



11-1-1998

What is a Central City in the United States? Applying a Statistical Technique for Developing Taxonomies

Edward W. Hill

Cleveland State University, e.hill@csuohio.edu

John F. Brennan

Cleveland State University

Harold L. Wolman

University of Maryland - Baltimore County

Follow this and additional works at: https://engagedscholarship.csuohio.edu/urban_facpub



Part of the [Economics Commons](#), [Other Public Affairs](#), [Public Policy and Public Administration Commons](#), and the [Urban Studies and Planning Commons](#)

How does access to this work benefit you? Let us know!

Publisher's Statement

(c) 1998 Sage Publications

Original Citation

Hill, E. W., Brennan, J. F., & Wolman, H. L. (1998). What is a Central City in the United States? Applying a Statistical Technique for Developing Taxonomies. *Urban Studies*, 35(11), 1935-1969. doi:10.1080/0042098983962

Repository Citation

Hill, Edward W.; Brennan, John F.; and Wolman, Harold L., "What is a Central City in the United States? Applying a Statistical Technique for Developing Taxonomies" (1998). *Urban Publications*. 0 1 2 3 14.
https://engagedscholarship.csuohio.edu/urban_facpub/14

This Article is brought to you for free and open access by the School of Urban Affairs at EngagedScholarship@CSU. It has been accepted for inclusion in Urban Publications by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.

**WHAT IS A CENTRAL CITY IN THE UNITED STATES?
APPLYING A STATISTICAL TECHNIQUE
FOR DEVELOPING TAXONOMIES**

Edward W. Hill, *Cleveland State University*
John F. Brennan, *Cleveland State University*
Harold L. Wolman, *University of Maryland*

This article was originally published in:

Hill, Edward W., et al. (1998). What is a Central City in the United States? Applying a Statistical Technique for Developing Taxonomies. *Urban Studies*, 35(11), 1935-1969.

Post-print standardized by MSL Academic Endeavors, the imprint of the Michael Schwartz Library at Cleveland State University, 2012



What is a Central City in the United States? Applying a Statistical Technique for Developing Taxonomies

Edward W. Hill, John F. Brennan and Harold L. Wolman

Summary. We test the null hypothesis that municipalities defined as central cities by the US Bureau of the Census in 1990 are homogeneous—a hypothesis we reject. Rather, we find that US central cities consist of 2 distinct subsets of municipalities that are aggregated from 13 cluster groupings. The article has two purposes. The first is methodological. We develop a method that uses cluster analysis to group US central cities; then we employ discriminant analysis to establish the statistical validity of those groups. We also develop techniques to minimise the role of judgement in selecting the appropriate cluster solution. The second purpose of the article is to test the substantive null hypothesis. Our rejection of the homogeneity assumption raises the spectre of specification error in research and public policies that assume homogeneity among central cities.

1. What Does ‘Central City’ Mean?

The power of the term ‘central city’ lies in the image it connotes. Say the word, and an icon of urban America is immediately constructed: large municipalities that are disproportionately poor and distressed, both socially and economically.¹ When the term is used as an image, its use incorporates the functions of America’s core municipalities (what is done within central cities) their conditions (or the social and economic outcomes from those functions that are disproportionately concentrated in stereotypical central cities), and the physical structure of stereotypical American metropolitan areas (a core

central city dominated by poor residential neighbourhoods surrounded by wealthier suburbs). This image is based largely on older central cities, most often located in the north-east and midwest US. A central city is typified as being the primary municipality of an expansive metropolitan area, consisting of a dense and dominant central business district surrounded by enclaves of the poor that often overlap with minority residential neighbourhoods.

Although that image is powerful, it does not apply equally well to all central cities in the nation. Most observers will agree—and

Edward W. Hill and John F. Brennan are at the College of Urban Affairs, Cleveland State University, Cleveland, OH 44115, USA. Fax: (216) 687-9277. E-mail: Ned@wolf.csuohio.edu (Edward W. Hill); Bren4452@wolf.csuohio.edu (John F. Brennan). Harold L. Wolman is at the Policy Sciences Graduate Program, University of Maryland, Baltimore County, 627 Administration, 1000 Hill Top Circle, Baltimore, MD 21250, USA. Fax: (410) 455 1172. E-mail: HDWolman@Potomac.net. Research support was received from Ohio’s Urban University Program. We thank Coit Cook Ford III, Peter Fisher, Bennett Harrison, Helen Liggett and Janet Smith for comments. We also benefited from the comments of the reviewers. Programming assistance was provided by Ellen Cyran of Cleveland State University’s Urban Center.

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

much research attests to the fact—that not all US central cities are the same (Bradbury *et al.*, 1982; James, 1990; Ladd and Yinger, 1989; Rusk, 1993; and Wolman *et al.*, 1994). Yet we frequently act as if they are, both in our research and in policy formulations. The reason this heterogeneous set of places tends to be discussed as if it is a homogeneous collection is that ‘central city’ is not just an image or a stereotype; it is a statistical artifact—created by the Bureau of the Census to operationalise the concept of central city. (See the Appendix for a history of the central city as a statistical concept in the US.) And, whenever any stereotype is operationalised, there is slippage.

Since 1983, the Bureau of the Census has used multiple criteria to identify municipalities as central cities. This definition recognises both the role of central cities as important employment nodes—which is consistent with the labour market basis of defining the extent of Metropolitan Statistical Areas (MSAs)—and the geographical spread of metropolitan areas (US Bureau of the Census, 1991, p. 356). The conflict between the image and statistical definition occurs because the image incorporates function, conditions or outcomes, and spatial structure, whereas the operational definition captures the functions (in the limited sense of the central city as a place of employment) and spatial structure in terms of the size of the residential population.

There is no problem with the way the Bureau of the Census quantitatively identifies central cities—the definition is clear, consistent and precise—but the problem lies with the way the operationalised concept is used. The Bureau of the Census has succeeded in establishing an economically based definition of central cities. However, the definition is then used for a purpose for which it was not intended, as a shorthand expression of social outcomes or concerns. Our own experience illustrates this.

