Technical Appendix to ***Kentucky Clusters***
Derivation of Benchmark U.S. Clusters

Edward J. Feser
Assistant Professor
Department of City and Regional Planning, CB 3140, University of North Carolina, Chapel Hill, NC 27599-3140; Voice: (919) 962-4768; Fax: (919) 962-5206; Email: feser@email.unc.edu; http://www.unc.edu/depts/dcrpweb/

Twenty-eight General Benchmark Clusters[1]

Our goal was to develop a consistent set of clusters comprised of linked export-oriented industries. We began by analyzing value-chain relationships via a factor analysis of 1992 U.S. input-output (I-O) data.[2] However, I-O data do not permit the statistical identification of meaningful value-chain ties between all private sector industries of interest to the study. A well-known limitation of U.S. I-O accounts is that services and other non-manufacturing sectors are poorly represented relative to manufacturing. Of some 500 hundred I-O industries in the table, 362 are in the manufacturing sector. Many of the remainder are agriculture, construction, and consumer services industries that are not of interest to our study.[3] There is thus simply is not enough information on inter-industry linkages among export-oriented non-manufacturing sectors to develop meaningful clusters via statistical factor analysis. Moreover, product linkages are often not an appropriate indicator of inter-industry ties between many human capital intensive non-manufacturing industries.

We therefore elected to confine our input-output analysis to the 367 manufacturing and selected export-oriented industries in health and business services that are reasonably well-represented in the table. We then organized the remaining industries in the study into clusters using information from an early analysis of joint labor requirements (described in Feser, 2000), a row-by-row inspection of input-output ties to other industries, other available sectoral studies, and judgement. While imperfect, this hybrid methodological approach allowed us to balance the need for a systematic and rigorous approach to cluster identification with the practical policy need to

---

[1] The 28 general benchmark clusters and 8 technology clusters were developed for a previous report, titled *High-Tech Clusters in North Carolina* (May 2000; see the reference list at the end of this document). The labor-based clusters based on ONET skills data were developed for this report and another study conducted concurrently for the North Carolina's Southeast, an economic development agency serving the southeast region of North Carolina.

[2] *Benchmark Input-Output Accounts of the United States*, 1992 (Washington, DC: U.S. Department of Commerce).

[3] In some cases, there are export-oriented industries included in the table that cannot be included in the analysis because of problems with their concordance to SIC industries or because subsequent data on their size and trends at the sub-state level are not available. The most important example is mining.

include important non-manufacturing industries in the analysis of state and sub-state economies.[4]

Factor Analysis of Inter-industry Linkages. We first developed a standard 491 by 491 inter-industry transactions matrix from the detailed industry by commodity benchmark input-output accounts. To identify a set of broad value chains from the raw input-output information, we conducted a factor analysis on a matrix of inter-industry linkages originally proposed by Czamanski (1974). Specifically, we formed–from the 491 by 491 matrix of inter-industry transactions, **A**–two matrices, **X** and **Y**, with elements:

$$x_{ij} = \frac{a_{ij}}{a_{+j}}, \quad y_{ij} = \frac{a_{ij}}{a_{i+}}$$

where a_{ij} is the dollar value of goods and services sold by industry *i* in some period to industry *j*, and a_{+j} and a_{i+} are total intermediate good purchases and sales, respectively, of industries *i* and *j* over the same period. x_{ij} is intermediate good purchases by sector *j* from *i* as a proportion of *j*'s total intermediate good purchases. A large value for x_{ij} implies that *j* depends on industry *i* as a source for a significant share of its total intermediate purchases. y_{ij} represents intermediate good sales from *i* to *j* as a proportion of *i*'s total intermediate good sales, such that a large value for y_{ij} implies that *i* depends on industry *j* as a market for a large proportion of its total intermediate good sales. The columns of **X** are the intermediate input purchasing pattern of each industry *j*, while the rows of **Y** are the intermediate output sales pattern of each industry *i*.

