# WORKING PAPERS

## RESEARCH DEPARTMENT

**WORKING PAPER NO. 03-23**
**AN ALTERNATIVE DEFINITION OF ECONOMIC REGIONS**
**IN THE U.S. BASED ON SIMILARITIES**
**IN STATE BUSINESS CYCLES**

Theodore M. Crone
Federal Reserve Bank of Philadelphia

October 2003

# FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall, Philadelphia, PA 19106-1574• (215) 574-6428• www.phil.frb.org

# An Alternative Definition of Economic Regions in the U.S. Based on Similarities in State Business Cycles

Theodore M. Crone

Federal Reserve Bank of Philadelphia

October 2003

# Abstract

Since the 1950s the Bureau of Economic analysis (BEA) has grouped the states into eight regions based primarily on cross-sectional similarities in their socioeconomic characteristics. This is the most frequently used grouping of states in the U.S. for economic analysis. Since several recent studies concentrate on similarities and differences in regional business cycles, this paper groups states into regions based not on a broad set of socioeconomic characteristics but on the similarities in their business cycles. The analysis makes use of a consistent set of coincident indexes estimated from a Stock and Watson-type model. We applied k-means cluster analysis to the cyclical components of these indexes to group the 48 contiguous states into eight regions with similar cycles. Having grouped the states into regions, we determine the relative strength of cohesion among the states in the various regions. Finally, we compare the regions defined in this paper with the BEA regions.

## An Alternative Definition of Economic Regions in the U.S.
## Based on Similarities in State Business Cycles

### I. Introduction

Since the mid-19th century the Bureau of the Census has aggregated state data into multi-state regions (Bureau of the Census 1994). Currently the Bureau divides the 50 states into four regions that are further subdivided into nine divisions. Except for the addition of Alaska and Hawaii to the Pacific division in the 1950s, the composition of the nine census divisions has remained unchanged since 1910. After the 1950 census, however, an interagency committee within the Department of Commerce reviewed the definition of census regions and divisions in an effort to identify six to 12 groups of contiguous states based on their socioeconomic homogeneity. This review resulted in several suggestions for the regrouping of states, but the proposed changes were never adopted by the Census Bureau because of the lack of acceptance by the data users.

Despite the Census Bureau's rejection, the Bureau of Economic Analysis (BEA) adopted one of the committee's proposed groupings of the states as its definition of multi-state regions with one modification.[1] This division of the states into the eight BEA regions has not been adjusted since its introduction in the 1950s. (See the first two columns in Table 1 and Figure 1.) Except for the trend in per capita income (1929 to 1950), the economic variables used by the interagency committee describe a state's economic profile at a point in time. The other economic factors considered in defining regions were the composition of income by source in 1950, the level of per capita income in 1950, and the industrial composition of the workforce in 1950. The non-economic factors used to group the states included population density and growth, the racial and ethnic composition of the population, the infant mortality rate in 1949, and telephones per capita in 1950. The final grouping of states into regions was based on both the economic and non-economic factors.

Economists have tended to use the BEA regions for empirical analysis as the best generally accepted grouping of states into regions. And several recent studies of regional trends and cycles have been based on the BEA regions, even though the states were grouped primarily on the basis of similarities at a point in time. In the late 1970s and early 1980s BEA regions were used to examine the regional effects of monetary and fiscal policy (Toal, 1977; Garrison and Chang, 1979; and Mathur and Stein, 1980). Since the late 1990s another set of papers on the regional effects of monetary policy have appeared using the BEA regions (Carlino and DeFina, 1998; Kouparitsas, 2001; and Owyang and Wall, 2003). These papers use standard vector autoregression (VAR) models. Two papers, published in the 1990s, use VAR analysis to investigate the origin and propagation of regional income and employment shocks (Carlino and DeFina, 1995; and Clark, 1998).[2] Finally, two recent papers have examined the comovement of income across BEA regions (Carlino and Sill, 2001; and Croux, Forni, and Reichlin, 2001).

Since these articles focus on business-cycle phenomena, multi-state regions based on similarities at a point in time may not be the appropriate set of observations. In this paper we concentrate on economic homogeneity among the states and use the common patterns in the states' economies over business cycles as the criterion for grouping them into regions. This alternative definition of regions is likely to provide a better grouping of states for research on differences in cyclical behavior across regions.

The remainder of the paper is divided into five sections. Section II describes the construction of the composite state indexes that form the basis for grouping states into regions and the decomposition of the indexes into trend and cyclical components. Section III presents the results of k-means cluster analysis of the 48 contiguous states based on the cyclical component of the state indexes. Section IV discusses the cohesion indexes developed by Croux et al. for the newly defined regions. In section V the regions resulting from the cluster analysis are compared with the BEA regions. Section VI concludes the paper.

## II. Construction of State Coincident Indexes and Their Decomposition into Trend and Cyclical Components

In the 1940s the Department of Commerce began publishing three composite indexes for the national economy: the indexes of leading, lagging, and coincident indicators.[3] Of the three indexes, the composite index of coincident indicators is the most important for tracking the business cycle. This index is constructed from four monthly data series: the number of jobs in nonagricultural establishments, real personal income (minus transfer payments), the index of industrial production, and manufacturing and trade sales adjusted for inflation. While the composite index of coincident indicators has tracked national business cycles fairly well, it can be criticized for not being based on a formal mathematical or statistical model (Koopmans 1947).

In the late 1980s James Stock and Mark Watson (1989 and 1991) provided a statistical basis for a composite index and developed an alternative index of coincident indicators for the U.S. The Stock/Watson model is based on the assumption that the observed indicators of the economy reflect a single, unobserved dynamic factor—the underlying "state of the economy." The Kalman filter is used to estimate the unobserved common factor. The assumptions of the model are set out in the following sets of equations.

The measurement equations (for the observed variables, X):

$$) x_{t=}" + \$(L)) c_{t+}: {}_t \tag{1}$$

And the transition equations:

$$( (L)) c_t = * + O_t \tag{2}$$

$$D(L): {}_t = {}_{,t} \tag{3}$$

where

        $x_t$ = the log of each of the observed variables,

        $c_{t=}$ the log of the state variable to be estimated, and

        L denotes the lag operator.

Equation (2) represents the law of motion for the state variable $c_t$ (the unobserved, underlying

state of the economy), which follows an autoregressive process. The idiosyncratic components

of the measurement variables (: ) are assumed to be uncorrelated with one another and also

follow an autoregressive process.

        Equations (1) and (2) are estimated using the standardized log difference of the observed

indicators and of the state variable.[4] Thus, " and * do not have to be estimated, and the

procedure provides an estimate of the standardized log difference of the latent dynamic factor.

The index level is determined by setting the index to 100 at a given date and applying the

monthly changes estimated from the system of equations. An average monthly increase is

reintroduced in the index by adding the weighted average increase of the components over the

estimation period.[5] For their alternative index, Stock and Watson used the same component

series that were used in the Commerce Department's index with one exception. In place of

nonfarm employment, Stock and Watson used total hours worked in the nonagricultural sector.