Our interest in the definition of central cities began with papers that two of us wrote about the income relationships between central cities and their suburbs in 1990, and

changes in that relationship between 1980 and 1990 (Hill and Wolman, 1997a, 1997b). We discovered that there was wide variation among places called central cities, in terms of their income relationships with their suburbs, and we had to adjust our selection criteria so that the central cities we used were appropriate for the hypotheses we were testing. One component of the stereotype of American central cities holds that their per capita income is lower than that of their suburbs. In working with the universe of 152 metropolitan areas with populations of at least 250 000 in 1980, we were surprised to find that, in 1990, central city per capita income exceeded suburban per capita income in 37 of these metropolitan areas—which is 24 per cent of the total. This is a rather large hint that the universe is not homogeneous.

We then inspected the list of places classified by the Bureau of the Census as central cities and saw a number that did not strike us as having the characteristics of a stereotypical central city. Although all the municipalities on the list appeared to be nodes of metropolitan area employment, Pasadena, California; Dearborn, Michigan, and Lynn, Massachusetts did not fit the image we had of central cities. Pasadena struck us as being a large, prosperous suburb of Los Angeles. Dearborn is a residential suburb of Detroit that contains the headquarters of the Ford Motor Company. Lynn is a decayed factory town that has been swallowed by the northward push of Boston’s suburbs. Left off the list of central cities, on the other hand, are extremely poor suburbs with large concentrations of social problems usually associated with central cities, but lacking large concentrations of employment, such as East Cleveland, Ohio, and Highland Park, Michigan (which is completely surrounded by the city of Detroit!)—even though Camden, New Jersey and East St Louis, Illinois, which are socially analogous to East Cleveland and Highland Park, are listed as central cities. In other words, our preconceived notion was built upon the stereotype of the social outcomes presumably contained in central cities, whereas these same cities were defined in

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

terms of their employment and residential functions.

The purpose of this article is to develop a rigorous, inductive methodology that—starting with the broader universe of heterogeneous central cities identified by the Census Bureau’s definition of a central city—permits us to identify various groups of more or less similar cities in a rigorous, inductive fashion according to their function, outcome and spatial structure, using cluster analysis, and then sets forth what distinguishes these various groups of homogeneous cities from each other, using discriminant analysis. We do not have an *a priori*, or prior, notion as to which central cities are distressed or not distressed, or similar and dissimilar; instead, we use the discriminant functions to characterise the differences among the clusters of central cities. We accept the Census Bureau’s definition as a starting point, since that is the construct around which data on cities are collected, and there is a solid theoretical rationale for the employment focus of that definition.² Although results of this procedure should provide a much better understanding of how, and along what dimensions, US central cities differ (and, consequently, should permit better and more sensitive research and policy making), our primary purpose in this article is to develop the methodology.

We discuss the methodology and variable selection in the next section. The third section is devoted to describing a technique that identifies the candidate cluster solutions. Cluster maps are provided, and the discriminant functions are discussed, in the fourth section. In the fifth section of the paper, we then interpret what differentiates the clusters of central cities.

2. Methodology

We classify US central cities into like groups in two stages. First, hierarchical cluster analysis is used to form groups from the universe of central cities, based on a number of relevant variables. We then employ discriminant analysis to assess the internal validity of the resulting clusters and, more

important, to identify the groups of variables that distinguish the clusters of cities.³ These two techniques use the same body of data but are conceptually different. Hierarchical cluster analysis is a mathematical rather than a statistical procedure. In cluster analysis, there is no dependent variable, and there are no meaningful descriptive or test statistics. Cases (in this research, our cases are central cities) are sorted into like groups. Discriminant analysis, on the other hand, is a statistical procedure that tests the goodness of fit of the prior group assignments. In this research, the prior groupings are the groups of cities formed by the cluster analysis. The groupings tested by the discriminant analysis form a multi-part, categorical, dependent variable, and there are meaningful test statistics.

The cluster analysis is discussed in the next part of this section. It is followed by a more complete consideration of the discriminant analysis in section 2.2. The selection of variables used in both procedures is discussed in section 2.3.

2.1 Cluster Analysis

We selected agglomerative hierarchical cluster analysis as the grouping procedure, because it is an effective tool for identifying distinct groupings within a population (Everitt, 1993, pp. 6-7). This mathematical technique is an operation that begins with the same number of clusters as there are observations (in our case, the 508 central cities in the US) and proceeds to group similar observations together in a systematic fashion, until the final cluster contains all the observations. Groups are constructed by minimising the variance of squared Euclidean distances for each variable between cities. These distance coefficients are derived from standardised variable scores (z scores).⁴ The number of stages in the process is one less than the number of observations. The key to the analysis is first to identify the candidate cluster solutions from all the cluster solutions. We discuss this in section 3.

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

2.2 Discriminant Analysis

Although cluster analysis identifies the existence of groups within populations, discriminant analysis tests the statistical validity of those groupings (Klecka, 1980; Hair *et al.*, 1987). Therefore, discriminant analysis requires *a priori* groups of observations, and the cluster analysis provides those groupings. More important, discriminant analysis also identifies the variables (and groups of variables) that drive the classification process. This allows us to discuss the typologies that the clusters represent from the data, rather than to look at place names and allow our prejudice or intuition to drive the cluster labelling.

Stepwise discriminant analysis was used to assess the relationships between variables and groups of variables by introducing variables into the analysis one at a time. Stepwise testing is legitimate in this case, because there are no *a priori* hypothesised relationships between the descriptor variables and a particular cluster of central cities. As we use the technique, discriminant analysis is equivalent to an analysis of variance that tests for statistically significant differences in variables across the clusters (Varady and Lipman, 1994, p. 531). Discriminant analysis yields a series of discriminant functions (one less than the number of groups) that are somewhat akin to factors in factor analysis. For each discriminant function, the mean discriminant score can then be calculated for each of the prior groups (the different groups classified by the previously conducted cluster analysis). The mean score for each cluster group is then used to derive a z score for each cluster—thus measuring the number of standard deviations that cluster is from the mean discriminant score for all central cities on that function. The clusters with z scores of 2 or more are judged to be highly associated with the discriminant function.

