On the Ellison-Glaeser geographic concentration index

Abstract. I use confidential employment data to investigate the empirical properties of a recent industry geographic concentration index (and related index of industry co-agglomeration) proposed by Ellison and Glaeser (1997). The results show that Ellison and Glaeser's theoretical finding that their concentration measures are robust to differences in the level of spatial aggregation in the underlying employment data does not generally hold in practice. This implies that sensitivity testing for alternative spatial units should accompany any analysis with the concentration measures.

Edward J. Feser
Assistant Professor
Department of City and Regional Planning
CB 3140, New East Building
University of North Carolina
Chapel Hill, NC 27599-3140
Tel: (919) 962-4768 Fax: (919) 962-5206
E-Mail: feser@email.unc.edu

June 2000

On the Ellison-Glaeser geographic concentration index[1]

1. Introduction

The most common measures of industrial geographic concentration include those akin to the Gini coefficient of income concentration (e.g., Isard *et al.* 1998, Krugman 1991) and the four-firm concentration ratio typical in studies of industrial organization (e.g., Enright 1990). In a recent paper, Glenn Ellison and Edward Glaeser (1997) propose an alternative:

$$\gamma \equiv \frac{\sum_{i=1}^{M}(s_i - x_i)^2 - \left(1 - \sum_i x_i^2\right)\sum_k z_k^2}{\left(1 - \sum_i x_i^2\right)\left(1 - \sum_k z_k^2\right)} = \frac{\sum_{i=1}^{M}(s_i - x_i)^2 - \left(1 - \sum_i x_i^2\right)H}{\left(1 - \sum_i x_i^2\right)(1 - H)} \tag{1}$$

where s_i is region i's share of the study industry's employment, x_i is the industry's share of overall manufacturing employment, and z_k is plant *k*'s share of the study industry's total employment. $\sum_k z_k^2 \equiv H$ is the Herfindahl index of the industry's plant size distribution. The authors use published data for U.S. counties and states to explore the spatial concentration of manufacturing sectors. Given the lack of enterprise-level data in most research contexts, they propose a method suggested by Schmalensee (1977) for estimating the measure's Herfindahl component.

While the Ellison-Glaeser index is similar to the standard concentration ratio and Gini-based measures in that can be constructed with readily available data, it differs importantly in that it is derived from an explicit theory of firm location behavior. As such, one can demonstrate that it possesses a number of desirable properties. The index defines concentration as agglomeration above and beyond what we would observe if plants simply chose locations randomly (as opposed to a uniform spatial distribution). It controls for differences in the size distribution of establishments among different industries, thereby accounting for the fact that spatial concentration is partly driven by industrial concentration. And it is, in principle, robust to differences in the level of spatial aggregation at which industry data are available. The authors note that those features mean that 'one may compare with more confidence. . .the concentration of American and European industries, the concentration of high- and low-tech industries, and the changes in levels of concentration over time (p. 890-1).' Researchers will undoubtedly move to carry out such comparisons (e.g., see Maurel and Sédillot 1999).

This note sheds light on the likely empirical validity of comparisons with the index by exploring its behavior under 'controlled conditions.' I use confidential establishment-level employment data for Tennessee and North Carolina to test for the presence of two possible 'layers' of error in the measure as it would most likely be applied by researchers.[2] The first is associated with the construction of the measure's Herfindahl component from *Census* counts of plants by size category. The second concerns the degree to which the index is sensitive to the use of data for alternative spatial units. The use of micro data for virtually all manufacturing enterprises in the two states eliminates possible confounding effects from a third layer of potential error—the necessity, in most cases, of estimating suppressed cells in published *Census* data. My results indicate that while the Schmalensee proxy for the Herfindahl component is highly effective, changes in the index with the use of alternative spatial units of analysis can introduce non-trivial ambiguities in the usual application. Users should thus exercise considerable caution when employing the index in comparative studies of the geographic concentration of industry.