        In a recent expansion of the Stock and Watson model, Alan Clayton-Matthews (2001)

developed a C++ program that allows for measurement equations in which some indicator

variables ($x_t$) are observed on a quarterly rather than a monthly basis. The measurement equation

for the quarterly variable becomes

        ) $x_{t=}$ " + \$(L)$\Omega$(L)) $c_{t+}$ : $_t$

Since "t" indexes months, we need to reinterpret $\Delta x_t$ for the quarterly variables, so that

$$\Delta x_t = x_t - x_{t-3}$$

Since $\Delta c_t$ is at a monthly frequency

$$\Omega(L) = 1 + 2L + 3L^2 + 2L^3 + L^4$$

Crone (2003) used the expanded version of the Stock/Watson model to produce a set of consistent indexes for the 50 states. Each state index was constructed from the same set of indicator variables: nonfarm employment, average hours worked in manufacturing, the unemployment rate, and quarterly real wage and salary income. The timing of the index for each state is set to coincide with the timing of employment by including only the contemporary value of the latent variable in the measurement equation for employment. Also the latent variable ($c_t$) follows an AR(2) process in the model for each of the states. For consistency among the states, the long-run increase in each state index was set equal to the long-run growth in the state's real gross state product (GSP), the most comprehensive measure of state output. The use of GSP to set the long-run growth in the state's index avoids the inconsistency of having long-term growth in the index determined by a combination of variables that are weighted differently from state to state.[6]

Since the time of Burns and Mitchell (1946) it has been widely recognized that a key characteristic of business cycles is the comovement of major economic variables. The Stock and Watson model estimates a single dynamic factor that captures the comovement of the component indicators (see Diebold and Rudebusch 1996). The final form of the state indexes, however, includes not only the cyclical movement in the state economies but also trend growth. To examine the commonality of business cycles across states, we must isolate the cyclical component from the trend component.

We use the band-pass filter developed by Baxter and King (1999) to decompose the log of the state indexes into trend, cyclical, and irregular components. We follow the standard assumption that the cyclical component includes movements in the index with frequencies between six quarters (18 months) and 32 quarters (96 months). Our indexes begin in 1979, and we use the data through 2002.[7] As one would expect, the traditional recessions and growth recessions in the national economy are reflected in the cyclical component of the state indexes. But there are also significant differences in the timing, depth, and number of downturns from state to state, so the identification of regions composed of states with similar business cycles is a potentially useful contribution to the regional literature. As an example of a state index and its components, Figures 2a-2c show the coincident index for New Jersey, its trend component, and its cyclical component.

### III. Defining Regions Using Cluster Analysis on the State Indexes

Having identified the cyclical components of the states' composite indexes, we need to employ some method of pattern recognition to group the 48 contiguous states by similar cycles. An obvious choice is cluster analysis, which can produce either hierarchical or partitional clusters. Since we do not posit any hierarchical relationship among the states in terms of business cycles, we chose to separate the states into non-hierarchical, partitional clusters. For a large number of observations (48 in our case) and any reasonable number of clusters, it is not feasible in practice to examine every possible set of clusters. The number of possible sets (S) of G non-empty groups of n observations is determined by the following formula:[8]

$$S(n,G) = \frac{1}{G!} \sum_{g=1}^{G} (-1)^{G-g} \left( \frac{G!}{g!(G-g)!} \right) g^n$$

For 48 states partitioned into eight non-empty clusters the number of possibilities is approximately $5.5 \times 10^{38}$. Given this number of potential sets of clusters, reallocation algorithms have been developed to identify the best set of clusters based on some predetermined criterion. We use the k-means partitional clustering method.

The k-means clustering technique separates the observations into a predetermined number of clusters (k) based on minimizing some measure of dissimilarity among the observations in each cluster. Our measure of the dissimilarity of observations in a cluster is the squared Euclidean distance from the center of the cluster (Gordon 1999).[9]

$$\sum_{m \in c} \sum_{i=1}^{N} (X_{im} - \overline{X_{ic}})^2$$

where

$X_{im}$ = variable i (i = 1 . . . N) for observation m (m = 1 . . . M), and

$\overline{X_{ic}}$ = the center of cluster c to which observation m is assigned or the average $X_i$ for all the observations in cluster c.  In practice, an iterative procedure is used to find the minimum squared distance from the cluster center for each observation (see Hartigan 1975). Initial cluster centers are designated, and observations are assigned to the respective clusters based on the minimum distance criterion. Alternatively, the observations can be randomly assigned to the predetermined number of clusters and the initial centers are calculated based on that random assignment. Once all the observations have been assigned to a cluster and the initial cluster centers have been calculated, observations are reassigned based on the minimum squared distance from those initial centers. New cluster centers are then calculated based on the

reassignment. This process is repeated until no observation is reassigned during an iteration. The cluster program is then terminated, and the final clusters are established.

There are two major indeterminacies with k-means clustering. First, there is no accepted criterion for determining the optimal number of clusters. Second, the final clustering depends on the designation of the initial centers. The local minimum determined by the iterative procedure is not necessarily a global minimum.

Since our objective is to define a set of eight regions composed of contiguous states to compare with the BEA regions, we limited the number of predetermined clusters (k) to eight or fewer. A region will consist of two or more contiguous states that belong to the same cluster even if they are geographically separated from other states in that cluster. Therefore, it is possible for the k-means algorithm based on fewer than eight clusters to produce eight regions that contain all 48 contiguous states. We make a distinction between "clusters," or the groups of states resulting from the application of the k-means algorithm, and "regions," or groups of two or more contiguous states belonging to the same cluster. In the case of eight predetermined clusters and no stand-alone states, each region will constitute a distinct cluster.[10] In the case of a predetermined number of clusters that is less than eight, the states in two or more regions must belong to the same cluster, i.e., they must have similar business cycles. In our search for a set of eight regions that contained all 48 contiguous states, we repeated the exercise described below three times with a predetermined number of clusters (k) from six to eight.

For each predetermined number of clusters, we used the random selection option to choose the initial clusters and their centers. All the observations in the data set were randomly assigned to clusters whose centers were then calculated. Using data sets with known clusters, Peña et al. (1999) show that this random initialization outperforms other initialization methods in producing the correct clusters. The proper clustering is more likely to result with random initialization, but it is not a *guaranteed* outcome. And there is no generally accepted statistic to

gauge the significance of the outcome of a *single* clustering exercise based on random initialization.

To select the best set of clusters from a number of clustering exercises with different initial centers, we applied the k-means algorithm 10,000 times to our business-cycle data. We then searched over the resulting 10,000 sets of clusters to determine the set that minimized the average squared distance for the 48 states from their respective cluster centers or, equivalently, the set that minimized the sum of the squared distances.

$$Min \left\{ \sum_{c} \left[ \sum_{m \in c} \sum_{i=1}^{N} (X_{im} - \overline{X_{ic}})^2 \right] \right\} \qquad (4)$$

If we measure the squared differences of the states' business-cycle components from the center of their regions as defined by the BEA, the value of the sum in expression (4) is 1.668. To be considered better than the BEA definition, any alternative definition of regions must produce a value of this sum that is significantly lower than 1.668.

We began our search for an alternative definition of regions by running 10,000 iterations of the k-means clustering algorithm with eight as the predetermined number of clusters. We then identified the set of clusters that satisfied condition (4).[11] This set of clusters, shown in Figure 3, is unacceptable as an alternative definition of regions because it contains six stand-alone states. Michigan represents a single-state cluster. West Virginia and Nevada form a two-state cluster, but they are not contiguous. Minnesota, Mississippi, and Wyoming are not contiguous with any other states in their respective clusters. The other 42 states form eight regions of two or more contiguous states, and three of those regions have similar business cycles; i.e., the states in the three regions belong to the same cluster.