For readers who are unfamiliar with discriminant or factor analysis, but use regression techniques, it may be useful to think of discriminant analysis as being analogous

to a multinomial logit or probit equation, where linear combinations of two or more independent variables are used to describe the behaviour of a single, multiple-category, dependent variable (Hair *et al.*, 1987). The discriminant analogue to the dependent variable would be each prior grouping—which, in our case, is one of the cluster groupings provided by the cluster analysis—and the independent variables would be the sets of descriptor variables found in each of the discriminant functions. However, the composition of each discriminant function is not known *a priori*; as is also true of the factors in factor analysis. As discussed above, the strength of the statistical association of each discriminant function with each prior group is assessed by examining the z score of each cluster group, evaluated at the mean for that group, for a particular discriminant function. This is similar to using t -tests to determine the strength of the statistical relationship between independent variables and the dependent variable of a regression equation.

Discriminant analysis generates a suggested group assignment for each case, based upon an aggregate discriminant score, and indicates whether that generated assignment corresponds to the prior group assignment. The aggregate discriminant score for each central city is calculated by summing the weighted scores for each discriminant function, in which each score's weight is based upon the percentage of the overall explained variation accounted for by that discriminant function. The resultant weighted score is used to assign the city to its discriminant grouping; this is the grouping that is generated by statistical criteria, as opposed to the cluster grouping, which is the grouping that is suggested by mathematical criteria. This is analogous to generating a fitted regression equation and comparing the estimated, or fitted, values with the observed values. Just as the mean squared error of the fitted versus actual values is the basis for determining the overall goodness of fit in a forecasting regression equation, the overall fit of the discriminant analysis is measured by the percentage of cases in which the discriminant

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

group assignment corresponds to the prior group assignments (in our case, the percentage of the cases in which the discriminant group assignments correspond with the cluster group assignments). This percentage is called the 'hit ratio' and is the measure of overall goodness of fit of the cluster groupings.

Combining cluster and discriminant analysis provides a number of tools to assess the homogeneity assumption in the classification of central cities. First, the cluster analysis provides a mathematical assignment of central cities into cluster groups that are hierarchical; central cities that are assigned to groups at earlier stages of the procedure stay together in subsequent stages as larger, more heterogeneous groups are formed. That is why the procedure is considered hierarchical. Secondly, discriminant analysis statistically tests the internal validity of each group, and the 'hit ratio' in the discriminant output provides an indicator of the goodness of fit of the original cluster assignments. (We also use *t*-tests of the mean discriminant scores of the paired cluster and discriminant groupings to determine whether there are statistically meaningful differences between each group.) Finally, discriminant analysis identifies the set of variables most highly associated with the *z* scores of each discriminant function, thus indicating which sets of associated variables are statistically significant descriptors of each grouping.

2.3 Variable Selection

Variable selection is critical to our procedure, because the cluster analysis minimises the within-group variance based on all the variables included in the analysis and cannot distinguish between variables statistically. Therefore, variables that do not have theoretical reasons for inclusion will distort the results of the clustering process. Previous research on the classification of cities, regions and metropolitan areas has used economic, social and demographic variables as the basis for classification. In the introduction to the *City Classification Handbook*,

Berry and Smith (1972a, pp. 1–2) stated that there are a number of reasons to classify urban places:

To some classification is a means of data exploration, either to determine convenient ways of summarizing information, to find new and potentially useful hypotheses, or to produce a universally true typology. To others, classification provides a means of facilitating hypothesis-testing or model-fitting. Yet others are concerned with developing improved modes of prediction, using subgroups rather than an entire population as guides to an efficient sampling plan, elements for which predictions are made, or guides to the selection of analogs or other forms of comparative cases.

Berry's work had been constructed on a rich tradition in the US that dates to the seminal work of Chauncy Harris (1943), in which he primarily used industrial specialisation data to order US cities. Nelson (1955), Hart (1955), Jones and Forestall (1963) and Forestall (1967) followed with similar classification approaches, adding occupational data to the industrial data. By and large, the overriding purposes of these studies were to analyse the spatial organisation of US cities in the context of central place theory and to identify the hierarchy of urban places that is derived from central place theory. The results were a depiction of US cities, ordered spatially and by economic function.

The purpose of the research on urban classification changed in the early 1970s. The focus shifted from testing a theoretical framework—central place theory and the hierarchy of urban places—to analysing the spatial concentration of various social problems, in the Chicago School's ecological tradition: entering variables into the analyses that measure social outcomes. Although the results of these empirical investigations are important theoretically, they had clear public policy purposes. These studies were conducted as ways of identifying places where social problems were concentrated. All of these

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

schemes used economic, social and/or demographic variables as the basis for their classifications. Berry (1996) reviews most of these classification efforts; of those reviewed, only two—Berry and Smith (1972) and Noyelle and Stanback (1983)—presented functional classification schemes that had spatial policy relevance. Keeler and Rodgers (1973) also used several social and demographic variables to classify metropolitan areas.

The variables we used in the cluster and discriminant analyses were selected to meet one of four sets of criteria. They had to be either:

- (1) one of the Bureau of the Census' statistical criteria for identifying central cities; or
- (2) descriptors of the role of central cities in the labour market; or
- (3) descriptors of central cities as locations of social outcomes of particular policy relevance; or
- (4) descriptors of the spatial structure of metropolitan areas.

An argument could be made that each set of variables should be tested separately and in that way separate sets of central cities generated—one that identifies types of central cities by function, the second by social outcome, and the third by spatial structure. Although doing this form of 'marginal' analysis is interesting in its own right, such an approach would not meet our purpose. We are interested in building a typology of central cities that combines the statistical definition of the concept (the economic and residential function) with the popular image of central cities (largely defined by social outcome), controlling for spatial structure. This means that we need to include all four sets of variables in our analysis. The full list of variables used and the form these variables take in the analyses are given in Table 1.

