

Louisville's Peer Cities Revisited

An Urban Studies Institute Report

Introduction

On October 31, 2013, Urban Studies Institute (USI) staff presented a proposal to the Greater Louisville Project (GLP) Board to revisit the composition of the peer cities to which Louisville has benchmarked its socio-economic performance for almost 20 years.

The justification for the proposal was that the original work by Paul Coomes and Barry Kornstein was intended to identify Louisville's economic peers for the purpose of comparative economic performance. Louisville's willingness to benchmark its economic performance brought national attention to the practice and reflected well on the region. As benchmarking became accepted as common practice for broader purposes, those same set of peers were used to assess noneconomic outcomes like children's health, educational attainment, arts and cultural attractions, and immigration.

USI staff posed two questions to the GLP Board:

1. If USI repeated the original Coomes-Kornstein methodology using the variables that produced the original fourteen peer cities, but with the most recent data, would the composition of the cities change appreciably?
2. If USI added more social, demographic and labor market variables to the original variables, how would the composition of the peer cities change from the original model run with the most recent data?

Round One: Repeat the Original Methodology

Following the GLP Board's authorization to proceed, USI staff collected the most recent data on the 77 variables used by Coomes and Kornstein to produce the current peer cities list. The unit of analysis was metropolitan statistical areas (MSAs). Coomes noted in his summary that since the purpose of the peer list was economic performance benchmarking, the MSA was the appropriate unit of analysis. MSAs are simply economic regions comprised of interdependent counties, derived by commuting patterns and other economic behaviors that tie the geography together.

The same statistical processes were used as well. Factor analysis takes a large number of variables (77) and reduces them into a smaller number of variables that reflect certain dimensions, or factors, associated with economic performance. Once the factors have been identified, cluster analysis was performed to see which of the 381 MSAs had factors most similar to Louisville. (Please see the technical notes at the end of this report for more precise information on the statistical processes used to produce the results.)

Louisville's Peer Cities [Original]

Birmingham AL
 Charlotte, NC
 Cincinnati OH
 Columbus OH
 Dayton OH
 Greensboro NC
 Indianapolis IN
 Jacksonville FL
 Kansas City MO
 Memphis TN
 Nashville TN
 Omaha NE
 Raleigh NC
 Richmond VA

Louisville's Peer Cities [2013 Data]

Cincinnati OH
 Columbus OH
 Des Moines IA
 Eau Claire, WI
 Fargo ND
 Indianapolis IN
 Kansas City MO
 La Crosse WI
 Lincoln NE
 Madison WI
 Mankato MN
 Milwaukee WI
 Minneapolis MN
 Nashville TN
 Rapid City SD
 Sioux Falls SD
 St. Louis MO
 Topeka KS

The original group of peer cities was limited to 14, but USI staff chose to let the statistics suggest a cutoff that was near 14. Five of the original peer cities were also on the new data list; thirteen were new peers.

Round Two: Adding More Non-economic Variables

Since the purposes to which the peer cities had been used included non-economic comparisons, USI staff added 92 variables to the original 77 and performed the analysis again. Some of the 92 new variables described the local workforce, but most were demographic and social. The result was a more balanced group of variables for the factor analysis, reflecting both social and economic dimensions of the Louisville MSA.

Once again, no specific number of peers was determined prior to the statistical process. The “best fit” solution resulted in 28 peer cities. Eleven of the expanded data cities were original

peers; eight were also in the list of cities that resulted from the analysis produced by using 2013 data on the original Coomes-Kornstein factors.

Louisville's Peer Cities [Expanded Data]

1	Atlanta GA	15	Houston TX
2	Austin TX	16	Indianapolis IN
3	Birmingham AL	17	Kansas City MO
4	Cedar Rapids IA	18	Knoxville TN
5	Charlotte NC	19	Lexington KY
6	Cincinnati OH	20	Little Rock AR
7	Columbia SC	21	Memphis TN
8	Columbus OH	22	Minneapolis MN
9	Dallas-Fort Worth TX	23	Nashville TN
10	Denver CO	24	Oklahoma City OK
11	Des Moines IA	25	Omaha NE
12	Grand Rapids MI	26	Raleigh NC
13	Greensboro NC	27	St. Louis MO
14	Greenville SC	28	Tulsa OK

Focus Group: December 12, 2013

Just as Coomes used judgmental techniques to narrow his original list of possible peers, the GLP Board hosted key community leaders to review the two new sets of peer cities described above. The meeting was facilitated by Karen Wunderlin and attended by the following key stakeholders:

- Liz Alkire, Community Foundation of Louisville
- Susan Barry, Community Foundation of Louisville
- Kevin Connelly, Center for Nonprofit Excellence
- Maggie Elder, Metro United Way
- Michael Gritton, KentuckianaWorks
- Caroline Heine, Seed Capital Kentucky
- Jeff Mosely, Mayor's Office
- Susan Overton, Greater Louisville Inc.
- Stephen Reily, IMC Licensing
- Theresa Reno-Weber, Mayor's Office
- Ted Smith, Mayor's Office
- Carl Thomas, Gheens Foundation
- Dustin Wallen, Mayor's Office
- Mary Gwen Wheeler, 55,000 Degrees

➤ Mimi Zinniel, Olmsted Parks Conservancy

The group discussed different objectives and strategies for selecting peer cities, all of which were recorded for future use. One consensus conclusion emerged: the peer cities should not reflect cities that are at either end of the population size spectrum, either very large or very small.