2. Data and procedures

I conduct two basic tests of the index in an application to three digit Standard Industrial Classification (SIC) manufacturing sectors. The first compares the agglomeration measure calculated with a true Herfindahl index (constructed as $\sum_k z_k^2 = H$, where z_k is the ratio of plant *k*'s employment to total industry employment as reported in the ES-202 data) with a measure computed using the Schmalensee proxy employed by Ellison and Glaeser.[3] The second test compares the agglomeration measure constructed with employment data aggregated at three progressively lower levels: commuter zones, counties, and zip codes.[4] If **(** is reasonably robust to changes in areal boundaries, its magnitude should vary little in a substantive sense when calculated with for data for zip codes, counties, or commuting sheds. The robustness of the index to changes in areal boundaries is particularly important if it is to be used for interregional and international comparisons.

In fact, however, the magnitude of the index is not likely to remain unchanged with changes in the geographic units of analysis. The model from which the measure is derived imposes heavy restrictions on the spatial force of natural advantages and spillovers: the former are independent and the latter are only realized if plants locate in the same region. In practice, there is little reason to believe that spillovers and natural advantages are contained within arbitrary administrative boundaries. Ellison and Glaeser therefore conjecture that the value of the index will likely increase with the level of spatial aggregation: “an estimate of (that is computed from county-level data (and hence reflects only the added probability with which pairs of plants locate in the same county) would be expected to be smaller than an estimate that is computed from state-level data and reflects the additional colocations due to spillovers felt at some distance and to correlated natural advantages (1997, p. 901).”[5]

With this in mind, I compare—within each state and between the two states—differences in index magnitude as well as changes in orderings that arise with the use of data for the three levels of aggregation. That permits an assessment of the degree to which meaningful comparisons of industrial agglomeration between regions can be made with (. Even though the magnitude of the agglomeration measure may increase with area aggregation, we should be able to make the same qualitative conclusions regarding the relative degree of concentration of a given sector in the two comparison states, regardless of the level of aggregation in the data. That simply represents what might be described as a kind of areal unit transitivity:

$${}_a\gamma_j^m > {}_b\gamma_j^m \Rightarrow {}_a\gamma_j^n > {}_b\gamma_j^n \qquad (2)$$

where *m* and *n* index different levels of data aggregation, *j* indexes the industry, and *a* and *b* index the study regions. The expression (2) is actually a modest criterion. I describe a more realistic one below.

3. Results

Although any positive value of the concentration measure (indicates concentration associated with spillovers and/or natural advantage, Ellison and Glaeser emphasize that values of (above

0.05 indicate a high degree of concentration while those below 0.02 indicate only minor concentration.[6] For the 459 four digit U.S. manufacturing sectors, they find a mean (of 0.051 (see Table 1). In only thirteen industries are plants more evenly distributed than would be expected at random. While the prevalence of concentration is striking, the median (is 0.026 and, in roughly half of the sectors, (is less than 0.024 (the index ranges from a low of -0.013 to a high of 0.630). Overall, the degree of concentration among three digit industries within both Tennessee and North Carolina is much lower. Mean concentration levels are below zero in both states while the ranges of values taken by the index are considerably wider. The number of sectors with concentration values below 0.024 are 81 (of 134) in Tennessee and 73 (of 128) in North Carolina.[7] That finding is fully consistent with the notion that the index should increase with the level of spatial aggregation.

3.1 Comparison of Herfindahl Indices

Using H' to denote the Ellison-Glaeser proxy for industry j's plant Herfindahl, I use a simple measure of relative difference $(D = |H' - H|/H)$ to compare the 'estimated' measure with an 'actual' Herfindahl calculated from establishment level data. High values of *D* may generate errors in (.