For any two industries (A and B) with the column vectors of **X** defined as $\mathbf{x}_A$, and $\mathbf{x}_B$ and the row vectors of **Y** defined as $\mathbf{y}_A$ and $\mathbf{y}_B$, four correlations on the sales and purchasing vectors of any two industries may be derived (again, following Czamanski): 1) $r(\mathbf{x}_A @ \mathbf{x}_B)$ measures the similarity in input purchasing patterns of industries A and B; 2) $r(\mathbf{y}_A @ \mathbf{y}_B)$ measures the degree to which A and B possess similar output selling patterns, i.e. the degree to which they sell goods to a similar mix of intermediate input buyers; 3) $r(\mathbf{x}_A @ \mathbf{y}_B)$ measures the degree to which the buying pattern of industry A is similar to the selling pattern of industry B, i.e. the degree to which industry A purchases inputs from industries in which B supplies (a second-tier linkage); and 4) $r(\mathbf{x}_B @ \mathbf{y}_A)$ measures the degree to which the buying pattern of industry B is similar to the selling pattern of industry A, i.e. the degree to which industry B purchases inputs from industries in which A supplies. A linkage matrix, **L**, comprising the largest of these four correlations for each pair of sectors, summarizes the degree of linkage between and among all 491 sectors.

We reduced **L** by eliminating agriculture, forestry and fishing; mining; construction, wholesale and retail trade; finance, insurance and real estate; and personal services. The following are the 367 IO sectors included in the factor analysis; each IO sector can be matched to SIC industries, in

[4] Note that 1996 study *Targeting North Carolina Manufacturing* focused exclusively on the manufacturing sector in part because our initial attempts to include non-manufacturing in the IO-based analysis proved unsuccessful.

most cases in direct one-to-one correspondence:

```
I130100 I130200 I130300 I130500 I130600 I130700 I140101 I140102 I140105 I140200 I140300 I140400
I140500 I140600 I140700 I140800 I140900 I141000 I141100 I141200 I141301 I141302 I141401 I141402
I141403 I141501 I141502 I141600 I141700 I141801 I141802 I141803 I141900 I142002 I142004 I142005
I142101 I142102 I142103 I142104 I142200 I142300 I142400 I142500 I142600 I142700 I142800 I142900
I143000 I143100 I143201 I143202 I150101 I150102 I150103 I150200 I160100 I160200 I160300 I160400
I170100 I170600 I170700 I170900 I171001 I171100 I180101 I180102 I180201 I180202 I180203 I180300
I180400 I190100 I190200 I190301 I190302 I190303 I190304 I190305 I190306 I200100 I200200 I200300
I200400 I200501 I200502 I200600 I200701 I200702 I200703 I200800 I200901 I200903 I200904 I210000
I220101 I220102 I220103 I220200 I220300 I220400 I230100 I230200 I230300 I230400 I230500 I230600
I230700 I240100 I240400 I240500 I240701 I240702 I240703 I240705 I240706 I240800 I250000 I260100
I260200 I260301 I260302 I260400 I260501 I260601 I260602 I260700 I260802 I260803 I260806 I270100
I270201 I270202 I270300 I270401 I270402 I270403 I270404 I270405 I270406 I280100 I280200 I280300
I280400 I290100 I290201 I290202 I290203 I290300 I300000 I310101 I310102 I310103 I310200 I310300
I320100 I320200 I320300 I320400 I320500 I320600 I330001 I340100 I340201 I340202 I340301 I340302
I340303 I340304 I340305 I350100 I350200 I360100 I360200 I360300 I360400 I360500 I360600 I360701
I360702 I360800 I360900 I361000 I361100 I361200 I361300 I361400 I361500 I361600 I361700 I361900
I362000 I362100 I362200 I370101 I370102 I370103 I370104 I370105 I370200 I370300 I370401 I370402
I380100 I380400 I380501 I380600 I380700 I380800 I380900 I381000 I381100 I381200 I381300 I381400
I390100 I390200 I400100 I400200 I400300 I400400 I400500 I400600 I400700 I400800 I400901 I400902
I410100 I410201 I410202 I410203 I420100 I420201 I420202 I420300 I420401 I420402 I420500 I420700
I420800 I421000 I421100 I430100 I430200 I440001 I440002 I450100 I450200 I450300 I460100 I460200
I460300 I460400 I470100 I470200 I470300 I470401 I470402 I470404 I470405 I470500 I480100 I480200
I480300 I480400 I480500 I480600 I490100 I490200 I490300 I490500 I490600 I490700 I490800 I500100
I500200 I500300 I500400 I510102 I510103 I510104 I510400 I520100 I520200 I520300 I520400 I520500
I530200 I530300 I530400 I530500 I530700 I530800 I540100 I540200 I540300 I540400 I540500 I540700
I550100 I550200 I550300 I560100 I560200 I560300 I560500 I570100 I570200 I570300 I580100 I580200
I580400 I580600 I580700 I590100 I590200 I590301 I590302 I600100 I600200 I600400 I610100 I610200
I610300 I610500 I610601 I610603 I610700 I620101 I620102 I620200 I620300 I620400 I620500 I620600
I620700 I620800 I620900 I621000 I621100 I630200 I630300 I640101 I640102 I640104 I640105 I640200
I640301 I640302 I640400 I640501 I640502 I640503 I640504 I640700 I640800 I640900 I641000 I641100
I641200 I660100 I730104 I730112 I730302 I770200 I770305
```

