlation between industries i and j in terms of purchase patterns

$Corr(\mathbf{b}_{i,}\mathbf{b}_j)$: correlation between industries i and j in terms of sales patterns

$Corr(\mathbf{a}_{i,}\mathbf{b}_j)$: correlation between the purchase pattern of i and sales pattern of j

$Corr(\mathbf{b}_{i,}\mathbf{a}_j)$: correlation between the sales pattern of i and purchase pattern of j.

Note that the last two correlations show the extent to which sectors that buy the output of one industry supply inputs to the other industry (Czamanski and Ablas 1979). In other words, they capture the indirect purchase-sales relationships between industry pairs.

In the third step, a new symmetric correlation matrix $\mathbf{S}$ is constructed by selecting the largest of the above four correlation coefficients for each industry pair, and a PCA is performed on this matrix to identify industry clusters. These clusters thus take into account indirect purchase-sales relationships between industry pairs as well as similarities in sales and purchase patterns.

Identification of domestic industrial clusters

Recall that a domestic industrial cluster is a group of linked local industries which might also belong to larger clusters that cross region boundaries. The method for identifying domestic industrial clusters is the same as the approach used for identifying local clusters except for a modification in the first step. The transaction matrix $\mathbf{T}$ now incorporates not just intra-regional inter-industry flows, but also inter-regional flows recorded in the exports matrix ($\mathbf{E}$) and the imports matrix ($\mathbf{I}$) of the I-O framework.

Each element E_{ij} of the exports matrix shows the total export of commodity j made by industry i to destinations within the US. Similarly, each element I_{ij} of the imports matrix represents the amount of commodity i imported from within the US for use as inputs to industry j. These export/import matrices are used along with the make and use matrices to derive the following domestic inter-industry transaction matrix:

$$\mathbf{T} = \left(\mathbf{M}+\mathbf{E}\right)\times\mathbf{C}^{-1}\times\left(\mathbf{U}+\mathbf{I}\right)$$

where the elements of the diagonal matrix $\mathbf{C}$ represent the total production of each commodity in the region. The domestic transaction matrix $\mathbf{T}$ given by Equation (2) thus accounts for both local and non-local

inter-industry transaction matrices, we do not provide a more detailed description of PCA here. Instead, we refer the interested reader to Czamanski (1974), Bergman and Feser (1999), and Tinsley and Tinsley (1987).

commodity trade activities of the industries in the region. Domestic industrial clusters, which reflect the position of Chicago's industries in the larger national economy, are then identified by performing PCAs based on $\mathbf{T}$.

Remarks on Principal Components Analysis (PCA)

We prefer PCA to other data reduction methods, such as factor analysis, mainly because it is the method of choice for descriptive and exploratory research where the primary focus is on summarizing data sets.[8] Furthermore, PCA allows an industry sector to be a member of more than one cluster. The goal of PCA is to transform the set of original variables to a more manageable set comprising of only a few factors.

In order to identify clusters, we use the Kaiser criterion—a criterion based on the eigenvalues of each factor. Eigenvalues in this context explain how much each factor contributes to explaining the common variance and is calculated as the sum of all squared factor loadings for a factor. More specifically, all factors with eigenvalues greater than 1.0 are initially selected as candidate clusters. If, however, a selected factor with an eigenvalue close to 1.0 is composed of seemingly unrelated industry sectors, then it is dropped from the final pool of clusters. For better interpretability of the PCA solution without changing the underlying mathematical principles, the initial factors are rotated using an orthogonal Varimax rotation.[9]

Sectors affiliated with each cluster are identified using their corresponding factor loadings, which represent the correlation of each sector with that cluster. In this paper, we use a factor loading value of 0.35 as the cut-off value for identifying cluster affiliation. Strong cluster affiliation is indicated by loadings of 0.7-1.0, median cluster affiliation by loadings of 0.5-0.7 and weak cluster affiliation by loadings of 0.35-0.5.

Though PCA is a valuable tool for identifying clusters using I-O tables, it does have a number of limitations. Most importantly, it does not use any external criterion that would allow us to test the goodness of received outcomes. For instance, in regression analysis, we can check the correlation between observed and predicted dependent variables to examine the interrelationship among a given set of variables (Tabachnick and Fidell, pp. 608). A second problem is the difficulty associated with choosing the rotation method. Though all rotations account for the same amount of variance in the final solution, the received factors may vary slightly. Third, nothing can be said about the importance of cluster members for regional economies when factor loadings are used as the sole cluster affiliation criteria. This limitation creates difficulties in making meaningful and policy-relevant interpretations of the PCA results. In other words, since factor loadings do not necessarily reflect inter-industry connectivity, industry employment and output, the ranking of sectors within a cluster based on factor loadings alone might not be meaningful from an economic perspective. Last, the ordering of derived industry clusters based on eigenvalues does not necessarily imply that clusters with higher eigenvalues are more important to regional economies. Despite the fact that higher eigenvalues usually associate with larger numbers of cluster members, nothing can be said about the position of a cluster within a regional economy. To improve the

---

[8] According to Tabachnick and Fidell (2007): *Principal components analysis is an empirical approach, whereas factor analysis and structural equation modeling tend to be theoretical approaches* (pp. 25).
[9] An orthogonal rotation has been given preference as it yields independent and uncorrelated factors (Tinsley and Tinsley, 1987).

interpretation of selected clusters and cluster members, we therefore do a key industry sector identification exercise as described below.

Key industry sector identification[10]

Once we have derived the final set of industrial clusters based on the **A** and **B** matrices, the last step involves rank ordering the industry sectors following the key sector criteria proposed by Beyers (1976). As indicated earlier, Beyer's approach involves deriving the index of the power of dispersion and the index of the sensitivity of dispersion.

The index of the power of dispersion identifies the sectors that have a greater than average degree of backward connectivity with the economy. Using the Leontief inverse matrix, $\left(\mathbf{I}-\mathbf{A}\right)^{-1}$, the index of dispersion for sector j (U_j) can be directly calculated as:

$$U_j = \frac{\frac{1}{m}\sum_i \alpha_{ij}}{\frac{1}{m^2}\sum_j \sum_i \alpha_{ij}} \qquad (j = 1, 2, \ldots, m) \tag{3}$$

where α_{ij} is the $(i, j)^{th}$ element of $\left(\mathbf{I}-\mathbf{A}\right)^{-1}$, and m denotes the number of rows in the matrix. Conceptually rooted in the Chenery and Watanabe (1958) linkage measures, U_j can be described as the normalized version of the average output multiplier for sector j. The numerator in Equation (3) represents what Hazari (1970) describes as the average output multiplier, while the denominator takes care of the normalization to allow for inter-industry comparisons. A sector with greater than average backward linkage is, therefore, identified by a U_j value greater than unity.