In fact, a comparison of the estimated and actual Herfindahl indices for three-digit industries in the two states strongly supports Schmalensee's original finding that the assumption of a uniform distribution of plants over given size categories offers an effective proxy for the true establishment size distribution. On average, the two indices differ by only 2.1 percent in Tennessee and 1.4 percent in North Carolina (see Table 2). In both states, the actual and estimated Herfindahls differ by 3 percent or less in four-fifths of the industries. The largest differences are for the motor vehicles and equipment (SIC 371) and jewelry, silverware, and plated ware sectors (SIC 391) in Tennessee (D = 0.32 and 0.17, respectively).

Although H' closely approximates the actual size distribution, small errors in H can, in principle, lead to large shifts in the concentration measure. Also, it is not necessarily the case that the largest errors in H will generate the largest errors in γ. Because of the wide range of index values across sectors, Table 3 characterizes the distributions of absolute (rather than relative) differences $D^{\gamma} = \gamma - \gamma'$ where γ' is the agglomeration index γ calculated with the estimated Herfindahl. Given the very low mean and median differences, the errors in the Herfindahl measure are clearly limited in their impact on the concentration measure. That is further confirmed by inspecting each industry in each state for qualitative changes in γ. For example, setting $\gamma \geq 0.020$ as the threshold at which an industry is considered concentrated, in only two Tennessee sectors and one North Carolina sector do findings change with the use of the estimated Herfindahl (Table 4).

Two of three of the sectors in Table 4 possess relatively high Herfindahl's. Ellison and Glaeser report that a simulation-based estimated standard error for their concentration index depends primarily on the value taken by H', finding that when H' is large (e.g., exceeding 0.1), the errors in γ can be substantial. One could argue that when most employment in a given industry is concentrated in one or a few plants (implying a high value for H'), the concept of spatial concentration loses its practical meaning (Ellison and Glaeser 1997). If we assume, on that basis, that the concentration indexes for both Tennessee sectors in the table are unreliable, we are left with only one sector–SIC 281 in North Carolina–that yields a different qualitative finding when the estimated Herfindahl is used. Even then, the difference between the two concentration measures for the sector–0.022 and 0.017–is too small to be meaningful. We can therefore conclude that the estimated Herfindahl is not a significant source of error in the index (in the present samples.

3.2 Spatial Aggregation of the Data

I next use employment data for three different levels of spatial aggregation to evaluate the empirical properties of the agglomeration measure three ways: 1) the ordering of index magnitudes across spatial units of analysis in each state; 2) the ordering of sectors according to index magnitudes (most to least concentrated) within each state; and 3) industry by industry comparisons of concentration between states.

Changes in index orderings by sector. Columns two and four in Table 5 report the share of sectors within each state for which the agglomeration index (follows one of six possible orderings. Twenty-four percent of sectors in Tennessee and 29 percent in North Carolina follow the pattern predicted by Ellison and Glaeser: a progressive increase in the magnitude of the indicator as employment data are aggregated from lower (zip codes) to higher (commuter zones) levels. The concentration measure for many sectors (30 percent in each state) follows the opposite pattern, highest for zip codes and lowest for commuter zones. And roughly another one-third of the industries' index in each state follows one of four other possible orderings. If we restrict the comparison to those sectors where *H*// < 0.1, we observe a result slightly more consistent with Ellison and Glaeser's prediction: the index increases with areal aggregation in 36 and 53 percent of sectors in Tennessee and North Carolina, respectively.

Ellison and Glaeser also construct a closely related *co*-agglomeration index, $(^c$, for examining the degree of joint concentration among a set of *r* sectors:

$$\gamma^{c} \equiv \frac{\left[\sum_i (s_i - x_i)^2 \Big/ \left(1 - \sum_i x_i^2\right)\right] - H^c - \sum_{j=1}^{r} \hat{\gamma}_j w_j^2 (1 - H_j)}{1 - \sum_{j=1}^{r} w_j^2} \tag{3}$$

where s_i is redefined as area *i*'s share of total employment in a group of r industries, H^c is the aggregate Herfindahl for the group $(H^c = \sum_j w_j^2 H_j)$, w_j is the *j*th industry's share of aggregate employment in the *r* industries, H_j is the *j*th industry's Herfindahl index, and $\hat{\gamma}_j$ is the value of the concentration index (1) for industry j. It is worthwhile examining whether the values co-agglomeration measure are more consistent across spatial units than those for **(**.