To derive the clusters, we wrote a program using SAS software that repeated the factor analysis on the linkage matrix **L** over 50 times, changing the number of components subject to rotation by 1 on each iteration.[5] We began by rotating 10 components and ended by rotating 50.[6] The program aided the analysis by sorting the rotated factors by loading, attaching detailed SIC industry labels to each I-O sector, and producing scree plots and tables reporting the relative proportion of variance explained by each component and the size of the associated eigenvalues.[7] After inspecting each set of results in terms of indicators of fit, economic plausibility, and general interpretability, we selected a final model revealing 26 components.

Associated with a given component is a reduced set of variables–in this case, sectors–which constitute its key statistical element. The indicator for determining this element is the individual loading, which is a measure of the relative strength of the relationship between a given variable and each derived component. The magnitudes of the loadings therefore determine the

[5] All calculations were performed using SAS Interactive Matrix Language (IML) software as well as the SAS FACTOR procedure.

[6] Our selection of the range 10-50 was based on findings from factor analysis of 1987 manufacturing linkages reported in *Targeting North Carolina Manufacturing*.

[7] In contrast to the approach used in *Targeting North Carolina Manufacturing*, we used a Promax rotation. Promax produces an orthogonal pre-rotation (equivalent to Varimax) followed by an oblique rotation. Oblique rotation is favored when a high degree of overlap across factors is expected.

membership in each cluster. Only sectors with loadings of 0.35 or higher were identified with a particular cluster. Based on that criterion, 23 of the 26 components could be interpreted in a straightforward manner by examining underlying input-output linkages. For example, the packaged food products cluster is comprised of 45 industries, including various food products sectors (within SIC 20), packaging machinery (SIC 3565), food products machinery (SIC 3556), and paper (SIC 2676). Three clusters lacked a discernable economic logic and appeared to be statistical artifacts; two of the three were comprised of only three industries. Note that the clusters are not mutually exclusive; any given underlying industry may be a member of multiple clusters.

Treatment of Pharmaceuticals. The 23 clusters derived from the input-output data constitute the core of the benchmark clusters. However, a number of industries failed to load highly on any group. The most important of these is pharmaceuticals or drugs manufacturing, which unfortunately constitutes only a single sector in the I-O table. Rather than exclude this important industry in subsequent economic analyses, we opted to consider pharmaceuticals its own "cluster." Thus SIC 283 (IO code 290100) is added to the group of 23 I-O-based clusters.

Other Remaining Export-Oriented Industries. As noted above, early tests showed that the inclusion of some important higher-end producer services industries with potential to export could not be included in the I-O analysis. Because they are significant and important suppliers to a very wide range of industries–and because their sub-industries are rarely delineated in the I-O table itself–their inclusion in the matrix **L** for the factor analysis tended to produce large and obscure factors.[8] They include industries such as legal services, transportation, shipping, securities, insurance, banking, and advertising. Rather than include them in the benchmark set as individual sectors (unlike the case of pharmaceuticals, we had no additional information about their relationship to other industries from the factor analysis itself), we elected to use supplementary information from an analysis of industry labor requirements (reported in Feser, 2000), a row-by-row inspection of I-O linkages, other industry studies, and judgement to help determine their membership in either the existing 24 I-O-based clusters or entirely new groupings.