The first group of variables in Table 1 captures the economic function of central cities; there are three sets of variables within this group. The first set contains the three

variables that are derived from the Census qualifying variables. We expected these variables not to be strong discriminators among the universe of central cities—because, by definition, variation among these places is limited. The second set of variables measures the industrial composition of employment among residents of the central cities and acts as proxy variables for the demand side of the local labour market. It is expected that central cities with larger shares of their residents employed in manufacturing will cluster together and be somewhat more disadvantaged than will be central cities with strong bases in the other industrial groups.

The third set of variables in the first group records the male and female labour force participation rates, the occupational distribution of residents of the central cities, and the distribution of the terminal educational attainment of adults. Together, these variables approximate the supply side of the local labour market. It is expected that this set of variables will provide a wider and better array of discriminating variables than will the other two sets, because it better reflects the purpose of residential neighbourhoods in a regional economy; they are pools of labour (Hill and Bier, 1989; Teitz, 1989). It is expected that healthier central cities will be associated with more of their residents being employed in symbolic analytical occupations (managerial, professional, technical and sales) and general service occupations, whereas distressed central cities will have a larger share of residents employed in the other occupational groupings. A similar set of expectations holds for the educational attainment of the adult population: the larger the proportion of highly educated residents, the less economically distressed the central city.

The second major group of variables contains the social outcomes that are disproportionately concentrated in central cities, are of interest to public policy, and form a large part of the popular and political image of cities. The first two variables concern population changes in the central city and in the metropolitan area from 1980 to 1990. We

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

Table 1. Variables used in the analysis

Economic function

Bureau of the Census' central city qualifying variables

The ratio of total employed in central city to employed residents of central city

The logarithm of the central city population

The percentage of employed central city residents that is employed outside the central city

Demand side of the labour market

Percentage of manufacturing industry jobs

Percentage of wholesale or retail industry jobs

Percentage of service industry jobs

Percentage of health, education and government industry jobs

Supply side of the labour market

Female labour force participation rate for the central city

Male labour force participation rate for the central city

Percentage of symbolic analyst occupations: managerial and professional, technical and sales occupations

Percentage of general service occupations: administrative support and service occupations

Percentage of machine and precision producer occupations

Percent labourer occupations

Percentage of population over 25 where less than 9 years of schooling is highest educational attainment

Percentage of population over 25 where some high school is highest educational attainment

Percentage of population over 25 where high school diploma is highest educational attainment

Percentage of population over 25 where some college is highest educational attainment

Percentage of population over 25 where associate's degree is highest educational attainment

Percentage of population over 25 where bachelor's degree is highest educational attainment

Percentage of population over 25 where graduate degree is highest educational attainment

Social outcome

The percentage change in population in the central city from 1980 to 1990

The percentage change in population for MSA from 1980 to 1990

The central city to MSA median household income ratio

The logarithm of central city per capita income

The poverty rate for the central city

The unemployment rate for the central city

Median monthly central city renter costs

Spatial structure

Percentage of the population that is African-American

The logarithm of the central city population density

The logarithm of the MSA population

The percentage of the MSA population that resides in the central city

Average travel time of residents employed in the central city

Median age of the single-family housing stock for the central city

Notes: Data obtained from 1980 and 1990 STF Data Files from the Bureau of the Census or from unpublished data provided by the Bureau of the Census. All data are for the year 1990 unless otherwise noted.

expect that clusters of healthy metropolitan areas will be typified by population growth and that those central cities with growing populations will be located in fast-growing metropolitan areas. However, there are some growing metropolitan areas that contain declining central cities, and these should form separate clusters.

The next three outcomes are the major focus of public concern: income. The ratio of central city to metropolitan average incomes should be smaller in less distressed central-cities and wider in more distressed central cities. Additionally, better-off central cities should be typified by larger absolute average incomes, as measured by the logarithm of per

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

capita income. We also expect that more distressed central cities are characterised by higher poverty and unemployment rates. We included the median monthly rental housing costs in the central city, assuming that more distressed central cities should have lower rental costs. Yet those central cities that are less distressed but have higher rental housing costs present policy problems different from those that are more distressed with lower rental costs.

The last group of variables we entered into the models measure the spatial structure of the central cities. We expect that the more distressed central cities will have larger shares of African-American residents in their populations. We expect this result because the African-American population is, on average, poorer than the majority population; having a larger share of African-American residents within a city's population implies, *ceteris paribus*, that the city's population will be poorer. We expect that more dense central cities will be worse off than less dense central cities, because density is a proxy variable for the economic 'age' of the city. And we expect that, the older the central city, the less attractive it is for modern employment and living. We also included the median age of the single-family housing stock as a way of capturing another dimension of the age of the central city. We entered the size of the metropolitan area as a variable, because it is an important spatial descriptor, and we expect that large metropolitan areas will have different characteristics from those of smaller places. We expect that, the smaller the fraction of the metropolitan area's population that resides in the central city, the worse-off that city will be, assuming that this is associated with either long-term population flight or the 'inelasticity' of the central city, to use Rusk's (1993) phrase. We also assume that, the longer their average travel times to their workplaces, the worse off the central city.

3. Selecting the Candidate Cluster Solutions

There is no purely objective method to determine the optimal or 'correct' cluster solution.

The critical question is when to stop clustering (Aldenderfer and Blashfield, 1984; Everitt, 1993). Everitt (1993, p. 44) indicates that analysis of the agglomeration schedule, specifically the change in the agglomeration coefficient, is the most commonly employed guide to halting the clustering. (The agglomeration coefficient is the sum of the within-group variance of the two clusters combined at each successive stage.) Simply put, a 'marked' increase in the value of the agglomeration coefficient between two stages indicates that heterogeneous clusters are being combined. The result is a greater increase in total variance. At this point, an assessment should be made as to whether the optimal cluster solution has been reached.