But first note that Ellison and Glaeser suggest a rescaling of **(**^{c} to ease interpretation:

$$\lambda \equiv \frac{\gamma^{c}}{\sum_j w_j \hat{\gamma}_j} \equiv \frac{\gamma^{c}}{\gamma^{r}} \tag{4}$$

In principle, a value of **8**=0 means that spillovers and/or natural advantages for a group of *r* industries are sector specific, i.e., agglomeration within the group is driven by tendencies for plants within specific sectors to co-locate geographically. Alternatively, **8**=1 implies that any natural advantages are perfectly correlated for all industries and spillovers between firms are not related to industry. In practice, the interpretation of λ is ambiguous because of the possibility of negative values in the numerator or denominator. Either the numerator or denominator in (4) must therefore be reported in addition to λ. Since the value of **(**^{r} (the weighted average of estimated concentration for *j* industries in the group) is a useful in itself, I report both **(**^{r} and **(**^{c}.

Co-agglomeration indices for all two-digit level SIC sectors in the two states, where the sectors *j* in (3) are the three-digit level SIC industries that comprise each two-digit SIC sector, are listed in Tables 6 and 7. Evidence of co-agglomeration clearly increases with spatial aggregation. No $(^c$ exceeds 0.02 in either state when employment data are aggregated to zip codes. At the county level, $(^c$ exceeds 0.02 in seven of twenty sectors in Tennessee and one sector in North Carolina. Eight sectors in Tennessee and four sectors in North Carolina show evidence of co-agglomeration among sub-industries at the commuter zone level. Yet, like (, there are still several two-digit sectors (one in Tennessee, three in North Carolina) that fail to follow a pattern of progressive increase in index magnitude with increases in the size of the areal unit.

Orderings within each state. Table 8 reports the value of (for commuter zones, counties, and zip codes for industries that satisfy two criteria: 1) a value of *H*/ of less than or equal to 0.1 in both states; and 2) a value of (exceeding 0.02 for at least one of the three areal units of analysis in one or both of the states.[8] The table indicates that those sectors with the highest concentration on one level of analysis (e.g., commuter zones) tend to be among those with high levels of concentration with other units of analysis (e.g., zip codes). But, again, there are some non-trivial exceptions. In Tennessee they include the meat products (SIC 201, the sixth most concentrated industry with zip codes is not concentrated when commuter zones or counties are used as the units of analysis) and girls’ and children’s outerwear (SIC 236; among the least concentrated sectors with commuter zones and the most concentrated with counties). Similar patterns are observed for the North Carolina electronic components (SIC 367) and commercial printing (SIC 275) sectors.

Changes in orderings between states. Orderings between states are also relatively consistent, but again with a few important exceptions. In four out of twenty-four sectors in Table 8, the state registering the highest level of concentration changes to a substantive degree with at least one change in level of data aggregation. In two of those cases, two shifts occur from the highest to lowest level of aggregation. An example is yarn and thread mills (SIC 228). One would conclude SIC 228 is more concentrated in Tennessee with commuter zone data, North Carolina with county data, and Tennessee with zip code data.

4. Assessment

Based on Table 8, if analysts restrict their between-state comparisons to sectors with estimated Herfindahls below 0.1 and concentration index values above 0.02, they will satisfy the weak areal transitivity criterion in expression (2) roughly 80 percent of the time. That is, two analysts using different spatial units of analysis are likely to come to the same conclusion regarding the relative concentration of a given industry in the two states about four-fifths of the time. If sectors are not screened according to the Herfindahl and level of concentration, the rate at which the criterion is satisfied is much lower (approximately 50 percent).