Similarly, the index of the sensitivity of dispersion for sector i (U_i) can be directly calculated from the elements β_{ij} of the forward-linkage inverse matrix $\left(\mathbf{I}-\mathbf{B}\right)^{-1}$ as:

$$U_i = \frac{\frac{1}{m}\sum_j \beta_{ij}}{\frac{1}{m^2}\sum_i \sum_j \beta_{ij}} \qquad (i = 1, 2, \ldots, m)$$

Indices of the sensitivity of dispersion can be interpreted as the normalized versions of average supply multipliers.[11] As before, a cutoff value of unity is used to identify industry sectors with greater than average forward linkages.

We use the above two indices along with industry-specific employment information to rank order the component sectors of each cluster according to their economic importance.[12] The ranking involves grouping the sectors in the cluster into three categories: i) strongly linked key sectors consisting of components that have above average backward as well as forward linkages (i.e., both linkage indices are greater than 1.0), ii) moderately linked sectors which have above average links with the economy in only one direction (i.e., only one index is greater than 1.0), and iii) weakly linked sectors whose linkage indices are below average in both directions. Within each category, cluster components are

[10] This step is common to both local and domestic clusters.

[11] For example, the supply multiplier for sector i is the change in total output in the whole economy resulting from a dollar increase in the availability of value added inputs for that sector.

[12] Note that the PCA factor loadings are also used in some approaches to rank component sectors. But while the factor loading for each sector does show the extent to which the sector is correlated with the cluster, it does not convey any information regarding the economic interactions between the sector and the cluster.

additionally ranked according to employment levels.

Apart from identifying key sectors, this ranking also allows us to identify sectors that have very weak links with the cluster and are also relatively unimportant in terms of employment. Such sectors are potential candidates for exclusion from the cluster. Some of these candidate sectors are removed from the cluster by utilizing a combination of subjective a priori knowledge about the linkages in the local economy and the sector ranking information. Thus the revision of the PCA clusters using the key sector identification process enables us to address, to some extent, the problem of seemingly unrelated industry sectors within clusters. [13]

FINDINGS AND RESULTS

In this section, we present results that highlight the similarities and differences between domestic and local industrial clusters identified by the various PCAs. Starting with an overview of the list of domestic and local clusters, we proceed to discuss industrial clusters that emerged in both frameworks – the common clusters. In logical sequence we then discuss those clusters that showed up in only one of the two frameworks presented – the unique and the analogous clusters. Then, we present a detailed discussion of a few key clusters identified in the two (local and domestic) frameworks.

Overview of identified clusters

Altogether, we ran three different variants of principal components factor analysis (i.e., R-mode, Q-mode, and symmetric correlation matrix) on each of the two frameworks (local and domestic). The PCA results from these three types of analyses were used to derive the initial set of local and domestic industrial clusters. A summary of the PCA results is presented in Table 1 below.

As shown in table 1, the six individual principal components factor analyses results differ in terms of the number of identified factors. For both the local and domestic frameworks, the R-mode and Q-mode analyses have identified slightly more factors than the PCAs based on the symmetric correlation matrices. For instance, the total numbers of factors in the R-mode and Q-mode analyses are 24 and 30 respectively in the local framework, and 28 and 29 respectively in the domestic framework. On the other hand, the corresponding number of factors derived from the symmetric correlation matrix is only 16 in the local framework and 15 in the domestic framework.

While the factors identified by the PCAs represent specific industry groupings, not all such factors constitute valid industrial clusters. Hence, as a first step in the cluster identification process, we attempted to eliminate irrelevant factors by assigning the identified factors to three different groups. The first group, which we refer to as "single-industry factors", consists of factors that include three or fewer industries. As such factors contain too few industries to be considered valid clusters; they are not included in the final list of industrial clusters. The second group, labeled "undefined factors" in Table 1, includes factors that appear to be collections of a random mix of industries. Hence, these factors too are not considered valid industrial clusters. Only the remaining factors are used for deriving the final set of industrial clusters. This last group of factors is refereed to as "significant factors" in Table 1.

[13] We would like to thank an anonymous reviewer for suggesting this approach to addressing the issue of spurious cluster affiliation.

Table 1: Summary results of six variants of principal components factor analysis

Variant	Number of significant factors	Number of single-industry factors	Number of undefined factors
Local Framework			
R-Mode	18	3	3
Q-Mode	20	7	3
Sym. Correlation Matrix	13	1	2
Domestic Framework			
R-Mode	20	5	3
Q-Mode	21	4	4
Sym. Correlation Matrix	12	1	2

Source: Author's tabulation

Taking a closer look at the significant factors, we observed that some of them consisted of a distinctly large number of component industries.[14] More specifically, the PCA on the symmetric correlation matrix and the transaction matrix (R-mode) resulted in two such mega-factors: one manufacturing and one service. For instance, the PCA on the symmetric correlation matrix using the local framework results in one manufacturing factor composed of as many as 144 component industries of which 122 are manufacturing sectors and 22 are service sectors. A second mega-factor is composed of 68 component industries: 63 service sectors and five manufacturing sectors. Thus these two mega-factors combine as many as 212 of the 223 industry sectors considered in this study.[15] The same patterns hold for the R-mode analysis for both the local and domestic frameworks. Although these two identified mega-factors make perfect sense, we decided to exclude them from the final pool of relevant clusters since apart from classifying industries as manufacturing or service sectors, they do not provide much information that might be valuable for potential cluster-based economic development policies. Having decided which information, i.e., factors, to use from the original six sets of PCA outcomes, in the second step we put together two final sets of industrial clusters—one using the local framework and the other using the domestic framework. These two sets of industrial clusters are shown in Table 2. This table also shows, for each cluster, the total number of component industries identified by the PCA as well as the revisions made using the key sector identification method discussed earlier.