Table 2 contains the partial agglomeration schedule for the cluster analysis performed on the central cities (the last 33 out of all 507 stages are reported). The first column of the table lists the stage of the cluster solution. The second column gives the number of clusters in that solution. The agglomeration coefficient is listed in the third column of the table. We use the data contained in the fourth and fifth columns of Table 2 to help select the candidate cluster solutions. The fourth column is the percentage change in the value of the agglomeration coefficient from the previous stage. The fifth column is the percentage change, of the percent change, found in column four. To understand better the meaning of these two numbers, we retreat to calculus to find an analogy.

The percentage change in the fourth column is the rate of change in the agglomeration coefficient from one stage to the next; in other words, it is the slope of the agglomeration schedule. The percentage in the fifth column is the rate of change in that slope coefficient, making it the measure of acceleration in changes of the agglomeration schedule. Keeping with the calculus analogy, we label the fourth column the first derivative of the agglomeration schedule and label the fifth column the second derivative of the agglomeration schedule.

The decision rule for selecting the candidate cluster solutions is: when there is a

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

Table 2. Partial agglomeration schedule for cluster analysis

Stage	Clusters in the solution	Agglomeration coefficient	First derivative (%) ^a	Second derivative (%) ^b
475	33	6 620		
476	32	6 705	1.3	
477	31	6 792	1.3	0.0
478	30	6 880	1.3	0.5
479	29	6 971	1.3	2.3
480	28	7 064	1.3	0.6
481	27	7 162	1.4	4.1
482	26	7 260	1.4	- 1.3
483	25	7 363	1.4	3.5
484	24	7 469	1.4	2.2
485	23	7 590	1.6	12.2
486	22	7 718	1.7	3.8
487	21	7 848	1.7	- 0.4
488	20	7 990	1.8	8.0
489	19	8 142	1.9	5.1
490	18	8 300	1.9	2.0
491	17	8 463	2.0	0.8
492	16	8 639	2.1	6.2
493	15	8 815	2.0	- 1.8
494	14	9 008	2.2	6.7
495	13	9 201	2.1	- 1.7
496	12	9 484	3.1	43.6
497	11	9 773	3.1	- 0.8
498	10	10 092	3.3	6.8
499	9	10 439	3.4	5.6
500	8	10 826	3.7	7.6
501	7	11 296	4.3	17.3
502	6	11 772	4.2	- 2.9
503	5	12 307	4.5	7.8
504	4	13 089	6.4	39.8
505	3	14 038	7.2	14.1
506	2	15 042	7.2	- 1.4
507	1	17 238	14.6	104.2

^a The percentage change in the agglomeration coefficient from the previous stage, given in column 3.

^b The percentage change of the percentage change in the agglomeration coefficient, given in column 4.

‘marked’ increase in the agglomeration coefficient, the previous stage of the cluster solution is a candidate solution. The challenge is to determine what constitutes a ‘marked’ increase. We identify the stages in which there are large changes in the first and second derivatives as candidate cluster solutions. Based on the first and second derivatives of the agglomeration schedule, there are three candidate solutions, at 2 clusters, 5 clusters and 13 clusters. These are indicated by large increases in the agglomeration

coefficients at the first, fourth and twelfth stages. The clusters where there are ‘marked’ increases in the first and second derivatives are indicated by bold type in the table, as are the data from the previous stage of the agglomeration schedule. We then use a combination of the first and second derivatives of the agglomeration schedule, output from the discriminant analysis, and face validity to choose among these candidate solutions.

First we examine the first and second derivatives in Table 2. The largest deriva-

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

tives were produced by the following moves (in descending order of the derivative, listing the first derivative and then the second): 2 clusters to 1 (14.6 per cent and 104.2 per cent), 5 to 4 (6.4 per cent and 39.8 per cent), and 13 to 12 (3.1 per cent and 43.6 per cent). Therefore, the largest heterogeneous groupings occur in movements from 2 to 1 cluster and then from 13 to 12 clusters. The second indicator of the validity of the cluster solutions comes from the 'hit ratio' of the discriminant analysis. The hit ratio is the percentage of the central cities for which the cluster and discriminant group assignments agree. Because the clustering is hierarchical, it is reasonable to expect that the hit ratio will increase as the number of clusters is reduced. These results are given at the bottom of Figure 1.⁵ The hit ratios are all in excess of 85 per cent. The 2-cluster solution has a hit ratio of 90 per cent, the 5-cluster solution produces a hit ratio of 90 per cent, and the 13-cluster solution has a hit ratio of 85 per cent.

As we mentioned above, deciding on the number of clusters to examine is, at root, a qualitative decision. There is a trade-off between changes in the agglomeration schedule (the first and second derivatives), the percentage in the variation in the clusters explained by the data (the hit ratio) and the number of clusters. However, the most important guide is that the resulting cluster solution has face validity. A good example of this line of reasoning is found in Gittleman and Howell (1995, p. 424). They examined clusters of jobs to test the dual labour market hypothesis. Because there is no quantitative rule for determining where the cluster grouping should stop, they decided that more compelling than any mechanical rule is the qualitative determination that at "various stopping points, the cluster analysis produces groups that are meaningful, particularly in light of previous theoretical and empirical work". Our decision rules lead us to prefer the 2- and 13-cluster solutions. However, we present results from all three candidate solutions to help distinguish among the resulting groupings of central cities.

4. Interpreting the Results

The null hypothesis is that the set of cities that the Bureau of the Census labels 'central cities' is homogeneous. Given the methods we use, this means that there would be one cluster of central cities, and the cluster would be confirmed by having low first and second derivatives of the agglomeration schedule in the move from 2 clusters to 1. The results clearly reject the null hypothesis. The universe of places called central cities is not homogeneous. At a minimum, there are two distinct groups of central cities: one healthy, the other distressed.