But more realistic criterion of spatial robustness is the following:

$$_a\gamma_j^m > {}_b\gamma_j^m \Rightarrow {}_a\gamma_j^m > {}_b\gamma_j^n$$

In a two region example the index should retain its ordering with changes in the spatial units of analysis in *one* (not both) of the study regions. Such a case more closely approximates what we require in practice, since types of administrative areas at any level (metropolitan areas, counties, zip codes, etc.) are not necessarily consistent in average size within countries or states much less between them (witness counties in California versus Virginia). In only two cases in Table 8 is the criterion satisfied (SIC 204 and SIC 267). Thus in the usual application, the results of a comparison of relative concentration across regions will very much depend on the spatial units of analysis, even if analysts restrict their comparisons to concentrated sectors with relatively low estimated Herfindahl's.

But how do those findings square with the theoretical independence of the agglomeration measures (in expected value) from the level of geographic concentration, as demonstrated by the authors? The source of the problem is the conceptualization of the geographic spread of natural advantages and spillovers. The location model and associated indices effectively view continuous spatial phenomena (spillovers and natural advantages) as discrete by confining those phenomena to arbitrarily defined areas. Furthermore, the model ignores potential interactions between those areas (or discrete units). In this sense, the Ellison-Glaeser indices are no worse than similar Gini- and location quotient-based concentration measures. Indeed, they are probably better because they account for plant size differences and their derivation from a clear model of firm location aids interpretation. But like Gini and other measures subject to the modifiable areal unit problem, they cannot be used without a high degree of caution as well as sensitivity testing with alternative spatial units of analysis.[9]

In the final analysis, the Ellison-Glaeser measures are probably most effectively used for exploring the unique spatial characteristics of natural advantages and spillovers for given industries in particular regions, rather than comparing relative industrial concentration across regions per se (as some researchers may attempt to do). The authors suggest, for example, that spillovers in some industries in some regions may be tightly confined in space (e.g., exhibited by a steep distance decay function), while in other sectors in other locations, spillovers might exhibit a broader ‘spatial reach’ (e.g., a gradual distance decay function). Use of the measures would focus on explaining shifts in index magnitude at alternative spatial scales rather than making hard and fast comparisons of the level of geographic concentration in different regions or countries.

Notes

1. I am grateful to Glenn Ellison for helpful comments on an earlier version of this paper. The usual disclaimer applies.
2. The ES-202 data, maintained by each state’s respective labor department, reports monthly employment for any business establishment covered by employment security law (as part of the Covered Wages and Employment program of US Bureau of Labor Statistics). The data constitute some 96 percent of enterprises in the two states, a share that is probably higher for the manufacturing sector.
3. For the Herfindahl proxy, employment shares are estimated from plant count and employment data for the ten establishment size categories reported in the *U.S. Census of Manufactures* by assuming a uniform distribution of plant sizes over each size range. The distribution is centered on the mean, bounded by the closest endpoint of the size range.
4. Commuter zones are aggregations of counties based on commuter flows reported in the 1990 *Census of Population* (developed by the Economic Research Services of the U.S. Department of Agriculture). There are 25 (21) commuter zones, 95 (100) counties, and 616

(641) zip codes in Tennessee (North Carolina). Each type of spatial unit is of similar average size in the two states. Boundary maps for each state are available on request.

5. This in itself implies the measure is not, strictly speaking, robust to spatial aggregation. The purpose of this note is to assess the degree of robustness.

6. I refer here only to the index as calculated using commuter sheds as the units of analysis for North Carolina and Tennessee and states for the U.S. (as reported in Ellison and Glaeser).

7. While there are 140 three-digit SIC sectors in U.S. manufacturing, there are one or fewer establishments in several sectors in the two states.

8. Values of (for all 134 sectors in Tennessee and all 128 sectors in North Carolina are available on request.

9. The modifiable areal unit problem is a longstanding issue in geography and regional science (see Cressie 1991).

References

Cressie, N. A. C. 1991. *Statistics for Spatial Data*. Chichester: John Wiley.