[14] The National Industry Cluster Template study by Feser and Bergman (2000) also highlights the fact that the number of component sectors varies a lot among the clusters; from 116 in the metalworking cluster to four in the tobacco cluster (pp. 6).
[15] These are the number of cluster components that were identified before making revisions using results from the key industry sector analysis.

Table 2: Local versus Domestic Industrial Clusters

I. Local Framework

No.	Cluster Description	Number of Original Cluster Components	Number of Revised Cluster Components	Total Revised Cluster Employment	Percent Change in Cluster Components
1	Retail	61	47	1,314,698	23.0
2	Computer-based Support	23	20	937,648	13.0
3	Construction incl. Building Servies	18	15	683,771	16.7
4	Personal Services	8	7	572,799	12.5
5	Financial Services	14	10	444,196	28.6
6	Transportation	14	14	304,841	0.0
7	Insurance	4	4	251,587	0.0
8	Metalworking	26	22	177,982	15.4
9	Telecom	6	5	142,017	16.7
10	Automotive	10	8	127,169	20.0
11	Food	26	26	112,383	0.0
12	Broadcasting-Arts-Sports	11	8	99,462	27.3
13	Electronics	11	9	70,108	18.2
14	Nonmetallic Mineral Products	12	9	62,017	25.0
15	Metal Foundries	7	7	57,486	0.0
16	Chemical-Plastic-Rubber	15	13	29,575	13.3
17	Printing	4	4	17,288	0.0
18	Textile	5	5	10,444	0.0

II. Domestic Framework

No.	Cluster Description	Number of Original Cluster Components	Number of Revised Cluster Components	Total Revised Cluster Employment	Percent Change in Cluster Components
1	Retail	78	49	1,910,307	37.2
2	Computer-based Support	33	28	1,101,432	15.2
3	Construction incl. Building Servies	23	20	748,574	13.0
4	Professional Services	10	8	483,666	20.0
5	Financial Services	12	9	439,258	25.0
6	Metalworking I	36	26	431,734	27.8
7	Travel and Sightseeing	3*	3*	338,133	0.0
8	Automotive	11	8	167,804	27.3
9	Transportation	6	6	159,042	0.0
10	Food	24	23	86,509	4.2
11	Nonmetallic Mineral Products	11	11	72,901	0.0
12	Electronics	9	9	70,369	0.0
13	Paper	11	11	69,545	0.0
14	Metalworking II	10	10	61,729	0.0
15	Plastic	12	12	61,435	0.0
16	Broadcasting-Arts-Sports	12	8	46,758	33.3
17	Wood Products	10	9	35,352	10.0
18	Chemical and Rubber	9	9	23,608	0.0
19	Utilities-Insurance	8	5	17,038	37.5
20	Textile	5	5	6,638	0.0

Note: A total of eleven clusters were common to both frameworks.

Source: Implan Pro and Bureau of Labor Statistics.

Table 3: Exclusive Clusters Comparison

I. Local Framework: Cluster Description	Number of Cluster Components	Total Cluster Employment	II. Domestic Framework: Cluster Description	Number of Cluster Components	Total Cluster Employment
a) Unique Clusters			a) Unique Clusters		
Personal Services	7	572,799	Professional Services	8	483,666
Telecom	5	142,017	Travel and Sightseeing	3*	338,133
Printing	4	17,288	Paper	11	69,545
			Plastic	12	61,435
			Wood Products	9	35,352
b) Analogous Clusters			b) Analogous Clusters		
Metalworking	22	177,982	Metalworking I	26	431,734
Metal Foundries	7	57,486	Metalworking II	10	61,729
Chemical-Plastic-Rubber	13	29,575	Chemical and Rubber	9	23,608
Insurance	4	251,587	Utilities-Insurance	5	17,038

Source: Implan Pro and Bureau of Labor Statistics.

Table 3 lists the Personal Services cluster as a unique <u>local cluster</u> with a strong focus on providing services locally only to individuals, households, and businesses. No equivalent cluster has been identified in the domestic framework. The seven components of this cluster include locally oriented sectors such as Health Care Services (621), Social Assistant including Child Care Service (624), Personal Care Services (8121) and Photographic Services (54192) [17]. Of particular regional economic importance is Health Care Services, which, apart from having the largest employment in the cluster, also has strong ties to the regional economy with above average backward linkage (U_j=1.06) and strong forward linkage (U_i=1.12). The other two local unique clusters—Telecom and Printing—too focus strongly on local business support. In addition to supporting local businesses, Telecom—which includes Computer and Telephone Manufacturing (3341-2), Telecommunications (5133), Management of Companies and Enterprises (55), Data Processing Services (5142) and Information Services (5141)—illustrates a characteristic common to many clusters with mixed manufacturing and non-manufacturing components, namely the distinctly stronger presence of manufacturing sectors among the key industries. For instance, the only Telecom sector that qualifies as a key industry is Computer and Telephone Manufacturing. None of the service sectors qualifies as a key sector. Furthermore, among the service sectors in the cluster, only Telecommunications has one above-average index (U_i=1.05); the forward and backward linkage indices for all the other sectors are smaller than 1.0.

The <u>domestic framework</u> shows five unique industrial clusters, namely Professional Services, Travel and Sightseeing, Wood Products, Paper, and Plastic, all of which have a non-local market base. The Paper cluster, in particular, is a good example of an industrial cluster with significant cross-boundary inter-industry linkages. Its component sectors such as Forestry and Logging (113), Pulp, Paper, and Paperboard Mills (32211, 32212, 32213), and Paperboard Container Manufacturing (32221) have large

[17] North American Industry Classification System (NAICS) codes in parenthesis.

sensitivity of dispersion indices (U_i). This suggests that these sectors have great export potential since a part of the supply chain of domestic clusters is located outside the study region. Professional Services is another unique cluster that deserves special attention. With an employment level of 484 thousand, it is clearly an important cluster for the Chicago economy. Its significance to the Chicago economy is also indicated by the fact that four of its eight component industries—Real Estate (531), Legal Services (5411), Architectural and Engineering Services (5413), and Environmental and Other Technical Consulting (54162, 54169)—meet the key sector criteria.