In defining *geographic regions* using cluster analysis, one can increase the compactness of the clusters and reduce the number of stand-alone states by including a set of proximity variables among the dimensions on which the states are clustered.[12] A simple set of proximity variables would be a set of binary variables, one for each state, in which the state and each of its neighbors is given a value of one and all the other states are given a value of zero. We introduce a more refined measure of proximity by including a variable for each state that reflects the similarity of its business cycle with the business cycle of each of its neighbors. From the 10,000 iterations based on random initial centers in the first stage of this analysis, we calculate the relative frequency of each pair of states clustering together. We then create a 48x48 matrix of proximity measures in which the proximity for neighboring states is measured by the relative frequency with which the two states cluster and the measure for non-neighboring states is zero.[13] We add these measures of proximity to the monthly business-cycle components from the state indexes and repeat the k-means clustering algorithm another 10,000 times with randomly selected initial centers.[14] Among these 10,000 sets of clusters we search over those that have only eight regions and no stand-alone states. Based on condition (4), we choose as best the one that minimizes the average squared distance of the business-cycle components of the 48 states from their respective cluster centers, which are also the centers of the regions in this case. The value of the sum in expression (4) for this set of regions is 1.492.[15] Based on a standard F-test, this is significantly lower (at the .01 level) than the sum of the squared distances from the centers for the BEA regions.[16]

We repeated this entire exercise with k equal to six and k equal to seven. We considered only those sets of clusters that resulted in eight regions and no stand-alone states. For k equal to six, the lowest value of the sum in expression (4) based on the regions' centers is 1.490. This is

10

not significantly different (at the .05 level) from the sum of squared differences when k is set equal to eight. For k equal to seven the lowest value of the sum in expression (4) based on the region's centers is 1.426. This is significantly different at the .01 level from the lowest values when k is set equal to six or eight. Among all the sets of clusters over which we searched, this set of regions satisfied condition (4) under the constraint that there be only eight regions and no stand-alone states. The proximity matrix that generated the clusters for k equal to seven is shown in Figure 4, and the eight regions that satisfy condition (4) are shown in Figure 5.

Table 1 compares the eight BEA regions with the eight regions defined by our cluster analysis. In two cases the regions identified by the cluster analysis are identical to the regions defined by the BEA: New England and the Mideast. The Southeast, Great Lakes, and Far West under the new definition are very similar to the BEA regions. Most of the changes occur in the BEA's Southwest, Plains, and Rocky Mountain regions.

## IV. Cohesion within the Regions

Having identified eight regions of contiguous states using cluster analysis, we calculate a cohesion index of state business cycles for each of the regions and compare it to the cohesion index for the corresponding BEA region. Croux et al. (2001) have developed a cohesion index for groups of two or more members. They define the dynamic correlation between two variables, x and y, at frequency $\lambda$ as:

$$\rho_{xy}(\lambda) = \frac{C_{xy}(\lambda)}{\sqrt{S_x(\lambda)S_y(\lambda)}}$$

where $S_x(\lambda)$ and $S_y(\lambda)$ are the spectral density functions of x and y and $C_{xy}(\lambda)$ is the cospectrum of x and y.

The dynamic correlation between x and y within the frequency band $\Lambda$ becomes

$$\rho_{xy}(\Lambda) = \frac{\int_\Lambda C_{xy}(\lambda)d\lambda}{\sqrt{\int_\Lambda S_x(\lambda)d\lambda \int_\Lambda S_y(\lambda)d\lambda}}$$

The authors demonstrate that this is identical to the static correlation between two series that have been properly pre-filtered with a suitable two-sided filter such as the band-pass filter. Thus the static correlation of the cyclical components of the state indexes corresponds to the dynamic correlation of the indexes in the frequency band of 18 to 96 months.

Croux et al. also propose a cohesion index for groups with more than two members. This cohesion index is a weighted average of the pairwise dynamic correlations for each possible pair in the group. We calculate a cohesion index for each of our eight regions by weighting the pairwise dynamic correlations of the business-cycle components in each region by the average gross state product in 2000 for the pair of states. To bound the cohesion indexes by −1 and +1 we adjust the weights in the group to sum to one. In our case, the formula for the cohesion index for each region is:

$$\frac{\sum_{i<j}[(GSP_i + GSP_j)\rho_{ij}]}{\sum_{i<j}(GSP_i + GSP_j)}$$

where the N members of any region are indexed

i and j = 1 . . . N

These indexes produce a relative measure of cohesion among the state business cycles in the regions. The cohesion indexes for both the BEA regions and the eight regions derived from our cluster analysis are reported in Table 1.[17]

The cohesion indexes for the eight regions resulting from the cluster analysis range from a high of 0.91 for New England to a low of 0.62 for the Mountain/Northern Plains region. A comparison of the cohesion indexes for the regions identified in this paper using cluster analysis with the indexes for the BEA regions generally confirms the superiority of the new set of regions. In general, our cluster-generated definition of regions improves the cohesion of states within the regions.

**V. A Comparison of the Regions Based on Cluster Analysis and the BEA Regions**

Table 1 compares the regions based on cluster analysis with the BEA regions state by state. The New England region is identical under the two definitions. The region accounts for 6 percent of the total GSP of the 48 contiguous states, and since it contains the same states under either definition, its cohesion index is the same in both cases (0.91). New England also has the highest cohesion index under either definition.

The Mideast region is also identical under the two definitions. It accounts for 18.2 percent of the GSP of the 48 contiguous states and has a cohesion index of .88.

The BEA's Southeast region is the largest of the BEA regions economically, accounting for 22.0 percent of the 48-state GSP. In our alternative definition of regions, two states in the Southeast are assigned to other regions. West Virginia is assigned to the Great Lakes region, and Louisiana to the Energy Belt. The reassignment of these two states raises the cohesion index for the Southeast from 0.74 under the BEA definition to 0.86 under our definition. The reassignment lowers the percentage of the 48-state GSP in the Southeast region to 20.1 percent.

Under the cluster-based definition, the Great Lakes region adds two states that are not included in the BEA definition: West Virginia and Minnesota. The inclusion of these two states lowers the cohesion index for the Great Lakes slightly from 0.83 to 0.81. The proportion of the 48-state GSP in the new Great Lakes region rises to 17.9 percent from 15.6 percent.

Three of the seven states in the BEA's Plains region are reassigned to neighboring regions in the cluster-based definition. Minnesota is assigned to the Great Lakes region, and North and South Dakota are assigned to the Mountain/Northern Plains region. The loss of these three states from the Plains region reduces the region's percentage of the 48-state GSP to 4.2 percent from 6.5 percent. The cohesion index for the Plains region increases from .73 to .83.

Two of the BEA's Rocky Mountain states combine with North and South Dakota to form the smallest region under the cluster-generated definition of regions (Mountain/Northern Plains). This region accounts for only 1.0 percent of the 48-state GSP. And its cohesion index is only 0.62, slightly lower than the cohesion index for the BEA's Rocky Mountain region (0.63).

Based on our cluster analysis, seven states stretching from Louisiana to Wyoming combine to form an Energy Belt region. For five of these states (Louisiana, Texas, Oklahoma,

New Mexico, and Wyoming), oil and gas production represents more than 5 percent of GSP. Three of the states (Texas, Oklahoma, and New Mexico) are part of the BEA's Southwest region, which accounts for 10.6 percent of the 48-state GSP and has a cohesion index of .55. The Energy Belt under the new definition of regions accounts for 13.2 percent of the 48-state GSP and has a much higher cohesion index of 0.72.

The Far West region under the cluster-based definition differs from the BEA's Far West region only by the addition of Arizona. This raises the percentage of the 48-state GSP in this region from 17.9 percent to 19.5 percent, and it lowers the cohesion index for the region slightly from 0.77 to 0.76.

A comparison of the cohesion indexes for the BEA regions and their counterparts under the cluster-based definition supports the use of the cluster-based definition for business-cycle analysis. Two regions (New England and the Mideast) are identical and, therefore, have the same cohesion indexes. Three regions (the Southeast, the Plains, and the Energy Belt) have cohesion indexes much higher than their BEA counterparts, with the differences ranging from 0.10 to 0.17. Three of the newly defined regions (the Great Lakes, Mountain/Northern Plains, and the Far West) have cohesion indexes slightly lower than their BEA counterparts with the differences ranging from 0.01 to 0.02.

**VI. Conclusions**

Recent research on regional business cycles has depended almost exclusively on the BEA grouping of states into regions. This definition of regions was based on several measures of socioeconomic homogeneity among the sates around 1950. We have proposed a new definition of US regions based on the similarity of state business cycles from 1979 to 2002.