Ellison, G., and E. L. Glaeser. 1997. Geographic concentration in U.S. manufacturing industries: A dartboard approach. *Journal of Political Economy* 105: 889-927.

Enright, M. J. 1990. Geographic concentration and industrial organization. Unpublished Ph.D. dissertation, Harvard University.

Isard, W., I. J. Azis, M. P. Drennan, R. E. Miller, S. Saltzman, and E. Thorbecke. 1998. *Methods of Interregional and Regional Analysis*. Brookfield, VT: Ashgate.

Krugman, P. 1991. *Geography and Trade*. Cambridge, MA: MIT Press.

Maurel, F., and B. Sédillot. 1999. A measure of geographic concentration in french manufacturing industries. *Regional Science and Urban Economics* 29: 575-604.

Schmalensee, R. 1977. Using the *H*-index of concentration with published data. *Review of Economics and Statistics* 59: 186-93.

Table 1
Descriptive statistics, (

	United States	Tennessee	North Carolina
Mean	0.051	-0.087	-0.036
Median	0.025	0.009	0.014
Variance	0.006	0.689	0.252
Low	-0.013	-8.478	-5.313
High	0.630	0.447	0.477

Note: U.S. results are for 459 4-digit sectors calculated for states as reported by Ellison and Glaeser (1997). Tennessee and North Carolina results are for 3-digit sectors calculated for commuter zones.

Table 2
Herfindahl relative difference, *D*

	Tennessee	North Carolina
Mean	0.021	0.014
Median	0.009	0.008
High	0.328	0.081

Range	Share of sectors with relative difference between actual and estimated *H* in range	
.00	.02	.05
.00-.01	.37	.32
.01-.02	.30	.36
.02-.03	.11	.11
.03-.04	.05	.04
.04-.05	.05	.04
>=.05	.10	.08

Table 3
Absolute difference, *D*(

	Tennessee	North Carolina
Mean	0.001	0.003
Median	0.001	0.000
High	0.149	0.044
Range	Number of sectors with difference between (and (N in range	
.00	46	67
.00-.005	63	40
.005-.01	10	10
.01-.02	3	6
.02-.03	6	2
>=.03	6	1

Table 4
Qualitative shifts in findings with estimated Herfindahl

Sector / State		*H*	*H*N	(	(N
333 / TN	Primary nonferrous metals	0.541	0.504	-0.026	0.052
232 / TN	Men's and boys' furnishings	0.230	0.240	0.028	0.015
281 / NC	Industrial inorganic chemicals	0.088	0.093	0.022	0.017

SICs 232 in Tennessee and 281 in North Carolina are considered highly concentrated when true Herfindahl is used but fall under threshold with estimated Herfindahl. The opposite is true for SIC 333 in Tennessee.

Table 5
Differences in γ with alternative areal units

	Share of sectors			
	Tennessee		North Carolina	
Ordering of values	All sectors	HN<0.1	All sectors	HN<0.1
$\gamma^{cz} > \gamma^{c} > \gamma^{z}$	0.24	0.36	0.29	0.53
$\gamma^{z} > \gamma^{cz} > \gamma^{c}$	0.13	0.00	0.00	0.00
$\gamma^{c} > \gamma^{z} > \gamma^{cz}$	0.06	0.04	0.09	0.09
$\gamma^{z} > \gamma^{c} > \gamma^{cz}$	0.30	0.18	0.30	0.14
$\gamma^{c} > \gamma^{cz} > \gamma^{z}$	0.15	0.20	0.16	0.14
$\gamma^{cz} > \gamma^{z} > \gamma^{c}$	0.13	0.22	0.16	0.10