The second group of industrial clusters in Table 3, the *analogous clusters*, can best be described as hybrid clusters in that while they do appear in both local and domestic frameworks, they nevertheless differ significantly in the two frameworks in terms of the number of cluster components and total cluster employment. Take, for example, the clusters related to the metal and steel industries. Two individual metal and steel-related clusters were identified in each framework. But although Metalworking in the local framework is somewhat similar to Metalworking I in the domestic framework in that there are 13 sectors common to both clusters, the two clusters differ significantly in other respects. Not only are the remaining cluster components and the total number of components different for the two clusters, but their employment figures are also quite dissimilar. Similar observations can be made about the other analogous clusters as well.

Of further interest to the policy analysts in this subgroup of *analogous clusters* might be the fact that the local framework merges chemical, plastic, and rubber products into one Chemical-Plastic-Rubber cluster, while the domestic framework identified two distinct clusters, namely, Chemical and Rubber and Plastic. This example shows that adding the non-regional trade data to the local input-output framework helps us to identify more clear-cut clusters, particularly when grouping industries with smaller employment numbers.

Details of specific clusters

In order to better understand the differences between the domestic and local clusters, we now present in the remainder of this section a detailed discussion of three specific clusters—Retail, Financial Services and Food Manufacturing. The Retail cluster is, by far, the largest cluster in both frameworks not only in terms of the number of components, but also when considering total cluster employment. Furthermore, unlike all other clusters, Retail includes a significant number of both service and manufacturing industries. It is, therefore, interesting to evaluate how this cluster differs in the two frameworks and what the implications for economic development purposes are. As shown in Table 2, the local Retail cluster consists of 47 industries and employs 1,314,698 people. On the other hand, the corresponding cluster in the domestic context includes 49 industries and provides employment for 1,910,307 individuals. Details on the components of this cluster are presented in Table 4.

Table 4: Retail Cluster Comparison

Local Framework

Retail	**Employment**	**Factor Loadings**	**Power of Dispersion (U_j)**	**Sensitivity of Dispersion (U_i)**
Health care services	383,916	0.383	1.06	1.12
Professional services	98,484	0.470	1.01	1.07
Food and beverage stores	86,626	0.971	1.01	1.09
Nonstore retailers	55,217	0.971	1.08	1.13
Miscellaneous store retailers	42,291	0.971	1.11	1.15
Sporting goods, hobby, book and music stores	28,172	0.971	1.06	1.08
Other Plastics Product Manufacturing	24,123	0.565	1.07	1.21
Gasoline stations	13,732	0.971	1.04	1.02
Electrical Equipment Manufacturing	11,616	0.443	1.05	1.13
Other Electrical Equipment Mfg	10,047	0.674	1.01	1.11
Commercial and Service Industry Machinery	6,059	0.480	1.06	1.08
All Other General Purpose Machinery Mfg	6,018	0.420	1.00	1.06
Cutlery and Handtool Manufacturing	4,704	0.521	1.00	1.02
Material Handling Equipment Mfg	4,434	0.394	1.05	1.09
HVAC and Commercial Refrigeration Equipment	3,839	0.764	1.00	1.04
Electric Lighting Equipment Mfg	3,495	0.638	1.06	1.02
Paint and Coating Manufacturing	3,263	0.866	1.06	1.23
Asphalt Paving, Roofing Materials and Other Petrol	2,591	0.854	1.35	1.28
Ag, Construction, and Mining Machinery Mfg	2,574	0.625	1.08	1.12
Plastics Pipe, Fittings, and Profile Shapes	2,188	0.557	1.00	1.01
Hardware Manufacturing	2,019	0.659	1.02	1.04
Other Textile Product Mills	2,005	0.384	1.03	1.06
Pulp, Paper, and Paperboard Mills	1,365	0.718	1.12	1.19
Adhesive Manufacturing	1,093	0.468	1.11	1.08
Social assistant, incl. child day care services	91,205	0.522	0.99	1.00
General merchandise stores	76,705	0.971	0.95	1.01
Clothing and clothing accessories stores	51,083	0.971	0.97	1.02
Motor vehicle and parts dealers	48,935	0.971	0.96	1.03
Furniture and home furnishings stores	20,164	0.971	0.98	1.01
All Other Fabricated Metal Products Mfg	10,971	0.596	0.99	1.08
All Other Miscellaneous Mfg	9,429	0.897	0.96	1.02
Household and Institutional Furniture Mfg	6,564	0.943	1.01	0.97
Metal Valve Manufacturing	6,423	0.846	0.96	1.03
Ornamental and Architectural Products	5,994	0.945	0.96	1.05
Machinery and equipment rental and leasing	4,995	0.362	0.97	1.00
Other Wood Product Manufacturing	3,666	0.952	0.98	1.03
Rubber Product Manufacturing	2,705	0.374	0.97	1.06
Cement and Ready-mix Concrete Mfg	2,503	0.956	0.99	1.03
Household Appliance Mfg	1,607	0.809	1.08	0.95
Lime and Gypsum Product Manufacturing	976	0.411	1.01	0.96
Abrasive Product Manufacturing	965	0.458	1.02	0.91
Clay Building Materials Manufacturing	871	0.708	1.03	0.81
Health and personal care stores	42,004	0.971	0.91	0.93
Personal care services	35,361	0.559	0.95	0.96
Warehousing and storage	33,748	0.422	0.91	0.88
Building material and garden supply stores	33,314	0.971	0.95	0.99
Electronics and appliance stores	24,640	0.971	0.87	0.85