15

We use the cyclical components of a newly developed set of consistent coincident indexes for the states as measures of the state business cycles. We apply k-means cluster analysis to these cyclical components to identify eight compact regions for the 48 contiguous states. Based on the squared distances of the state business-cycle components from their respective region centers, this new definition of regions is superior to the BEA's definition. There is also an improvement in the cohesion indexes for these newly defined regions compared to the indexes for the corresponding BEA regions.

The alternative definition of regions presented in this paper is theoretically a more appropriate set of regions to use in regional business-cycle analysis than the BEA definition because it is based on similarities in state business cycles. In practice, this alternative definition may also produce more significant results for the regional effects of fiscal and monetary policy, changes in exchange rates, or energy shocks because the states in the newly defined regions are more cohesive in terms of their business cycles.

**Endnotes**

[1]The modification was the combining of an Upper South and a Lower South region into one Southeast region.

[2] Clark uses census divisions rather than BEA regions in his analysis.

[3]In 1994, the Conference Board took over the production of these indexes.

[4] The average log difference over the sample period is subtracted from the log difference for each month, and the result is divided by the standard deviation of the log differences.

[5] The weights are determined by each component's contribution to the cyclical change in the composite index.

[6] Of course, one could set the long-run growth of the index to the long-run growth of one of the component variables, such as employment or real wages and salaries. But GSP comes closer to an output measure than any of the component variables. The consistent state indexes used in this paper differ from those used in Crone (1998/1999), which were composed of only the three monthly variables and whose average monthly increases were based on the weighted average of the monthly increases in the components.

[7] To obtain the cyclical component at the beginning and end of the series, we extend each state's index in both directions using an autoregressive model with 12 lags (Stock and Watson 1999).

[8] Jensen (1969), pp.1035-1036.

[9] Gordon (pp. 36-38) lists four other commonly used measures of dissimilarity.

[10] A stand-alone state is one that does not belong to the same cluster as any of its neighboring states.

[11] The value of expression (4) for this set of clusters is 0.95.

[12] See comments in Webster and Burrough, 1972. This will not guarantee compact clusters of geographically contiguous states.

[13] Abraham, Goetzmann, and Wachter (1994) and Goetzmann and Wachter (1995) group metro areas according to common features of their housing and commercial real estate markets based on the frequencies with which the metro areas cluster in repeated applications of the clustering algorithm.

[14] The 48 states are clustered on the basis of 330 variables, the 282 monthly business-cycle components (July 1979 through December 2002) and the 48 proximity variables.

[15] This is the 142nd lowest sum of the 10,000 sets of clusters generated by the k-means algorithm. The 141 sets of clusters with lower sums either produced more than eight regions or had one or more stand-alone states. In calculating the sums we used only the squared differences in the monthly business-cycle components, not the squared differences in the proximity variables.

[16] There are 13, 546 monthly business-cycle components for the 48 states in our sample, and the critical F-value for the sum of squared differences from the center at the 0.01 level is 1.0408. The value of the F-statistic for the difference between the sum for the BEA regions and the sum for the eight cluster regions is 1.118.

[17] We omit Alaska and Hawaii from the BEA's Far West region, since they are not included in our cluster analysis.

# References

Abraham, Jesse M., William M. Goetzmann, and Susan M. Wachter. "Homogeneous Groupings of Metropolitan Housing Markets," *Journal of Housing Economics*, 3 (1994), pp. 186-206.

Bureau of the Census. *Geographic Areas Reference Manual*. Washington: Department of Commerce, November 1994.

Baxter, Marianne, and Robert G. King. "Measuring Business Cycles: Appropriate Band-Pass Filters for Economic Time Series," *The Review of Economics and Statistics*, 81(1999), pp. 575-593.

Burns, Arthur F., and Wesley C. Mitchell. *Measuring Business Cycles,* NY: National Bureau of Economic Research, 1946.

Carlino, Gerald A., and Robert H. DeFina. "Regional Income Dynamics," *Journal of Urban Economics*. 37 (1995), pp. 88-106.

Carlino, Gerald A., and Robert H. DeFina. "The Differential Regional Effects of Monetary Policy," *The Review of Economics and Statistics*. 80 (1998) pp. 572-587.

Carlino, Gerald, and Keith Sill. "Regional Income Fluctuations: Common Trends and Common Cycles," *The Review of Economics and Statistics*. 83 (2001) pp. 446-456.

Clark, Todd E. "Employment Fluctuations in the US Regions and Industries: The Roles of National, Region-Specific, and Industry-Specific Shocks," *Journal of Labor Economics* 16(1998), pp. 202-229.

Clayton-Matthews, Alan. "DSFM Manual," mimeo, UMass Boston, April 2001.

Crone, Theodore M. "Consistent Economic Indexes for the 50 States," Federal Reserve Bank of Philadelphia, Working Paper 02-7R (June 2003).

Crone, Theodore M. "New Indexes Track the State of the States," *Business Review*, Federal Reserve Bank of Philadelphia, January/February 1994, pp. 19-31.

Crone, Theodore M. "Using State Indexes to Define Economic Regions in the US," *Journal of Economic and Social Measurement*. 25 (1998/1999), pp. 259-275.

Croux, Christophe, Mario Forni, and Lucrezia Reichlin. "A Measure of Comovement for Economic Variables: Theory and Empirics," *The Review of Economics and Statistics*, 83(2001), pp. 232-241.

Diebold, Francis X., and Glenn D. Rudebusch. "Measuring Business Cycles: A Modern Perspective," *The Review of Economics and Statistics*, 78(1996), pp. 67-77.

Garrison, Charles B. and Hui S. Chang. "The Effect of Monetary and Fiscal Policies on Regional Business Cycles," *International Regional Science Review*, 4 (1979), pp. 167-180.

Goetzmann, William N., and Susan M. Wachter. "Clustering Methods for Real Estate Portfolios," *Real Estate Economics*. 23 (1995),  pp. 271-310.

Gordon, A.D. *Classification*. 2nd edition, NY: Chapman & Hall, 1999.

Jensen, Robert E. "A Dynamic Programming Algorithm for Cluster Analysis," *Operations Research*, 17 (1969), pp. 1034-1057.

Hartigan, John A. *Clustering Algorithms*. NY: John Wiley, 1975.

Koopmans, Tjalling C. "Measurement Without Theory," *The Review of Economics and Statistics*, 29, 1940, pp. 351-394.

Kouparitsas, Michael A. "Is the United States and Optimum Currency Area? An Empirical Analysis of Regional Business Cycles," Working Paper 2001-22, Federal Reserve Bank of Chicago, 2001.

Mathur, Vijay K. and Sheldon Stein. "Regional Impact of Monetary and Fiscal Policy: An Investigation into the Reduced Form Approach," *Journal of Regional Science*, 20 (1980), pp. 343-351.

Owyang, Michael T., and Howard J. Wall. "Regional Disparities in the Transmission of Monetary Policy," Working Paper 2003-008A, Federal Reserve Bank of St. Louis, 2003.

Peña, J. M., J. A. Lozano, and P. Larrañaga. "An Empirical Comparison of Four Initialization Methods for the K-means Algorithm," *Pattern Recognition Letters*, 20(1999), pp. 1027-1040.

Stock, James H., and Mark W. Watson. "New Indexes of Coincident and Leading Economic Indicators," *NBER Macroeconomics Annual* (1989), pp. 351-94.

Stock, James H., and Mark W. Watson. "A Probability Model of the Coincident Economic Indicators," in *Leading Economic Indicators: New Approaches and Forecasting Records*. K. Lahiri and G.H. Moore, eds. Cambridge University Press, Cambridge, 1991, pp. 63-89.

Stock, James H., and Mark W. Watson. "Business Cycle Fluctuations in US Macroeconomic Time Series," in *Handbook of Macroeconomics*. Vol. IA, John B. Taylor and Michael Woodford, eds. Elsevier, New York: 1999, pp. 3-64.