Based on this approach, we assigned the remaining sectors to four new clusters: securities and insurance, banking and advertising, legal services (comprised of a single industry), and transportation, shipping, and logistics. The final set of 28 clusters reported in Appendix 1, Table 1 of *Kentucky Clusters* is thus comprised of 24 clusters developed from an analysis of 1992 U.S. input-output patterns and four clusters based loosely on input-output relationships, U.S. labor requirements patterns (as revealed in the *U.S. Staffing Patterns Matrix* for1994), and other sources of information on industry relationships (such as the *U.S. Industry and Trade Outlook*). The 28 are essentially value-chain clusters, since input-output relationships are at the core of their development.

[8] This is both a limitation of statistical factor analysis as well as a problem of data availability.

The primary disadvantages behind the combined qualitative and quantitative methodology are validity and interpretability. The fact that a single metric, source of data, and statistical methodology were not used to define the full set limits our ability to judge the validity of the findings. Yet at the same time, using inadequate input-output data or neglecting important sectors for the sake of a single methodological approach would have severely compromised the utility of the clusters for subsequent economic and policy analyses.

Eight Benchmark Technology-Intensive Value-Chain Clusters

To derive the eight technology-intensive clusters, we reduced **L** to represent only SIC sectors commonly identified as technology-intensive.[9] The following 109 I-O codes concord to 148 technology-intensive SIC industries:

```
I130100 I270100 I270201 I270202 I270300 I270401 I270402 I270403 I270404 I270405 I270406 I280100
I280200 I280300 I280400 I290100 I290201 I290202 I290203 I290300 I300000 I430100 I430200 I450100
I450200 I450300 I460100 I460200 I460300 I460400 I470100 I470200 I470300 I470401 I470402 I470404
I470405 I470500 I480100 I480200 I480300 I480400 I480500 I480600 I490100 I490200 I490300 I490500
I490600 I490700 I490800 I510102 I510103 I510104 I510400 I530200 I530300 I530400 I530500 I530700
I530800 I540100 I540200 I540300 I540400 I540500 I540700 I550100 I550200 I550300 I560100 I560200
I560300 I560500 I570100 I570200 I570300 I580100 I580200 I580400 I580600 I580700 I590100 I590200
I590301 I590302 I600100 I600200 I600400 I610603 I620101 I620102 I620200 I620300 I620400 I620500
I620600 I620800 I620900 I621000 I621100 I630200 I630300 I660100 I730104 I730112 I730302 I770200
I770305
```

We then repeated the factor analysis procedure described above on this reduced **L** matrix, producing repeated rotations beginning with 5 components and ending with 20. Using the same criteria as for the full set above, we selected a final model with 8 components. I-O sectors were generally included in the cluster if their loading was equal to or exceeded 0.35, though in some cases we eliminated sectors after inspecting underlying input-output patterns in each cluster.[10] The SIC concordance to the I-O clusters are reported in Appendix 1, Table 2 of *Kentucky Clusters*.

Seventeen Labor-Based Industry Clusters

Labor-based industry clusters are groups of industries that face similar labor requirements defined in terms of the relative utilization of internally homogenous skill groups. Two sectors that essentially utilize the same types of labor skills should fall into the same cluster. The problem is that data on the skill needs of individual sectors has historically been very scarce.

---

[9] The original set of technology-intensive SIC industries is from the North Carolina Employment Security Commission (and is reported in *Making Manufacturing and Technology Work for North Carolina*, North Carolina Alliance for Competitive Technologies, Raleigh, NC, December 1995). NC ESC's list derives from US BLS and NSF data on R&D expenditures and the share of scientists and engineers by sector. We amended the list to exclude SICs 7376 (computer facilities management), 7377 (computer rental and leasing), and 7378 (computer maintenance and repair). This has implications for the subsequent economic analysis in the study rather than the factor analysis since SIC 737 is represented by only a single I-O code (730104).

[10] For example, the carbon black industry (SIC 2895) loaded with the communications services and software cluster on the basis of a strong tie to only a single other cluster member. The industry's technology-intensity is also questionable.

Instead, the primary source of information on detailed labor utilization patterns by sector has been the U.S. *Staffing Patterns Matrix*, which reports the number of employees in each of 501 occupations across 241 industries (the subject of an analysis in Feser 2000). The aggregated nature of the staffing patterns matrix in terms of occupation (there are thousands of occupations defined in the *Dictionary of Occupational Titles*) means that there is probably greater homogeneity of skill within clusters than between them. Nevertheless, a limitation of the use of the staffing patterns data is that there are likely workers with similar skills (i.e., that may be transferred easily between occupations) classified in different nominal occupations. This introduces greater heterogeneity in industry labor requirements than there actually may be.