Table 4: Retail Cluster Comparison

Domestic Framework

Retail	**Employment**	**Factor Loadings**	**Power of Dispersion (U_j)**	**Sensitivity of Dispersion (U_i)**
Wholesale trade	258,035	0.662	1.07	1.09
Management consulting services	75,299	0.401	1.02	1.21
Management of companies and enterprises	63,317	0.757	1.23	1.18
Accounting and bookkeeping services	53,558	0.437	1.21	1.20
Advertising and related services	37,501	0.592	1.29	1.21
Warehousing and storage	33,748	0.582	1.32	1.23
Business support services	30,441	0.439	1.41	1.16
Other Plastics Product Manufacturing	24,123	0.565	1.12	1.00
Other support services	18,873	0.500	1.22	1.18
Automotive equipment rental and leasing	14,313	0.515	1.07	1.17
Machinery and equipment rental and leasing	4,995	0.533	1.37	1.19
Asphalt Paving, Roofing Materials and Other Petrol	2,591	0.591	1.17	1.07
Plastics Pipe, Fittings, and Profile Shapes	2,188	0.546	1.59	1.17
Adhesive Manufacturing	1,093	0.455	1.26	1.20
Automotive repair and maintenance, except car	65,433	0.421	1.07	0.88
Drycleaning and laundry services	16,118	0.436	1.02	0.76
Electrical Equipment Manufacturing	11,616	0.468	0.86	1.08
All Other Fabricated Metal Products Mfg	10,971	0.653	0.74	1.27
Other Electrical Equipment Mfg	10,047	0.707	0.81	1.19
Metal Valve Manufacturing	6,423	0.355	0.83	1.10
Cutlery and Handtool Manufacturing	4,704	0.496	0.78	1.04
HVAC and Commercial Refrigeration Equipment	3,839	0.416	0.73	1.04
Other Wood Product Manufacturing	3,666	0.657	1.51	0.99
Pump and Compressor Manufacturing	3,594	0.722	0.75	1.14
Plate Work and Fabricated Structural Products	3,518	0.372	0.74	1.07
Paint and Coating Manufacturing	3,263	0.450	0.74	1.16
Hardware Manufacturing	2,019	0.650	0.95	1.14
Other Textile Product Mills	2,005	0.403	0.76	1.14
Doll, Toy, and Game Mfg	1,370	0.449	0.73	1.08
Lime and Gypsum Product Manufacturing	976	0.393	0.73	1.43
Clay Building Materials Manufacturing	871	0.481	0.73	1.25
Sawmills, Plywood, and Engineered Wood Mfg	584	0.727	1.58	0.96
Health care services	383,916	0.474	0.73	0.33
Professional services	98,484	0.386	0.85	0.66
Social assistant, incl. child day care services	91,205	0.475	0.73	0.32
Food and beverage stores	86,626	0.954	0.91	0.61
General merchandise stores	76,705	0.954	0.85	0.58
Nonstore retailers	55,217	0.954	0.82	0.46
Clothing and clothing accessories stores	51,083	0.954	0.82	0.50
Motor vehicle and parts dealers	48,935	0.954	0.82	0.53
Miscellaneous store retailers	42,291	0.954	0.92	0.58
Health and personal care stores	42,004	0.954	0.97	0.77
Personal care services	35,361	0.408	0.74	0.34
Building material and garden supply stores	33,314	0.954	0.99	0.74
Sporting goods, hobby, book and music stores	28,172	0.954	0.97	0.63
Electronics and appliance stores	24,640	0.954	0.93	0.84
Furniture and home furnishings stores	20,164	0.954	0.79	0.45
Gasoline stations	13,732	0.954	0.83	0.60
Office Furniture and Fixture Mfg	7,367	0.437	0.82	0.47

Source: Implan Pro and Bureau of Labor Statistics.

In terms of sectoral composition, Table 4 shows that the local Retail cluster includes 18 service sectors and 29 manufacturing industries while the domestic Retail cluster has 28 service-related industries and 21 manufacturing industries. It is also interesting to note that, in both cases, this large cluster appears to consist of a variety of seemingly unrelated industries. This finding is consistent with the observations made by Feser and Bergman (2000) in their national cluster template study, where they too identified a large cluster consisting of many unrelated components.[18] And it is not a surprising finding considering that we are using an input-output framework that includes not just the commonly used manufacturing industries but also a wide range of service industries.

With respect to the economic importance of their cluster components, there are some striking differences between the local and domestic retail clusters. The local Retail cluster has 24 service and manufacturing industries—with a total employment of 800 thousand—that meet the key industry criteria based on U_i and U_j. Among these 24 key sectors, only five are retail sectors[19], while 17 are manufacturing sectors and two belong to the services category. It is interesting to note that although the remaining 23 sectors in this cluster rank low in terms of the key sector criteria, they have strong correlations (i.e., high factor loadings) with the Retail cluster. In fact, seven of these retail sectors have factor loadings of over 0.9.[20] In the local framework, the most important Retail cluster component is Health Care Services (621, 622). Not only is this sector the largest employer, but it also has strong linkages in the economy as indicated by its large U_i and U_j values. Hence, although it has a relatively low factor loading of 0.383, its membership in the Retail cluster can be explained through the value chain. Similarly, the same logic of value chain linkages and factor loading explains the cluster membership of many of the manufacturing sectors such as Other Plastic Product Manufacturing (32619) and Electrical Equipment Manufacturing (33531).

In contrast to the local Retail cluster, the domestic Retail cluster includes only 14 industries that satisfy the key industry criteria. None of these fourteen, however, is a retail sector. Also note that while Health Care services is the biggest employer in this framework as well, it no longer ranks as a key sector. Rather, the most important key industry now is Wholesale Trade (42), with a total employment of 258 thousand. Not only is this finding consistent with Chicago's role as a transshipment center of national importance, but it also makes a strong argument for the inclusion of inter-regional trade activities in applied cluster analysis to get a more complete picture of regional cluster activities.

Clearly, the full potential of the Retail cluster in strengthening Chicago's local economic base would not have been apparent if we had limited our analysis to identifying only local clusters without looking at domestic clusters. While the cluster components in the local Retail cluster gain in competitiveness by co-locating next to one another, it is the domestic cluster components including the Wholesale Trade and the Warehouse and Storage sectors that have the potential to boost Chicago's competitiveness through cross-boundary inter-industry trade.

The second cluster we focus on is Financial Services. Unlike the Retail cluster, this cluster differs only slightly in the two frameworks in terms of cluster components and employment. As shown in Table 5, the

[18] In their National Industry Cluster Templates study, Feser and Bergman identified as first cluster Metalworking which consists of as many as 116 individual cluster components.

[19] These five retail sectors are: Food and Beverage Stores (445), Nonstore Retailers (454), Miscellaneous Store Retailers (453), Sporting Goods, Hobby, Book, and Music Stores (451), and Gasoline Stations (447).