Toal, William D. "Regional Impacts of Monetary and Fiscal Policies in the Postwar Period: Some Initial Tests," Federal Reserve Bank of Atlanta, Working Paper Series, June 1977.

Webster, R., and P. A. Burrough, "Computer-Based Soil Mapping of Small Areas from Sample Data," *Journal of Soil Science,* 23 (1972), pp. 222-234.

| Table 1 | | | | | |
|---|---|---|---|---|---|
| **BEA Regions**<br>**(Percent 48-state GSP)**<br>**CI = cohesion index** | | | **Cluster Regions**<br>**(Percent 48-state GSP)**<br>**CI = cohesion index** | | |
| **Region** | **State** | | **State** | **Region** | |
| New England (6.0%) CI = .91 | Maine | | Maine | New England (6.0%) CI = .91 | |
| | New Hampshire | | New Hampshire | | |
| | Vermont | | Vermont | | |
| | Massachusetts | | Massachusetts | | |
| | Rhode Island | | Rhode Island | | |
| | Connecticut | | Connecticut | | |
| | | | | | |
| Mideast (18.2%) CI = .88 | New York | | New York | Mideast (18.2%) CI = .88 | |
| | New Jersey | | New Jersey | | |
| | Pennsylvania | | Pennsylvania | | |
| | Delaware | | Delaware | | |
| | Maryland | | Maryland | | |
| | | | | | |
| Southeast (22.0%) CI = .74 | Virginia | | Virginia | Southeast (20.1%) CI = .86 | |
| | North Carolina | | North Carolina | | |
| | South Carolina | | South Carolina | | |
| | Georgia | | Georgia | | |
| | Florida | | Florida | | |
| | Kentucky | | Kentucky | | |
| | Tennessee | | Tennessee | | |
| | Alabama | | Alabama | | |
| | Mississippi | | Mississippi | | |
| | Arkansas | | Arkansas | | |
| | Louisiana | | | | |
| | West Virginia | | | | |
| | | | | | |
| | | | West Virginia | | |
| Great Lakes (15.6%) CI = .83 | Michigan | | Michigan | Great Lakes (17.9%) CI = .81 | |
| | Ohio | | Ohio | | |
| | Indiana | | Indiana | | |
| | Illinois | | Illinois | | |
| | Wisconsin | | Wisconsin | | |
| | | | Minnesota | | |

| BEA Regions (Percent 48-state GSP) CI = cohesion index | | Cluster Regions (Percent 48-state GSP) CI = cohesion index | |
|---|---|---|---|
| Plains (6.5%) CI = .73 | Minnesota | | |
| | Missouri | Missouri | Plains (4.2%) CI = .83 |
| | Kansas | Kansas | |
| | Nebraska | Nebraska | |
| | Iowa | Iowa | |
| | South Dakota | South Dakota | Mountain / Northern Plains (1.0%) CI = .62 |
| | North Dakota | North Dakota | |
| Rocky Mountain States (3.2%) CI = .63 | Montana | Montana | |
| | Idaho | Idaho | |
| | Wyoming | | |
| | Utah | | |
| | Colorado | | |
| | | Louisiana | Energy Belt (13.2%) CI = .72 |
| | | Wyoming | |
| | | Utah | |
| | | Colorado | |
| Southwest (10.6%) CI = .55 | Texas | Texas | |
| | Oklahoma | Oklahoma | |
| | New Mexico | New Mexico | |
| | Arizona | | |
| | | Arizona | Far West (19.5%) CI = .76 |
| Far West (17.9%) CI = .77 | California | California | |
| | Nevada | Nevada | |
| | Washington | Washington | |
| | Oregon | Oregon | |