USI staff decided to repeat the cluster analysis after removing the smallest places (MSA population under 250,000) and largest places (population 5,000,000 or greater). Atlanta, Dallas-Fort Worth and Houston fell from the list of 28, leaving 25 peer cities.

Core County Analysis

After reviewing the results, USI staff recommended that the unit of analysis change from the MSA to the core county so that more health and quality of place variables could be added to the analysis. The recommendation was driven by two considerations. First, the peers were used at the MSA level for analysis of certain economic variables. For example, gross domestic product, a crucial measure of productivity, is only available at the MSA level. Most other economic measures were actually benchmarked at the county level, even though MSA level analysis determined the list of peers.

Second, the use of peers had expanded from economic analysis to demographic and social analysis. Once again, the most common unit of analysis was the county because most of the demographic and social variables used in benchmarking were available at the county level. Educational attainment, government spending, poverty measures, cost of living, health indicators, air quality, arts and culture indices, and many more were typically county-level data applied to peer cities that had been generated by a statistical process that used MSA level data.

Authorized to proceed, USI staff repeated the MSA level cluster analysis and expanded the size of the clusters around Louisville. That produced a list of 74 MSAs for which the core county was identified. Staff then collected data on 193 variables for the 74 core counties.

Factor analysis was repeated for the new variables and the results were used to create a new cluster of 12 peers. The number of peers was unforced; that is, the peers presented below are the product of a statistical process that clustered in a group of 13 with Louisville.

The following table presents the 13 peers and notes if they also were presented in the original peer group, the repetition of the Coomes model using more recent data and 2013 geography, and the expanded MSA model with the addition of new variables.

MSA Peers	Core County of MSA	Original List	Repeat List	Expanded List
Birmingham AL	Jefferson County	X		X
Columbus OH	Franklin County	X		X
Fort Wayne IN	Allen County			
Grand Rapids MI	Kent County			X
Greenville SC	Greenville County			X
Indianapolis IN	Marion County	X	X	X
Kansas City MO	Jackson County	X	X	X
Knoxville TN	Knox County			X
Memphis TN	Shelby County	X		X
Oklahoma City OK	Oklahoma County			X
Omaha NE	Douglas County	X		X
Tulsa OK	Tulsa County			X
Wichita KS	Sedgwick County			

After discussing the results, USI staff decided to further refine the county level model, as certain statistical problems can arise when there are a greater number of variables (193) than observations (74). Similar variables were collapsed into a single measure to reduce the number of variables. For example, the variables measuring the percent of the adult population who have earned a bachelor’s degree and the percent of the adult population who have earned a graduate degree were combined to form a single variable, capturing the percent of the adult population who have earned a bachelor’s degree or higher.

To further reduce the number of variables, staff used the anti-image of the correlation matrix to eliminate variables not adequately represented by the model. Variables scoring below 0.5 were removed from the model. Factor analysis was performed again, this time using 56 variables. The cluster analysis followed, producing the group of peer cities shown on the next page.

The statistical test of sampling adequacy was slightly improved by the reduction of variables as expected, though both models were above the threshold for significance.

At the direction of GLP, USI staff met to review the results of all four sets of peer cities and make a recommendation to the GLP board for a new set of peers. Responding to the preference of the focus group, the ideal number of peers was determined to be between 14 and 20. Additionally, the focus group recommended that size comparability should be considered in the selection of peer cities. USI staff ranked all cities that appeared in the four “rounds” by population of the MSA, population of the core county, and percentage of the population in the MSA residing in the core county.

MSA Peer	Original List	Re-Run MSA List	Expanded MSA List	First County List
Birmingham AL	X		X	X
Charlotte NC	X		X	
Chattanooga TN				
Cincinnati OH	X	X	X	
Cleveland OH				
Columbus OH	X	X	X	X
Evansville IN				
Grand Rapids MI			X	X
Greensboro NC	X		X	
Indianapolis IN	X	X	X	X
Kalamazoo MI				
Kansas City MO	X	X	X	X
Memphis TN	X		X	X
Peoria IL				
Pittsburgh PA				
Richmond VA	X			
St. Louis MO		X	X	
Wichita KS				X
Winston-Salem NC				

The criteria employed for selection of the recommended peers was prioritized this way. First, peers that had emerged in both the MSA level and core county level models were selected. The justification was that the factor analysis at both levels had clustered them around Louisville, so we could be relatively certain they were similar in many important respects.

Then the peers that appeared once in the MSA level analysis and once in the core county analysis were examined for size comparability and all were ultimately included in the peer list. That produced a list of 15 peers.