[20] These seven retail sectors are General Merchandise Stores (452), Clothing and Clothing Accessories Stores (448), Motor Vehicle and Parts Dealers (441), Furniture and Home Furnishings Stores (442), Health and Personal Care Stores (446), Building Material and Garden Supply Stores (444), and Electronics and Appliance Stores (443).

local cluster has ten components with a total cluster employment of 444 thousand. On the other hand, the domestic framework has nine component sectors and 439 thousand jobs. Funds, Trusts, and Other Financial Vehicles (525) and Other Accommodations (72119, 7212, 7213) are the only cluster components that belong exclusively to the local Financial Services cluster while All Other Miscellaneous Professional and Technical Services belongs exclusively to the domestic Financial Services cluster.

Table 5: Financial Service Clusters Comparison

Local Framework

Financial Services	**Employment**	**Factor Loadings**	**Power of Dispersion (U_j)**	**Sensitivity of Dispersion (U_i)**
Funds, trusts, and other financial vehicles	14,691	0.932	1.32	1.28
Transit and ground passenger transportation	34,501	0.378	1.03	0.97
Other accommodations	798	0.465	1.19	0.76
Securities, commodity contracts, investments	140,991	0.961	0.98	0.94
Monetary authorities and depository credit in	65,767	0.829	0.95	0.96
Legal services	58,454	0.399	0.89	0.92
Accounting and bookkeeping services	53,558	0.768	0.83	0.79
Nondepository credit intermediation and rela	33,263	0.597	0.87	0.87
Hotels and motels, including casino hotels	27,861	0.495	0.92	0.93
Automotive equipment rental and leasing	14,313	0.663	0.94	0.92
	444,196			

Domestic Framework

Financial Services	**Employment**	**Factor Loadings**	**Power of Dispersion (U_j)**	**Sensitivity of Dispersion (U_i)**
Securities, commodity contracts, investments	140,991	0.998	1.09	1.00
Legal services	58,454	0.386	1.08	1.06
Accounting and bookkeeping services	53,558	0.759	1.21	1.20
Transit and ground passenger transportation	34,501	0.371	1.04	1.04
Nondepository credit intermediation and rela	33,263	0.983	1.18	1.18
Hotels and motels, including casino hotels	27,861	0.501	1.19	1.08
Automotive equipment rental and leasing	14,313	0.644	1.07	1.17
All other miscellaneous professional and tech	10,551	0.527	1.22	1.31
Monetary authorities and depository credit in	65,767	0.980	0.85	0.90
	439,258			

Source: Implan Pro and Bureau of Labor Statistics.

The high degree of similarity between the two frameworks in the Financial Services cluster can be viewed as a local phenomena with equally important local and cross-boundary ties, a finding that highlights the important position of Chicago as a financial center in the Midwest and nationwide. Four of the ten local components are establishments directly engaged in financial transactions and/or in facilitating financial transactions. These are Funds, Trusts, and Other Financial Vehicles (525), Securities, Commodity Contracts, and Other Financial Investments and Related Activities (523), Monetary Authorities and Depository Credit Intermediation (521 and 5221), and Nondepository Credit Intermediation and Related Activities (5222 and 5223). Of these four financial sectors, only Funds, Trusts, and Other Financial Vehicles (525) is classifiable as a key industry sector in the local cluster with backward and forward linkage indices of 1.32 and 1.28 respectively. Although the remaining three financial sectors all have below average linkage measures in the local framework, two of them nevertheless stand out as important key sectors in the domestic framework.[21] Furthermore, while eight out of the nine cluster components in the domestic framework qualify as key sectors, only one of the eleven local cluster components meets the key industry criteria. The importance of the Financial Services cluster as whole is thus better captured by the domestic framework.

Our third focus cluster is the Food Manufacturing cluster shown in Table 6.[22] This is an example of a cluster where similarities as well as significant differences between local versus domestic industrial clusters are apparent.

A number of interesting observations can be made from Table 6. First, there is a high degree of similarity in the number of cluster components in the two frameworks, with 26 components in the local cluster versus 23 in the domestic cluster. Another similarity between them is that nineteen component sectors are common to both frameworks.

As for the differences between these two clusters, one interesting finding is that while the local framework shows strong linkages between food manufacturing industries and packaging plastic and paper industries, such linkages do not appear in the domestic cluster. This suggests that a substantial amount of packaging material is provided locally. Of the 26 industries in the local cluster, seven supply plastic and paper packaging materials—essential inputs to the food manufacturing industries. These are Laminated Plastics Plate, Sheet, and Shapes (32613), Plastics Packaging Materials, Film and Sheet (32611), Plastics Bottle Manufacturing (32616), Paper Bag and Coated Paper Manufacturing (32222), Other Converted Paper Manufacturing (32229), Other Plastics Product Manufacturing (32619), and Paperboard Container Manufacturing (32221). In addition, the Doll, Toy, and Game Manufacturing (33993) sector also provides some inputs to the Snack Food Manufacturing industry (31191), namely the toys included in kids meals sold by food manufacturing.

Another difference between the two frameworks is that, in contrast to the local cluster, the domestic cluster shows linkages between food manufacturing and agricultural production activities. The agricultural and fishing-related industries included in the domestic cluster are Crop Farming (111), Animal Production (112), Fishing and Hunting (114), and Animal Food Manufacturing (3111). Since agricultural activities play only a marginal role in the economy of the largely urban Chicago region, inputs such as crops and livestock need to be imported from outside the region. This could explain the existence of strong linkages between food manufacturing and agricultural production activities only in the domestic framework.

A third difference of significance relates to the economic importance of individual cluster components. As indicated in Table 6, seventeen of the 26 cluster components in the local framework qualify as key industry sectors giving the Food cluster an important place in the local economy. In the domestic framework, on the other hand, only one sector—Animal Production (112)—can be classified as a key industry sector.

[21] These two sectors are: Securities, Commodity Contracts, and Other Financial Investments and Related Activities (523) and Credit Intermediation and Related Activities (5222 and 5223).
[22] According to Porter (2003), food processing is one of three specializations in the Chicago region.