Table 6
Co-agglomeration index, 2-digit industries, Tennessee

		Commuter zones		Counties		Zip codes	
SIC	Industry	γ^{c}	γ^{r}	γ^{c}	γ^{r}	γ^{c}	γ^{r}
20	Food and kindred products	0.030	0.028	0.028	0.025	0.005	0.026
21	Tobacco products	0.258	-0.002	0.007	-0.092	-0.128	0.432
22	Textile mill products	0.044	0.046	0.027	0.023	0.001	0.005
23	Apparel and other textile products	0.018	0.031	0.013	0.016	0.003	0.027
24	Lumber and wood products	0.002	0.000	0.001	0.004	-0.009	0.049
25	Furniture and fixtures	0.049	0.075	0.031	0.035	0.006	0.018
26	Paper and allied products	0.072	0.024	0.065	0.022	-0.009	0.042
27	Printing and publishing	0.062	0.049	0.057	0.033	0.001	0.013
28	Chemicals and allied products	0.015	0.002	0.010	-0.003	0.009	0.126
29	Petroleum and coal products	0.256	0.011	0.235	0.017	-0.010	0.065
30	Rubber and misc. plastics products	0.003	-0.003	-0.001	-0.007	0.004	0.011
31	Leather and leather products	0.056	0.028	0.020	0.021	0.018	0.011
32	Stone, clay, and glass products	0.007	0.005	0.001	-0.007	-0.012	0.068
33	Primary metal industries	0.005	0.119	0.011	0.089	0.016	0.056
34	Fabricated metal products	0.016	-0.010	0.003	-0.002	0.003	-0.004
35	Industrial machinery and equipment	0.006	0.004	0.002	0.006	-0.004	0.012
36	Electronic and other electric equipment	0.015	0.004	-0.003	0.017	0.007	0.014
37	Transportation equipment	n/a	n/a	n/a	n/a	n/a	n/a
38	Instruments and related products	0.020	0.046	0.006	0.044	0.001	0.010
39	Miscellaneous manufacturing industries	0.032	0.006	0.000	0.042	-0.005	0.027
	Mean	0.051	0.024	0.027	0.015	-0.005	0.053
	Median	0.020	0.011	0.010	0.017	0.001	0.026
	High	0.258	0.119	0.235	0.089	0.018	0.432
	Low	0.002	-0.010	-0.003	-0.092	-0.128	-0.004

Note: Co-agglomeration measure is calculated for three digit sub-industries of each two digit sector. SIC 37 in Tennessee was comprised of only one industry (SIC 371) in the study period.

Table 7
Co-agglomeration index, 2-digit industries, North Carolina

		Commuter zones		Counties		Zip codes	
SIC	Industry	γ^c	γ^r	γ^c	γ^r	γ^c	γ^r
20	Food and kindred products	0.021	0.040	0.005	0.018	0.002	-0.005
21	Tobacco products	0.018	-0.065	0.011	-0.040	0.005	0.013
22	Textile mill products	0.009	0.023	0.005	0.020	0.001	0.006
23	Apparel and other textile products	-0.001	0.019	0.005	0.014	0.000	0.005
24	Lumber and wood products	0.000	0.030	0.009	0.020	0.002	0.008
25	Furniture and fixtures	0.035	0.162	0.001	0.064	0.002	0.030
26	Paper and allied products	0.002	0.007	-0.001	0.007	-0.001	-0.001
27	Printing and publishing	0.028	0.026	0.043	0.032	0.004	0.004
28	Chemicals and allied products	0.015	0.047	-0.002	0.021	-0.003	-0.002
29	Petroleum and coal products	-0.101	-0.158	-0.015	-0.007	-0.006	-0.013
30	Rubber and misc. plastics products	0.014	-0.021	0.007	-0.014	-0.001	-0.006
31	Leather and leather products	0.007	-0.064	0.018	-0.008	-0.008	-0.016
32	Stone, clay, and glass products	-0.007	-0.160	0.005	-0.020	0.000	-0.018
33	Primary metal industries	0.003	0.026	-0.008	0.085	-0.001	0.006
34	Fabricated metal products	0.002	-0.006	0.005	-0.004	0.001	-0.001
35	Industrial machinery and equipment	-0.006	-0.027	-0.001	-0.016	-0.006	-0.021
36	Electronic and other electric equipment	0.040	0.086	0.004	0.036	-0.002	0.001
37	Transportation equipment	0.002	0.037	0.002	0.015	0.003	-0.002
38	Instruments and related products	0.000	0.058	0.002	0.066	0.011	0.018
39	Miscellaneous manufacturing industries	0.006	0.008	0.002	0.003	-0.003	-0.005
	Mean	0.004	0.003	0.005	0.015	0.000	0.000
	Median	0.005	0.021	0.004	0.015	0.000	-0.001
	High	0.040	0.162	0.043	0.085	0.011	0.030
	Low	-0.101	-0.160	-0.015	-0.040	-0.008	-0.021