4.1 The Cluster Map

We use the candidate cluster solutions—the 2-, 5- and 13-cluster solutions—to map the relationships that exist among the various groups of central cities in Figure 1. The clusters are hierarchical, meaning that those appearing in earlier stages of the analysis remain together in later clusters. As the clustering progresses, each cluster becomes more heterogeneous, as indicated by the increases in the agglomeration schedule. Each of the candidate solutions we selected marks a stage in which the clusters at the next stage are much more homogeneous than they are at present stage. We use all three cluster solutions in the discussion of the statistical results that follows. To distinguish among these solutions and to recognize their hierarchical nature, we refer to the clusters in the 2-cluster solutions as two *groups* of central cities; we term each of the clusters in the 5-cluster solution as a *set* of central cities; and we call each of the 13 clusters just that—a *cluster*. When we link the hierarchical solutions in a cluster map (Figure 1), we are able to trace the contours of America's central cities. (The names given to the clusters help to provide an intuitive understanding of the clusters, and are explained later in the paper.)

We label one group of central cities 'stressed', the other we call 'healthy'. This

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

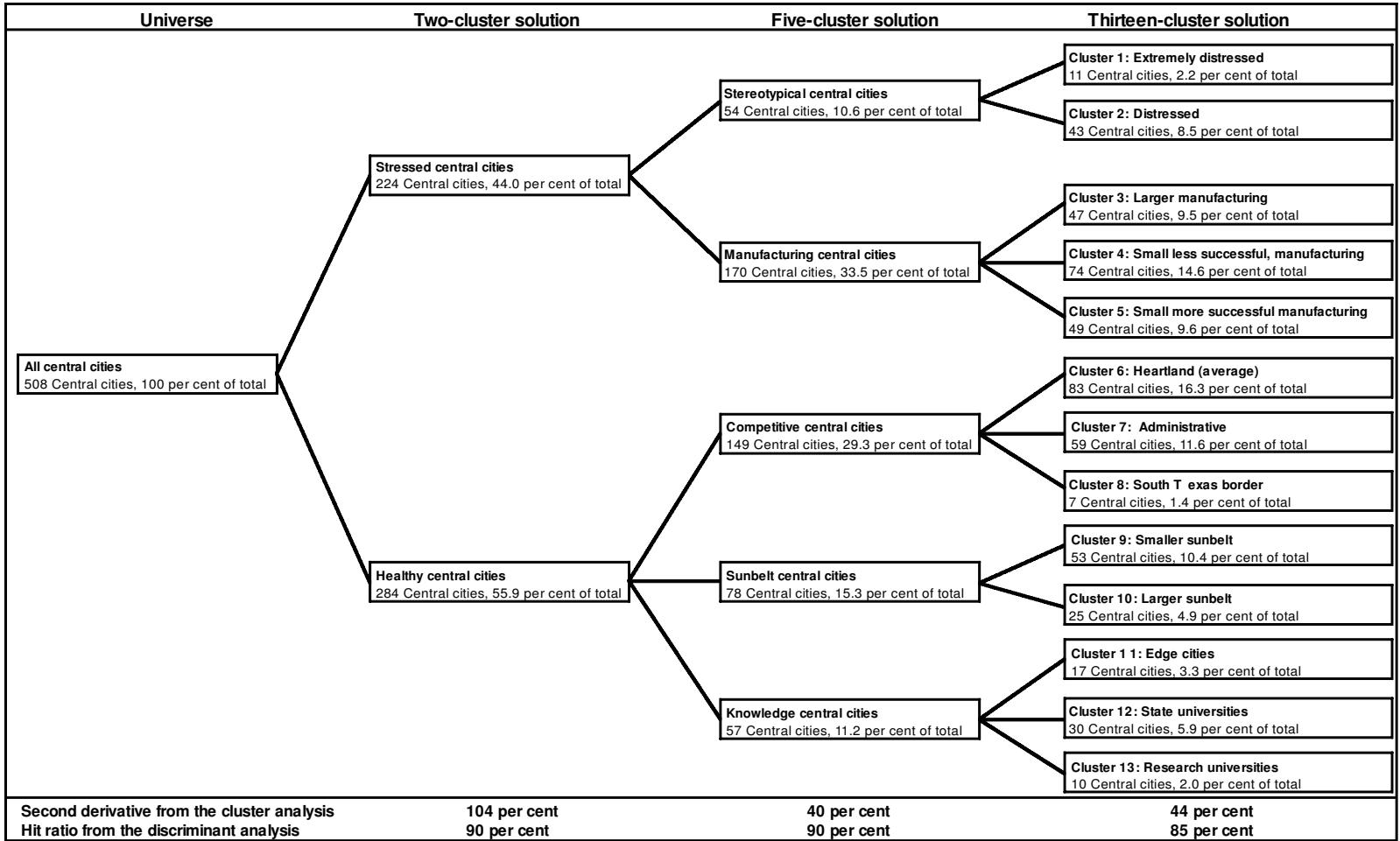


Figure 1. Cluster map.

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

distinction is drawn from the discriminant functions that describe the clusters that comprise each group. Those central cities that are stressed are typified by lower skilled and more manufacturing-based economic functions, they tend to house smaller proportions of the metropolitan area's population and—most importantly—they are typified by poorer social outcomes. The stressed communities have higher proportions of their populations with low levels of educational attainment and low incomes. Additionally, the stressed cities tend to play a smaller role in the economies of their metropolitan areas than do the healthier cities. Statistically, the clusters of stressed communities have higher discriminant scores than do the clusters of healthy central cities.

The group of stressed central cities is displayed along the top half of Figure 1. A bit less than half of US central cities (44 per cent) are stressed. Most prominent among these cities are the set of 54 cities, mostly large places, that form the stereotype of American central cities. These are coupled with the three clusters of central cities that form the set of manufacturing central cities to form the group of stressed central cities. These stressed central cities are contrasted with 284 municipalities that are healthier. Within the healthy group is a distinct set of 57 central cities we call knowledge-based cities, which are the most viable of all central cities in the country. They are joined by a set of Sunbelt growth poles and a large number of cities, constituting nearly 30 per cent of all central cities, that we term the competitive core of the economy.