Relatively recently, the Department of Labor's Employment and Training Administration undertook a project to characterize, in an easily accessible relational database, the basic skill, knowledge, and worker attributes of thousands of occupations. Called ONET, the database is designed to "allow users to profile similarities and differences across occupations and anticipate skill changes now and into the 21st century."[11] ONET, as the electronic replacement of the *Dictionary of Occupational Titles*, is organized into a conceptual model of six dimensions, each of which characterizes a dimension of a detailed set of occupations: worker characteristics (abilities, interests, and work styles); worker requirements (basic and cross-functional skills, general knowledge, and education); experience requirements (training, experience, licensing); occupation requirements (work activities and context, organizational context); occupation specific information (knowledge, skills, tasks, and machinery/equipment); and occupational characteristics (such as wages and the labor market outlook). Although ONET has already been through two releases, much of the data underlying the model have yet to be collected.

In the latest release, the six dimensions are defined by 132 variables describing the general skill and ability requirements of each of 1,122 occupations. Each occupation is rated in two ways: 1) on a scale of 1 (low) to 7 (high) according to the level of knowledge or skill for the given category, e.g., in the case of economics and accounting knowledge, from knowledge of simple number and accounting procedures (rated a 1) to knowledge of complex financial operations, regulations, agencies and advanced cost accounting procedures (rated a 7); and 2) on a scale of 0 (low) to 5 (high), according to the "importance" of the skill to the occupation in terms of centrality to the work involved and frequency of use.

The ONET variables provide us with a means to classify occupations into groups, or clusters, based on similarities in the skill set of the workforce holding those occupations. Assuming a concordance between ONET and the BLS *Staffing Patterns Matrix* can be established, the skill groups can be used to produce a revised staffing patterns matrix (what we call an industry/occupation matrix below) that reports the workforce *skill*–rather than *occupational*–requirements vectors of 241 industries. That revised matrix can then be used to cluster industries into groups based on utilization of similar skills, or, by extension, the utilization of joint labor pools.

[11] See www.doleta.gov/programs/onet/.

Data and procedures. The two basic data sources are the 1,122 row by 132 column ONET skills matrix, where cell s_{ij} is the product of the two variables: the level for skill type *j*, occupation *i*, v_{ij}, rated on a scale of 1 to 7; and the importance of skill type *j* to the occupation *i*, m_{ij}, rated on a scale of 0 to 5. Taking the product of the two dimensions has the effect of weighting more heavily a skill that is both of a high level as well as central to the occupation concerned. Other alternatives, such as adding v_{ij} and m_{ij} or taking the average of the two, tend to treat occupations scoring highly on one factor but low on the other similarly to occupations scoring at moderate levels of the scale on both dimensions. The multiplicative approach also helps to introduce additional variation in the ONET matrix, improving the ability of hierarchical clustering algorithms to identify distinct groups of occupations.

One of the biggest challenges in implementing the methodology we outlined above is establishing a concordance between various occupational data sources (e.g., ONET, *Staffing Patterns*, OES, etc.). We began by conducting a hierarchical cluster analysis on the 1,122 by 132 ONET skill matrix, resulting in 300 aggregated skill groups (or occupations that share the same skill profile). We then established a concordance between the 300 skill groups and the 506 occupations in the *Staffing Patterns Matrix*. Eliminating all non-exporting industries from the *Staffing Patterns Matrix* resulted in a reduced 300 row by 127 column industry/occupation matrix. Each cell in the industry/occupation matrix reports the share of total employment of the column industry accounted for by the given row occupational skill group. Conducting a second hierarchical cluster analysis on the industry/occupation matrix yielded 17 labor-based industry clusters. Comparing the industries in each cluster to their underlying skill profiles provided the information necessary to interpret each cluster. Figure 3.5 in *Kentucky Clusters* summarizes the methodology; the components of the labor-based industry clusters are reported in Appendix 1, Table 3.