# Figure 1

## BEA Regions
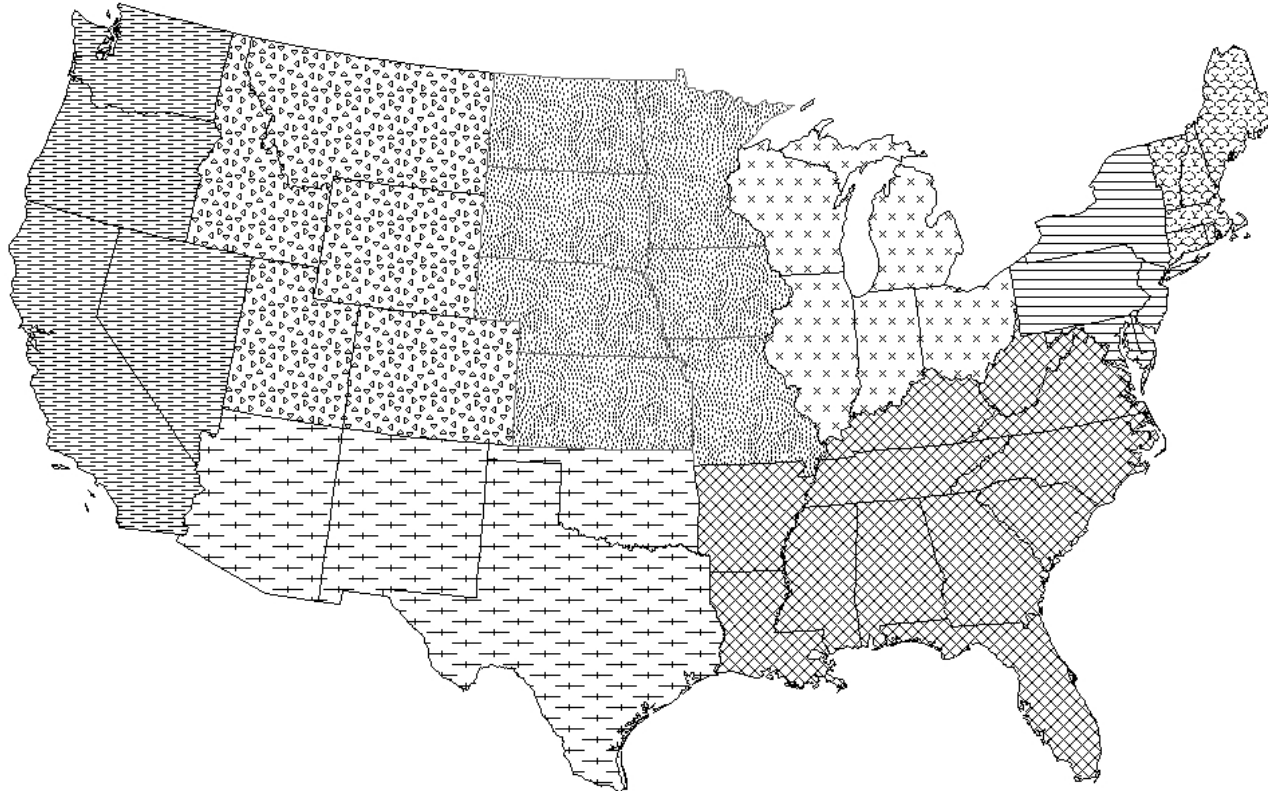
**Figure 2a**

**New Jersey Coincident Index**
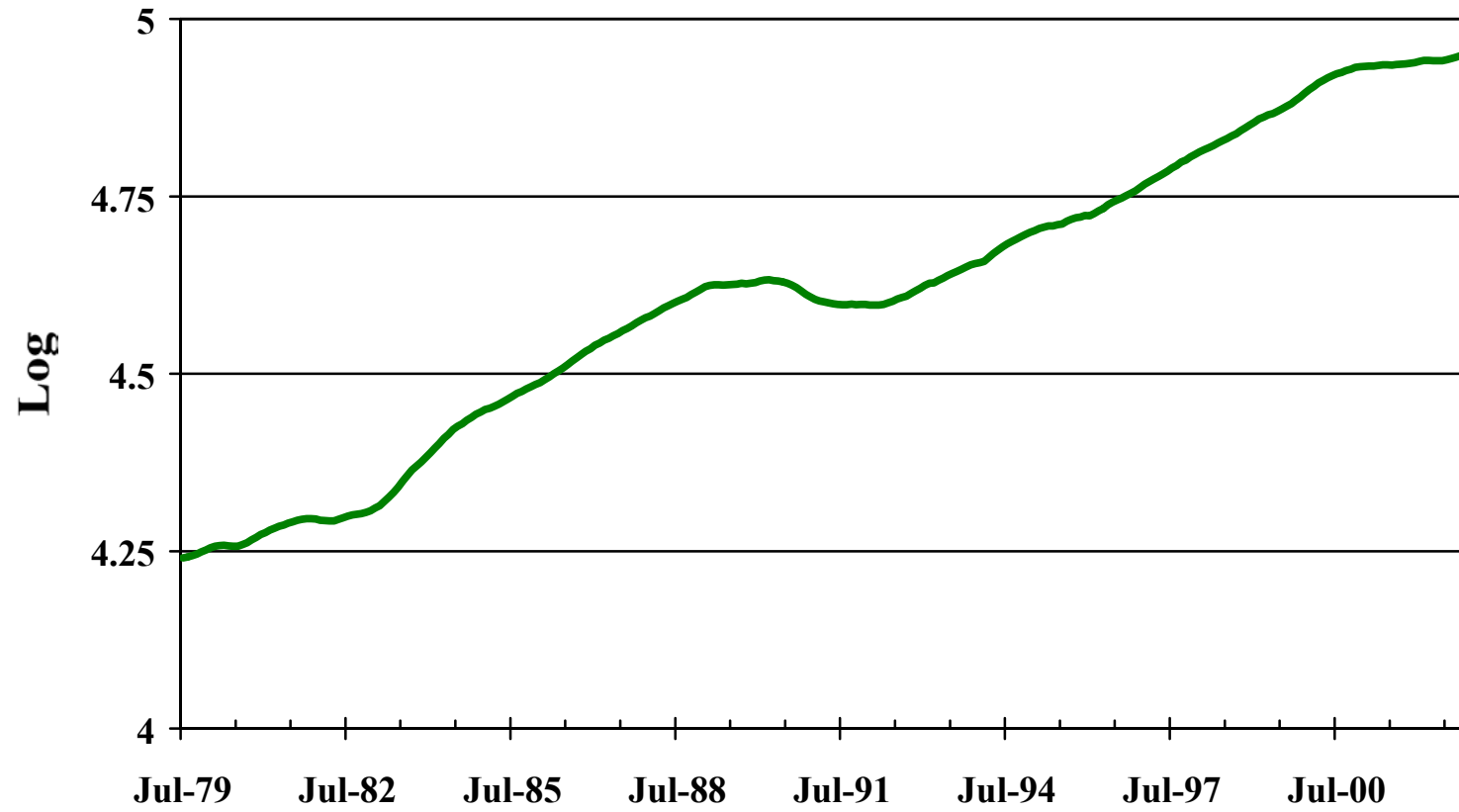
**Figure 2b**

**New Jersey Index: Trend Component**
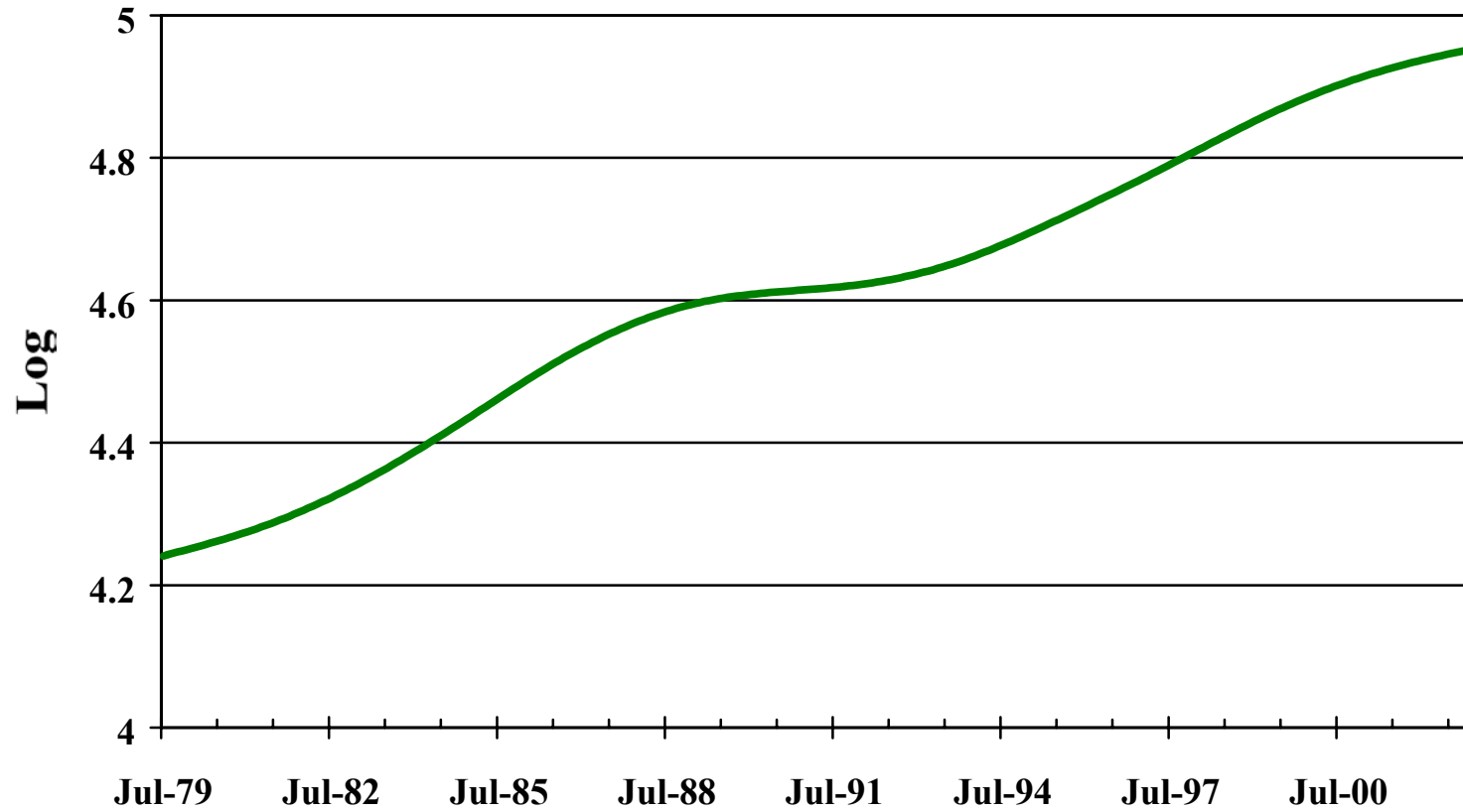
**Figure 2c**

**New Jersey Index: Cyclical Component**
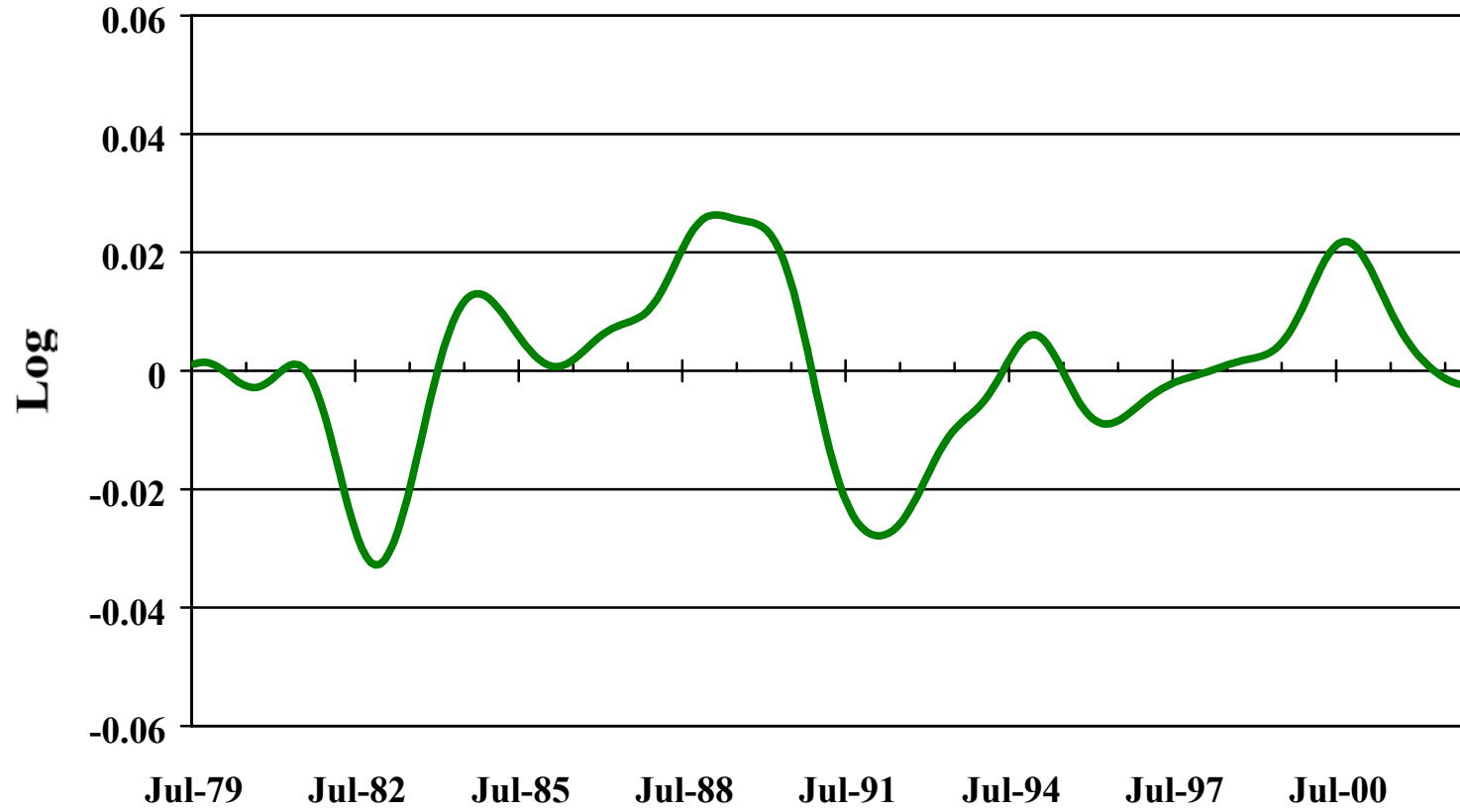
**Figure 3**

**Clusters of States Resulting from K-means Clustering (k=8)
by Business Cycle Components of State Coincident Indexes**
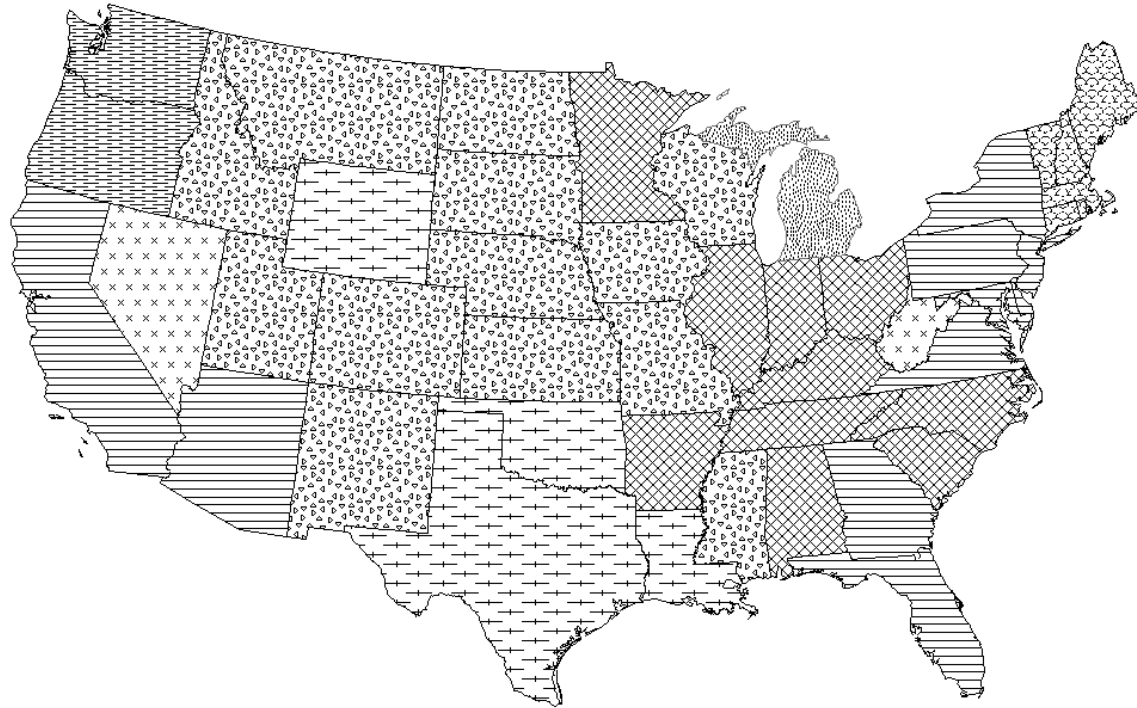
**Figure 4**
**Matrix of Relative Frequencies of Neighboring States Being in the Same Cluster**
**(10,000 Iterations, k=7)**[*]

| State | AL | AR | AZ | CA | CO | CT | DE | FL | GA | IA | ID | IL | IN | KS | KY | LA | MA | MD | ME | MI | MN | MO | MS | MT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AL | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0.143 | 0.155 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.819 | 0 |
| AR | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.000 | 0 | 0 | 0 | 0 | 0 | 0.793 | 0.808 | 0 |
| AZ | 0 | 0 | 1.000 | 0.451 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CA | 0 | 0 | 0.451 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CO | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.463 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CT | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.738 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| DE | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.797 | 0 | 0 | 0 | 0 | 0 | 0 |
| FL | 0.143 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0.773 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| GA | 0.155 | 0 | 0 | 0 | 0 | 0 | 0 | 0.773 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| IA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0.059 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.371 | 0.889 | 0 | 0 |
| ID | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.723 |
| IL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.059 | 0 | 1.000 | 0.100 | 0 | 0.147 | 0 | 0 | 0 | 0 | 0 | 0 | 0.033 | 0 | 0 |
| IN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.100 | 1.000 | 0 | 0.442 | 0 | 0 | 0 | 0 | 0.421 | 0 | 0 | 0 | 0 |
| KS | 0 | 0 | 0 | 0 | 0.463 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.875 | 0 | 0 |
| KY | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.147 | 0.442 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0.713 | 0 | 0 |
| LA | 0 | 0.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0.000 | 0 |
| MA | 0 | 0 | 0 | 0 | 0 | 0.738 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MD | 0 | 0 | 0 | 0 | 0 | 0 | 0.797 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 |
| ME | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 |
| MI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.421 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 |
| MN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.371 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 |
| MO | 0 | 0.793 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.889 | 0 | 0.033 | 0 | 0.875 | 0.713 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 |
| MS | 0.819 | 0.808 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.000 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 |
| MT | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.723 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 |

[*]Frequencies for neighboring states are in bold. Non-neighboring states are given a value of zero.