Note: Co-agglomeration measure is calculated for three digit sub-industries of each two digit sector.

Table 8

Agglomeration index by sector, Tennessee and North Carolina

(where *H/*<0.10 and (>0.02 for at least one spatial unit type)

		Tennessee				North Carolina			
			(				(		
SIC	Industry	*H/*	CZ	Cty	Zip	*H/*	CZ	Cty	Zip
201	Meat products	0.085	-0.001	-0.001	0.022	0.041	0.049	0.019	-0.002
204	Grain mill products	0.078	0.069	0.072	0.052	0.060	0.002	-0.001	-0.005
225	Knitting mills	0.049	0.067	0.030	0.019	0.006	0.040	0.018	0.009
228	Yarn and thread mills	0.080	0.112	0.036	0.010	0.013	0.055	0.053	0.001
232	Men's and boys' furnishings	0.016	0.037	0.022	0.006	0.020	0.027	0.009	-0.002
233	Women's and misses' outerwear	0.043	0.023	0.012	0.082	0.026	0.038	0.018	0.002
236	Girl's and children's outerwear	0.080	0.018	0.034	0.038	0.031	0.104	0.079	0.037
238	Misc. apparel and accessories	0.092	0.041	0.006	0.007	0.044	0.008	0.022	-0.005
239	Misc. fabricated textile products	0.042	0.037	-0.001	0.018	0.020	-0.005	0.008	0.004
241	Logging	0.008	0.027	0.010	0.005	0.006	0.081	0.031	0.006
242	Sawmills and planing mills	0.020	0.009	0.003	0.005	0.006	0.034	0.019	0.006
251	Household furniture	0.045	0.076	0.037	0.008	0.006	0.178	0.067	0.025
267	Misc. converted paper products	0.071	0.066	0.055	0.024	0.034	0.013	0.008	0.000
275	Commercial printing	0.030	0.048	0.031	0.003	0.009	0.023	0.031	0.003
279	Printing trade services	0.048	0.219	0.185	0.025	0.054	0.069	0.082	-0.001
327	Concrete, gypsum, plaster products	0.012	0.023	0.005	0.003	0.009	0.013	0.010	0.001
342	Cutlery, handtools, and hardware	0.093	0.035	-0.021	-0.008	0.052	-0.026	-0.003	0.002
344	Fabricated structural metal products	0.019	0.045	0.023	-0.004	0.017	0.006	0.002	0.001
349	Misc. fabricated metal products	0.027	0.024	0.003	0.001	0.025	-0.013	-0.004	-0.004
353	Construction and related machinery	0.085	0.129	0.133	0.006	0.058	0.029	0.003	0.000
355	Special industry machinery	0.058	-0.003	0.000	-0.016	0.014	0.061	0.069	0.007
356	General industrial machinery	0.049	-0.002	-0.004	0.005	0.029	0.028	0.012	0.004
362	Electrical industrial apparatus	0.082	-0.009	0.002	-0.002	0.063	0.064	0.038	0.006
367	Electronic components and accessories	0.088	-0.032	-0.011	0.001	0.037	0.091	0.015	-0.002