Table 6: Food Clusters Comparison

Local Framework

Food	Employment	Factor Loadings	Power of Dispersion (U_j)	Sensitivity of Dispersion (U_i)
Other Plastics Product Manufacturing	24,123	0.617	1.07	1.21
Bread and Bakery Product Manufacturing	14,925	0.987	1.02	1.14
Sugar and Confectionery Manufacturing	8,497	0.642	1.01	1.15
Beverage Manufacturing	6,897	0.919	1.10	1.17
Meat Processed from Carcasses	6,105	0.983	1.20	1.41
Fruit and Vegetable Preserving	5,878	0.983	1.08	1.23
Cookie, Cracker, and Pasta Manufacturing	4,672	0.846	1.03	1.11
Paper Bag and Coated Paper Mfg	4,263	0.884	1.06	1.20
Plastics Packaging Materials, Film and Sheet	3,308	0.916	1.05	1.18
All Other Food Manufacturing	2,706	0.636	1.08	1.19
Grain and Oilseed Milling	2,521	0.637	1.05	1.35
Dairy Product Manufacturing	1,872	0.945	1.02	1.33
Plastics Bottle Manufacturing	1,717	0.887	1.01	1.05
Seasoning and Dressing Manufacturing	1,514	0.960	1.08	1.13
Fishing and Hunting	780	0.982	1.01	1.03
Snack Food Manufacturing	692	0.927	1.10	1.16
Coffee and Tea Manufacturing	171	0.941	1.27	1.06
Paperboard Container Manufacturing	12,865	0.408	0.95	1.28
Animal, Except Poultry, Slaughtering	3,641	0.483	0.92	1.52
Other Converted Paper Manufacturing	1,486	0.875	0.99	1.03
Tortilla Manufacturing	908	0.925	1.01	0.95
Seafood Product Preparation	180	0.990	1.09	0.89
Doll, Toy, and Game Mfg	1,370	0.922	0.92	0.69
Flavoring Syrup and Concentrate Manufacturing	590	0.397	0.89	0.82
Laminated Plastics Plate, Sheet, and Shapes	364	0.933	0.95	0.57
Rendering, Meat Byproduct and Poultry Processing	336	0.977	0.90	0.99
	112,383			

Domestic Framework

Food	Employment	Factor Loadings	Power of Dispersion (U_j)	Sensitivity of Dispersion (U_i)
Animal Production	1,002	0.741	1.58	1.03
Paperboard Container Manufacturing	12,865	0.407	0.87	1.18
Crop Farming	6,175	0.706	0.83	1.22
Animal, Except Poultry, Slaughtering	3,641	0.955	1.10	0.72
Foam Product Manufacturing	3,033	0.682	1.21	0.98
Grain and Oilseed Milling	2,521	0.881	0.76	1.07
Dairy Product Manufacturing	1,872	0.946	1.08	0.72
Fishing and Hunting	780	0.983	1.11	0.87
Flavoring Syrup and Concentrate Manufacturing	590	0.398	1.41	0.96
Animal Food Manufacturing	547	0.693	0.74	1.19
Seafood Product Preparation	180	0.990	1.22	0.86
Bread and Bakery Product Manufacturing	14,925	0.987	0.88	0.63
Sugar and Confectionery Manufacturing	8,497	0.549	0.78	0.96
Beverage Manufacturing	6,897	0.929	0.76	0.63
Meat Processed from Carcasses	6,105	0.985	0.90	0.63
Fruit and Vegetable Preserving	5,878	0.945	0.74	0.88
Cookie, Cracker, and Pasta Manufacturing	4,672	0.838	0.79	0.49
All Other Food Manufacturing	2,706	0.672	0.82	0.51
Seasoning and Dressing Manufacturing	1,514	0.961	0.91	0.66
Tortilla Manufacturing	908	0.941	0.78	0.45
Snack Food Manufacturing	692	0.930	0.82	0.58
Rendering, Meat Byproduct and Poultry Processing	336	0.976	0.94	0.93
Coffee and Tea Manufacturing	171	0.942	0.99	0.64
	85,507			

Source: Implan Pro and Bureau of Labor Statistics.

CONCLUSIONS AND FINAL REMARKS

The literature on analytical industrial cluster identification methodologies has paid much attention to the definition of the study region. Specifically, when using input-output tables to identify industrial clusters based on inter-industry linkages, the question of whether to give preference to a national or to a regional input-output framework has been raised repeatedly in the relevant literature. Promoters of the regional framework, for instance Ó hUallacháin (1984), argue that with data confined to a single region, identified regional clusters represent the gateway to regional economic development. Advocators of the national framework on the other hand, for instance Feser and Bergman (2000), stress the importance of using national input-output tables. Here the focus is on identifying future economic development opportunities through an understanding of missing sectors in existing or potential clusters. But since these two approaches focus either solely on local or on national inter-industry transactions, they are often not able to identify regional clusters that have strong inter-industry linkages with other regions. The domestic clusters identified using our approach overcomes this limitation. We argue that in a framework of demand-side oriented economic development policies with focus on export promotion strategies (Eisinger, 1988), the identification of 'domestic' industrial clusters provides valuable insights of the competitive position of local industries in the larger national economy.

In this research, we applied principal components analysis to two conceptually different input-output frameworks: (i) local framework based on purely regional transactions, and (ii) domestic framework utilizing combined regional and non-regional transactions. This analytical approach allowed us to identify three different groups of clusters—common, unique and analogous—for both the local and domestic frameworks. In addition, we used the backward and forward linkage indices proposed by Beyers (1976) to evaluate the economic importance of individual cluster components and refine the clusters derived from the PCA results. We used the Chicago metropolitan area as a case study to demonstrate the applicability of this approach.

The clusters *common* to both the regional and domestic frameworks include clusters such as Retail, Construction and Building Services, Computer-based Support, and Financial Services. The second group, which we refer to as *unique* clusters, contains clusters that emerged either in the local framework (e.g., Personal Services) or in the domestic framework (e.g., Professional Services), but not in both. And finally, the third group, or *analogous* clusters, includes all clusters that show some similarities between the two frameworks, but which are still too different to be labeled as *common* clusters. Altogether, we have identified eighteen clusters in the local framework and twenty clusters in the domestic framework.

Most of the identified clusters were common to both local and domestic frameworks and illustrate the integration of Chicago in the national economy. The unique local industrial clusters correspond to the regional clusters that would be obtained using purely regional data. They represent industry groups that have strong interactions amongst themselves within the region. The unique domestic clusters, on the other hand, give insights similar to those obtained from the national framework proposed by Bergman and Feser and others, while taking into account the unique features of the local economy. They represent sectors which have strong regional as well as national linkages, and which could over time be induced through appropriate policy measures to increase their vertical and horizontal relationships within the region.