**Figure 4 (cont.)**

| State | AL | AR | AZ | CA | CO | CT | DE | FL | GA | IA | ID | IL | IN | KS | KY | LA | MA | MD | ME | MI | MN | MO | MS | MT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.539 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ND | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.062 | 0 | 0 | 0.927 |
| NE | 0 | 0 | 0 | 0 | 0.505 | 0 | 0 | 0 | 0 | 0.939 | 0 | 0 | 0 | 0.900 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.874 | 0 | 0 |
| NH | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.859 | 0 | 0.900 | 0 | 0 | 0 | 0 | 0 |
| NJ | 0 | 0 | 0 | 0 | 0 | 0 | 0.836 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NM | 0 | 0 | 0.015 | 0 | 0.897 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NV | 0 | 0 | 0.142 | 0.235 | 0 | 0 | 0 | 0 | 0 | 0 | 0.010 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NY | 0 | 0 | 0 | 0 | 0 | 0.181 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.021 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| OH | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.627 | 0 | 0.160 | 0 | 0 | 0 | 0 | 0.671 | 0 | 0 | 0 | 0 |
| OK | 0 | 0.000 | 0 | 0 | 0.138 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.001 | 0 | 0 |
| OR | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 | 0 | 0 | 0 | 0.003 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| PA | 0 | 0 | 0 | 0 | 0 | 0 | 0.622 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.513 | 0 | 0 | 0 | 0 | 0 | 0 |
| RI | 0 | 0 | 0 | 0 | 0 | 0.874 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.857 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| SC | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.137 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| SD | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.897 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.353 | 0 | 0 | 0 | 0.611 |
| TN | 0.676 | 0.528 | 0 | 0 | 0 | 0 | 0 | 0 | 0.366 | 0 | 0 | 0 | 0 | 0 | 0.670 | 0 | 0 | 0 | 0 | 0 | 0 | 0.417 | 0.600 | 0 |
| TX | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.991 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| UT | 0 | 0 | 0.019 | 0 | 0.906 | 0 | 0 | 0 | 0 | 0 | 0.695 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| VA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.154 | 0 | 0 | 0.653 | 0 | 0 | 0 | 0 | 0 | 0 |
| VT | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.795 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| WA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.003 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| WI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.929 | 0 | 0.085 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.003 | 0.425 | 0 | 0 | 0 |
| WV | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.074 | 0 | 0 | 0.027 | 0 | 0 | 0 | 0 | 0 | 0 |
| WY | 0 | 0 | 0 | 0 | 0.045 | 0 | 0 | 0 | 0 | 0 | 0.004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.041 |

**Figure 4 (cont.)**

| State | NC | ND | NE | NH | NJ | NM | NV | NY | OH | OK | OR | PA | RI | SC | SD | TN | TX | UT | VA | VT | WA | WI | WV | WY |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.676 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| AR | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.000 | 0 | 0 | 0 | 0 | 0 | 0.528 | 0.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| AZ | 0 | 0 | 0 | 0 | 0 | 0.015 | 0.142 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.019 | 0 | 0 | 0 | 0 | 0 | 0 |
| CA | 0 | 0 | 0 | 0 | 0 | 0 | 0.235 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CO | 0 | 0 | 0.505 | 0 | 0 | 0.897 | 0 | 0 | 0 | 0.138 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.906 | 0 | 0 | 0 | 0 | 0 | 0.045 |
| CT | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.181 | 0 | 0 | 0 | 0 | 0.874 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| DE | 0 | 0 | 0 | 0 | 0.836 | 0 | 0 | 0 | 0 | 0 | 0 | 0.622 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| FL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| GA | 0.539 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.137 | 0 | 0.366 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| IA | 0 | 0 | 0.939 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.897 | 0 | 0 | 0 | 0 | 0 | 0 | 0.929 | 0 | 0 |
| ID | 0 | 0 | 0 | 0 | 0 | 0 | 0.010 | 0 | 0 | 0 | 0.003 | 0 | 0 | 0 | 0 | 0 | 0 | 0.695 | 0 | 0 | 0.003 | 0 | 0 | 0.004 |
| IL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.085 | 0 | 0 |
| IN | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.627 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| KS | 0 | 0 | 0.900 | 0 | 0 | 0 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| KY | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.160 | 0 | 0 | 0 | 0 | 0 | 0 | 0.670 | 0 | 0 | 0.154 | 0 | 0 | 0 | 0.074 | 0 |
| LA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.991 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MA | 0 | 0 | 0 | 0.859 | 0 | 0 | 0 | 0.021 | 0 | 0 | 0 | 0 | 0.857 | 0 | 0 | 0 | 0 | 0 | 0 | 0.795 | 0 | 0 | 0 | 0 |
| MD | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.513 | 0 | 0 | 0 | 0 | 0 | 0 | 0.653 | 0 | 0 | 0 | 0.027 | 0 |
| ME | 0 | 0 | 0 | 0.900 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.671 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.003 | 0 | 0 |
| MN | 0 | 0.062 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.353 | 0 | 0 | 0 | 0 | 0 | 0 | 0.425 | 0 | 0 |
| MO | 0 | 0 | 0.874 | 0 | 0 | 0 | 0 | 0 | 0 | 0.001 | 0 | 0 | 0 | 0 | 0 | 0.417 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MS | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.600 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MT | 0 | 0.927 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.611 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.041 |

Figure 4 (cont.)

| State | NC | ND | NE | NH | NJ | NM | NV | NY | OH | OK | OR | PA | RI | SC | SD | TN | TX | UT | VA | VT | WA | WI | WV | WY |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.464 | 0 | 0.548 | 0 | 0 | 0.411 | 0 | 0 | 0 | 0 | 0 |
| ND | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.549 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NE | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.918 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NH | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.932 | 0 | 0 | 0 | 0 |
| NJ | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0.973 | 0 | 0 | 0 | 0.613 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| NM | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0.087 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.125 | 0 | 0 | 0 | 0 | 0 | 0 |
| NV | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0.233 | 0 | 0 | 0 | 0 | 0 | 0 | 0.004 | 0 | 0 | 0 | 0 | 0 | 0 |
| NY | 0 | 0 | 0 | 0 | 0.973 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0.630 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.126 | 0 | 0 | 0 | 0 |
| OH | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0.081 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.069 | 0 |
| OK | 0 | 0 | 0 | 0 | 0 | 0.087 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0.950 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| OR | 0 | 0 | 0 | 0 | 0 | 0 | 0.233 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.826 | 0 | 0 | 0 |
| PA | 0 | 0 | 0 | 0 | 0.613 | 0 | 0 | 0.630 | 0.081 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.281 | 0 |
| RI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| SC | 0.464 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| SD | 0 | 0.549 | 0.918 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TN | 0.548 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0.300 | 0 | 0 | 0 | 0 | 0 |
| TX | 0 | 0 | 0 | 0 | 0 | 0.125 | 0 | 0 | 0 | 0.950 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| UT | 0 | 0 | 0 | 0 | 0 | 0 | 0.004 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 | 0 | 0.044 |
| VA | 0.411 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.300 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0.055 | 0 |
| VT | 0 | 0 | 0 | 0.932 | 0 | 0 | 0 | 0.126 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0 |
| WA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.826 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 |
| WI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.000 | 0 | 0 |
| WV | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.069 | 0 | 0 | 0.281 | 0 | 0 | 0 | 0 | 0 | 0 | 0.055 | 0 | 0 | 0 | 1.000 | 0 |
| WY | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.044 | 0 | 0 | 0 | 0 | 0 | 1.000 |

**Figure 5**
**Regions Based on Cluster Analysis (k=7) of Business Cycle Components of State Coincident Indexes**
**And Weights for Neighboring States**