Staff then reviewed the cities remaining on the list. Quite a few appeared on one list but no others. Each of the remaining cities was evaluated for size and comparability. Eventually, Nashville and Pittsburgh were selected as additional peers. Nashville appeared twice on the MSA list, but did not appear on the core county list. Pittsburgh appeared only once, on the second core county list. Both Nashville and Pittsburgh have certain economic, demographic and structural characteristics that are comparable to Louisville and are close to Louisville geographically.

Pittsburgh joins Kansas City, St. Louis and Cincinnati to provide some higher population balance to the peer list that now includes smaller places like Greenville, Greensboro and Knoxville.

	MSA Population	Core County Population	Percent of MSA in Core County
Birmingham AL	1,136,650	660,009	58%
Charlotte NC	1,831,084	969,205	53%
Cincinnati OH	2,146,560	802,038	37%
Columbus OH	1,878,714	1,195,537	64%
Grand Rapids MI	785,352	614,462	78%
Greensboro NC	736,065	500,879	68%
Greenville SC	653,498	467,605	72%
Indianapolis IN	1,798,786	918,977	51%
Kansas City MO	2,064,296	677,377	33%
Knoxville TN	709,492	441,311	62%
Louisville KY	1,302,457	750,826	58%
Memphis TN	1,377,850	940,764	68%
Nashville TN	1,645,638	648,295	39%
Oklahoma City OK	1,296,565	741,781	57%
Omaha NE	886,348	531,265	60%
Pittsburgh PA	2,360,733	1,229,338	52%
St. Louis MO	2,819,381	1,318,610	47%
Tulsa OK	951,514	613,816	65%

GLP Board members were presented with USI's recommended list of peers at their January 30 meeting. After discussion, the Board voted to remove Pittsburgh from the list and retain the other sixteen cities.

USI appreciates the Board's consideration of the new peer list. GLP's advocacy of benchmarking to peers has served the region well in the past. USI is confident that benchmarking to the new peers will provide new insights to the area's opinion leaders.

Technical Notes

MSA Geography

Following the 2010 census the MSA geography and designations changed. Some MSAs that existed in 2000 no longer exist. Quite a few new MSAs were added in 2013 following the decennial census. Because we wanted to use the most current data (2012) we elected to use the 2013 MSA designations. That created missing values when data was collected from sources that have not yet converted to the new MSA definitions. The statistical treatment of missing values can be complicated. We chose a technique called means substitution to “create” observations for new MSAs. The alternative would have allowed only 88 MSAs in to the analysis. Means substitution permitted all 381 MSAs to enter the factor analysis.

Factor Analysis

Exploratory factor analysis is a common data reduction technique. It reduces a large number of variables into a smaller number of factors (groups of variables) that the researcher can use for further analysis. The idea is that joint variations in the observed variables are attributable to a smaller group of unobserved latent variables. Interdependencies among the observed variables can be used to reduce the set of variables in a dataset.

All observed variables on 381 Metropolitan Statistical Areas (MSAs) were introduced in the first rotation, or initial run. For the repeat of the original Coomes-Kornstein model, 77 variables were introduced. For the expanded MSA level model, 169 variables were introduced. For the core county model, 193 variables were introduced.

Each rotation produces a set of factors with unique factor loadings (correlation coefficients between the variables and the factors) and eigenvalues (the variance in all the variables accounted for by that factor).

We produced a scree plot of the positive eigenvalues after each rotation and looked for the “elbow” or the point at which the curve begins to level off to get a sense of the number of meaningful factors. Only those factors whose eigenvalues were greater than 1 were retained.

Two statistical tests confirmed the adequacy of the using factor analysis to represent the data. The Kaiser-Meyer-Olkin measure of sampling adequacy sets a threshold of 0.6. All models except the first core county model produced a KMO score between 0.70 - 0.86. The number of variables was reduced in the second core county model, as described in the text, and the resulting model scored above the appropriate threshold.

Bartlett’s Test of Sphericity tests the null hypothesis that the correlation matrix is an identity matrix (all diagonal elements are 1 and all off diagonal elements are 0) and thus no linear

relationship exists among the variables. The null hypothesis was rejected at the 99% confidence level for each model.

Cluster Analysis

Like factor analysis, cluster analysis is a data reduction tool that creates subgroups by examining interrelationships among variables. However, where factor analysis reduces the number of variables by grouping them into a smaller set of factors, cluster analysis reduces the number of cases by grouping them into smaller sets of clusters.

We performed a hierarchical cluster analysis for both MSA core county models. The MSA clustering used Ward's method and the core county used furthest neighbor. Both Ward's and furthest neighbor are considered space-dilating methods.

In the first run, each MSA or core county is assumed to be in a single cluster with itself. The squared Euclidian distance was calculated for each cluster. The two most similar clusters were fused and distances were recalculated. The fusion process was repeated multiple times until of the clusters fuse together to form one cluster containing all MSAs or core counties. The results show some statistically optimal number of clusters. While the researcher can force the cluster analysis to a certain number, USI staff elected not to, letting the number of peers in the Louisville cluster fluctuate based on the statistical results. This explains why the three processes resulted in a different number of peers each time.

USI maintains the list of variables and observations on those variables used for all four analyses. They are available on request. USI has also retains the statistical output from each of the four factor analyses and four cluster analyses. They are also available for inspection by any interested party.