The discriminant analysis indicates 'misclassified' cases as part of its analysis of the prior distribution; that is, it indicates which cases in the cluster groupings would be assigned to another cluster, based on the discriminant functions. We therefore have two groupings of cases: one from the cluster analysis (called the cluster groupings) and the other from the discriminant analysis of the 13-cluster solution (the discriminant groupings).

We do not use the discriminant groupings

directly in the analysis, because they are not hierarchical. The discriminant group assignments change with each prior distribution used; in progressing from the 13- to the 5- and then to the 2-cluster solutions, the prior distributions for each discriminant analysis will change. With changes in the prior distributions, the number of discriminant functions will change; as those functions change, so will the suggested discriminant group assignments. The advantage of using the cluster analysis is that it is hierarchical. We can follow a particular cluster of central cities from its assignment in the 13-cluster solution to its more heterogeneous set in the 5-cluster solution, to the even more heterogeneous cluster group in the 2-cluster solution.

We use the information from the discriminant analysis in three ways. First, we use the hit ratio from the discriminant analysis for each of the candidate cluster solutions to test the internal validity of the cluster solutions as a whole (as reported in Figure 1). Then we use the detailed results from the discriminant analysis associated with the 13-cluster solution to help array the clusters in the figures and tables. Finally, we perform *t*-tests to determine whether the means of the discriminant scores of each of the cluster groups were significantly different from the means of their paired discriminant groups in the 13-cluster solution. To do this, we calculate for each cluster (in both the 13- and 5-cluster solutions) the mean and standard deviation of the discriminant score associated with each central city in that cluster. We similarly calculate the mean score and standard deviation for its paired discriminant group. If the means are not significantly different from each other, there is no statistically meaningful difference between the two groups, providing internal validity to each of the cluster group assignments in the 13-cluster solution.

There is no statistically meaningful difference between any of the pairs of clusters at the 0.10 level of significance.⁶ Therefore, the cluster groupings and their paired discriminant groupings are statistically equivalent. Additionally, the *F*-test of the cluster assignments in the 13-cluster solution that is part of

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

the cluster analysis indicates that each cluster group is statistically different from the other cluster groups.⁷ These statistical tests demonstrate that the clusters are independent of each other (from the *F*-test) and that all the clusters are statistically valid (from the *t*-tests).

We use the information on the cluster and discriminant group assignments to construct Table 3.⁸ The columns consist of the cluster assignments in the 13-cluster solution; the rows are the assignments from the discriminant analysis. The diagonal elements in the table consists of those cases in which the assignments agree; the off-diagonal elements are cases in which the assignments differ. There are two ways to think about the off-diagonal elements. One is to consider these to be cases in which the cluster technique made mistaken assignments. This would be appropriate if the clusters were mutually exclusive objective categories and, in this article, they are not. However, it is also possible that these off-diagonal cases are the makings of latent or emerging clusters. Thinking in this way is desirable because cities are multi-dimensional, as is our assignment technique, and they form a continuum; however, that continuum is difficult to map in one-dimensional space.⁹ In fact, we are working in 12 dimensions in the discriminant analysis (the number of discriminant functions) and in 33 dimensions in the cluster analysis (the number of variables used).¹⁰

4.2 *The Discriminant Functions*

The discriminant analysis of the 13-cluster solution produced 12 discriminant functions (Table 4). Each of the discriminant functions is significantly different from zero at the 0.01 critical level. Table 4 first lists the proportion of the explained variation in the discriminate analysis that is accounted for by each discriminant function. For example, the first discriminant function is responsible for 29 per cent of the variation accounted for by the discriminant analysis. The table then lists the correlation between each variable in a dis-

criminant function and that function. For example, the correlation between the first variable listed in Function 1, the percentage of adults with a terminal graduate degree—has a negative correlation with the first discriminant function (-0.70). This means that there is an inverse relationship between the proportion of adults holding advanced degrees and the first function. After examining the median discriminant scores reported for each cluster group in Table 3 and the mean scores given in Note 8, it is clear that high discriminant scores are associated with economic distress and low-quality economic resources, whereas low (negative) scores are associated with economic success and high-quality resources.

The first two discriminant functions dominate the analysis and are positively associated with weak economic resources and outcomes. We call the first discriminant function a Low human capital function, because it consists of educational attainment and occupational variables. This function accounts for 29 per cent of the explained variation in the discriminant analysis. There is a negative correlation between this function and desirable educational and occupational traits, such as completing education beyond secondary school and the proportion of residents holding symbolic analytical positions. (This variable is defined in Table 1.) The second function highlights the central city's characteristics, and we call it a Decline function. As is true with the first function, the decline function is negatively correlated with what are usually seen as beneficial economic resources. This function is negatively correlated with two higher educational attainment variables: the rate of change in central-city and MSA populations, and the ratio of central city to MSA median family income. This function accounts for 21 per cent of the explained variation in the data.