Caveats. The ONET system is new and the data have not be subjected to extensive testing. Different clusters might be derived by using different subsets of skill variables or applying different manipulations to the resulting industry/occupation matrix. The labor-based clusters in *Kentucky Clusters* should therefore be regarded as preliminary. We are continuing to develop approaches to labor-based industry cluster analysis using ONET and other sources that should be useful for studies in Kentucky and other states.

Spatial Analysis Details

To study the location and spatial extent of value chain clusters in Kentucky, we use a simple and straightforward measure of geographic concentration developed by Getis and Ord (Getis and Ord, 1992; Ord and Getis, 1995). The Getis and Ord G_i^* statistic measures geographic concentration of employment in industry clusters by detecting patterns of spatial association among arbitrarily or administratively defined areas (in this case, counties). G_i^* is calculated for each individual county and is based on cluster employment levels both for the county itself as well as that county's immediate neighboring counties. In this way, G_i^* detects concentrations of cluster activity across counties boundaries, rather than just within individual counties (as is the case with simple measures such as the location quotient).

Specifically, the measure for county *i* for a given industry cluster is calculated as:

$$G_i^* = \frac{\sum_j w_{ij} x_j - W_i \bar{x}}{s\sqrt{(nS_{1i} - W_i^2)/(n-1)}} \qquad (1)$$

where x is total employment in the cluster, $\{w_{ij}\}$ is a symmetric 0/1 spatial weights matrix that defines neighboring counties to county i (note that $w_{ii} = 1$), W_i is sum of weights in $\{w_{ij}\}$, $\bar{x}$ is the mean of county cluster employment for the U.S. as a whole or $\bar{x} = (\sum_j x_j)/(n-1)$, $S_{1i} = \sum_j w_{ij}^2$, and $s^2 = \left(\sum_j x_j^2 / n-1\right) - (\bar{x})^2$. The first step in calculating the statistic is to develop a spatial weights matrix (w_{ij}) that summarizes the spatial association among neighboring counties. We used a binary adjacency matrix: neighboring counties (including the county itself) are assigned a weight of one while non-neighbors are given a weight of zero. The weights matrix essentially acts as a filter so that only cluster employment of neighboring counties are included in the calculation of local concentration. For simplicity, we assumed G_i^* is distributed approximately standard normal with a 95 percent significance value of 1.96.[12]

The skewed urban size distribution tends to limit the effectiveness of the *G* statistic when very large counties are compared to very small ones. Therefore, we used the natural log of employment in order to dampen the size distribution effect. We calculated the *G* statistic for all counties in 49 of 50 U.S. states using 1997 U.S. Bureau of Labor Statistics ES-202 data, obtained via special permission. The extensive data series permitted us to select a broad study region that included Kentucky's counties as well as the counties of its neighbor states, while avoiding the problem of data suppression common to small areas in data sources such as *County Business Patterns*.

References and Additional Sources

Bergman, E. M, E. J. Feser, and S. H. Sweeney. 1996. *Targeting North Carolina Manufacturing*. Report prepared for the North Carolina Alliance for Competitive Manufacturing (Raleigh, NC). The North Carolina Alliance for Competitive Technologies is no longer in operation. However, key elements of Targeting North Carolina Manufacturing appear in the following two publications:

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

[12] A high value of G_i^* for a given county indicates a clustering of high levels of value chain employment in that county and among neighboring counties. Low values of the statistic indicate a clustering of low values of employment in the value chain. In our maps in the subsequent section, we focus only on high values of the statistic since they are most meaningful in the present context.

Feser, E. J., and E. M. Bergman. 2000. National industry clusters: Frameworks for state and regional development policy. *Regional Studies* 34 (1): 1-20.

Feser, E. J. 2000. Sectoral labor requirements as a basis for interindustry and sectoral clustering. Paper presented at the Western Regional Science Association meetings, Kauai, Hawaii, February 2000. PDF copies available upon request.

Feser, E. J., and H. Renski. 2000. *High-Tech Clusters in North Carolina*. Report prepared for the North Carolina Board of Science and Technology (Raleigh, NC). PDF copy available at http://govhunt.gov.state.nc.us/govoffice/vision2030/index.html.

Getis, A., and J. K. Ord. 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis* 24: 189-206.

Ord, J. K., and A. Getis. 1995. Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis* 27: 286-306.