REFERENCES

Bergman, E. M. and Feser, E. J. 1999 *Industrial and Regional Clusters: Concept and Comparative Applications.* Web Book *in Regional Science*, Regional Research Institute, West Virginia University, http://www.rri.wvu.edu/WebBook/Bergman-Feser/contents.htm.

Beyers, W. B. 1976 Empirical Identification of Key Sectors: Some further Evidence. *Environment and Planning A* 8: 231-236.

Bradshaw T. K. and E. J. Blakely 1999 What are 'third wave' state economic development efforts? From Incentives to Industrial Policy Third Wave*. Economic Development Quarterly* 13 (3) 229-244

Botham, R., Gibson, H., Martin R. and Moore, B. 2001 Business Clusters in the U.K – A First Assessment. *A Report for the Department of Trade and Industry,* A Consortium led by Trends Business Research.

Chenery and Watanabe 1958 International Comparisons of the Structure of Production. *Econometrica* 26: 487.

Czamanski, S. 1974 *Study of Clustering of Industries*. Halifax, Nova Scotia, Canada: Institute of Public Affairs, Dalhousie University.

Czamanski, S. 1976 *Study of Spatial Industrial Complexes*. Halifax, Nova Scotia, Canada: Institute of Public Affairs, Dalhousie University.

Czamanski and L. A. de Q. Ablas 1979 Identification of Industrial Clusters and Complexes: A Comparison of Methods and Findings. *Urban Studies* 16: 61-80.

Doeringer, P.B. and Terkla, D.G. 1995 Business Strategy and Cross-industry Clusters. *Economic Development Quarterly* 9: 225-37.

Eisinger P. K. 1988 *The Rise of the Entrepreneurial State: State and Local Economic Development Policy in the United States.* Madison: The University of Wisconsin Press.

Feser, E. J. and Luger, M. 2002 Cluster Analysis as a Mode of Inquiry: It's Use in Science and Technology Policymaking in North Carolina. *European Planning Studies* 11.1: 1-14.

Feser, E. J. and E. M. Bergman 2000 National Industry Templates: A Framework for Applied Regional Cluster Analysis. *Regional Studies* 34.1: 1-19.

Fosler, R.S. 1992 State Economic Policy: The Emerging Paradigm. *Economic Development Quarterly* 6(1): 3-13.

Hanson, G. H. 2000 Scale Economies and the Geographic Concentration of Industry. Working Paper 8013, National Bureau of Economic Research. Cambridge: Massachusetts.

Hazari, B. R. 1970 Empirical Identification of Key Sectors in the Indian Economy. *The Review of Economics and Statistics* 52(3): 301-305.

Herbers, J. 1990 A Third Wave of Economic Development. *Governing*. 3(9): 43-50.

Hewings, G. J. D. 1974 The Effect of Aggregation on the Empirical Identification of Key Sectors in a Regional Economy: A Partial Evaluation of Alternative Techniques. *Environment and Planning A* 6: 439-453.

Hill, E. W. and Brennan, J.F. 2000 A Methodology for Identifying the Drivers of Industrial Clusters: The Foundation of Regional Competitive Advantage. *Economic Development Quarterly* 14: 67-96.

Hoover, E. M. 1948 *The Location of Economic Activity*. New York: McGraw-Hill.

Isard, W., Schooler, E.W. and Vietorisz, T. 1959 *Industrial Complex Analysis and Regional Development: A Case Study of Refinery-Petrochemical-synthetic-Fiber Complexes and Puerto Rico.* Cambridge: Technology Press of the Massachusetts Institute of Technology.

Latham, W.R. III 1976 Needless Complexity in the Identification of Industrial Complexes. *Journal of Regional Science* 16.1: 45-55.

Marshall, A. 1920, 8th. edition *Principles of Economics*. London: Macmillan.

Martin, R. and Sunley, P. 2003 Deconstructing Clusters: Chaotic Concept or Policy Panacea? *Journal of Economic Geography* 3: 5-35.

Munnich, L. W., Bau, M. M., Clark, J. J., Young, N. J. and Patel, A. M. 1998 *Southwest Minnesota Industry Cluster Study.* Humphrey Institute of Public Affairs and Minnesota Extension Service of the University of Minnesota.

Ó hUallacháin, B. 1984 The Identification of Industrial Complexes. *Annals of the Association of American Geographers* 73.3: 420-436.

Peters, D. 2004 Revisiting Industry Cluster Theory and Method for Use in Public Policy: An Example Identifying Supplier-based Clusters in Missouri. Presented at the Mid-Continent Regional Science Association, 35th. Annual Meeting, Madison, Wisconsin.

Porter, M. E. 2003 The Competitive Advantage of Regions. Presentation at The Indiana Leadership Summit, Indianapolis, Indiana, May 13th, 2003

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

Rasmussen, P. 1952 *Studies in Intersectorial Relations*. Amsterdam: North-Holland Publishing.

Rex, T. 1999 Prominent Industry Clusters Vary by County. *Arizona Business* 46.6: 6-8.

Roepke H. D., Adams, D. and Wiseman, R. 1974 A New Approach to the Identification of Industrial Complexes Using Input-Output Data. *Journal of Regional Science* 14.1: 15-29.

Rosenfeld, S. A. 1997 Bringing Business Clusters into the Mainstream of Economic Development. *European Planning Studies* Vol. 5 No. 1.

Streit, M. E. 1969 Spatial Associations and Economic Linkages between Industries. *Journal of Regional Science* 9:177.

Tabachnick, B. G. and L. S. Fidell 2007 *Using Multivariate Statistics.* Fifth Edition. Boston: Pearson Education, Inc.

Tinsley, H. E. A., and D. J. Tinsley 1987 Uses of factor analysis in counseling psychology research, *Journal of Counseling Psychology* 34, 414-24.

U.S. Census Bureau, 2006 Metropolitan and Micropolitan Statistical Area Definitions, http://www.census.gov/population/www/estimates/metrodef.html.

U.S. Department of Labor, Bureau of Labor Statistsics, 2006 Quarterly Census of Employment and Wages (QCEW), www.bls.gov/cew/home.htm.

vom Hofe, R. and Chen, K. 2006 Whither or not Industrial Cluster: Conclusions or Confusions. *The Industrial Geographer* 4: 2-28.