The third function is a High poverty function, accounting for 12 per cent of the explained variation. This function is positively correlated with the poverty rate (a social outcome) and with the percentage of the

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

Table 3. Classification of central cities by cluster and discriminant analyses (median discriminant score for each cluster group is given under each group name)

Cluster name, number and median discriminant score														
Cluster discriminant score:	All central cities 0.095													
Group discriminant score:	Stressed central cities 0.660							Healthy central cities -0.475						
Set discriminant score:	Stereotypical 0.985		Manufacturing 0.590			Competitive 0.000			Sunbelt -1.17		Knowledge -0.900			
Cluster discriminant score:	Extreme Cluster 1	Distressed Cluster 2	Larger Cluster 3	Small, Less success	Small, More success	Heartland Cluster 6	Administrative Cluster 7	Border TX Cluster 8	Smaller Cluster 9	Larger Cluster 10	Edge Cluster 11	State university Cluster 12	Research Cluster 13	Row Total
	1.540	0.800	0.890	0.700	-0.020	0.200	-0.640	0.510	-1.100	-1.220	0.010	-0.990	-1.650	
Discriminant Group														
Extreme distress 1	11													11
Distressed 2		36	1	1		1								39
Larger manufacturing 3		1	37	1							1			40
Small, less success, manufacturing 4		1	4	64	3	6								78
Small, more success, manufacturing 5			2	3	44	1					2			52
Heartland 6		2	3	5	2	65	5		2			2		86
Administrative 7		2				7	51		2					62
Border Texas 8								7						7
Smaller Sunbelt 9						2	1		45			1		51
Larger Sunbelt 10									4	23				27
Edge 11											14			14
State universities 12		1				1	2					27	1	32
Research 13													9	9
Column Total	11	43	47	74	49	83	59	7	53	25	17	30	10	508

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

Table 4. Correlations between the discriminating variables and the canonical discriminant functions

Functions	Correlation coefficient
<i>Function 1: Low human capital</i>	29.0
Percentage with graduate degree	- 0.70
Percentage with some high school	0.61
Percentage with terminal bachelors degree	- 0.58
Percentage in symbolic analytical occupations	- 0.51
Percentage in machinist and precision occupations	0.43
Percentage with terminal high school degree	0.35
Percentage in precision producers occupations	0.33
<i>Function 2: Decline</i>	21.0
Percentage increase in central city population 1980-90	- 0.47
Median age of the housing stock	0.45
Percentage increase in MSA population 1980-90	- 0.44
Central city to MSA median family income ratio	- 0.41
Percentage terminate education with some college	- 0.34
Percentage with associates degree	- 0.27
<i>Function 3: High poverty</i>	12.2
Poverty rate	0.44
Percentage with less than 9th grade education	0.43
<i>Function 4: Metropolitan interdependence</i>	10.1
Percentage out-commuters	0.50
Central city population as percentage of MSA population	- 0.47
<i>Function 5: Industrial city</i>	9.1
Percentage employed in manufacturing industries	0.50
Monthly housing cost for renters	- 0.47
Percentage employed in service industries	- 0.42
Percentage in service occupations	- 0.42
Logarithm of central city population density	- 0.38
<i>Function 6: Tight labour market</i>	7.4
Logarithm of central city per capita income	0.57
Female labour force participation rate	0.42
Average commuters' travel time	0.41
Percentage employed in labourer occupations	- 0.40
Male labour force participation rate	0.40
Logarithm of MSA population	0.39
<i>Function 7: Small employment base</i>	3.6
Ratio: employment in central city to employed residents	- 0.29
<i>Function 8: High unemployment rate</i>	3.1
Unemployment rate	0.44
<i>Function 9</i>	2.5
<i>Function 10: Low percentage African-American</i>	1.1
Percentage African-American	- 0.39
Percentage employed in wholesale and retail industries	0.38
<i>Function 11</i>	0.7
<i>Function 12: Large percentage public service employment</i>	0.3
Percentage employed in health, education and government occupations	0.51

All variables for the central city unless noted. The percentage of the total variance explained by each function is displayed in bold in the second column.

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

adult population with less than a 9th grade education.

The fourth function measures Metropolitan interdependence; it gauges the degree to which the central city is integrated into the metropolitan economy, and it is responsible for 10 per cent of the explained variation. The metropolitan interdependence function is positively correlated with the percentage of central city residents who commute out of that central city for work; it is negatively correlated with the percentage of the MSA's population that reside in that central city. Therefore, the function increases in value as the percentage of the MSA population living in that central city decreases (there is often more than one central city in a metropolitan area). In the aggregate, this discriminant function increases in value—with negative implications for residents—as the central city's economic importance within the MSA decreases.

The fifth function captures the essence of an Industrial central city and accounts for 9 per cent of the explained variation. This function is positively correlated with the percentage of the workforce employed in manufacturing industries; it is negatively correlated with the percentage employed in service industries and occupations, population density and the monthly housing cost for renters (the lower the rent, the higher the value of this function—a convincing indicator of decline).

The sixth function contains many of the indicators of a Tight labour market, accounting for 7.4 per cent of the explained variation. The labour market function is negatively associated with the percentage of the population employed as labourers; it is positively associated with central city per capita income, the female and male labour force participation rates, the population of the MSA and the average length of commuting time for central city residents. This function is different from the others, because a positive value is desirable; the function is associated with tight labour market conditions, good economic outcomes and a larger MSA.

The seventh function consists of one of the Bureau of the Census' qualifying variables and reflects a relatively Small employment base. This qualifying function is negatively correlated with the ratio of employment in the central city to the number of employed residents; it is a sign of a weak employment base in the central city relative to the number of workers who live there. However, a low ratio can apply to three types of municipalities: declining inner-ring suburbs with significant nodes of employment; industrial central cities with significant residential poverty populations; or emergent edge cities with a relatively small (but increasing) number of jobs, coupled with a well-off white-collar commuter workforce. Due to this indeterminacy, the function only accounts for 3.6 per cent of the explained variation.

The eighth function is positively associated with the High unemployment rates in central cities, accounting for a bit more than 3 per cent of the explained variation. The tenth function is correlated with the Low percentage of African-Americans in the population of the central cities and with the percentage of the population employed in wholesale and retail industries. It has little explanatory power. The twelfth function is positively correlated with high percentages of the central city workforce employed in Public service industries, including education and health care. The ninth and eleventh functions were not significantly associated with any particular variable, even though the functions as a whole were significantly different from zero in their effect on the clusters.¹¹

Discriminant analysis differentiates among the clusters, based on their association with each of the discriminant functions. For each discriminant function, we measure the association between the mean value of the discriminant scores for each cluster against the mean value of the discriminant scores for all cities on that function, by examining the *z*-scores for each cluster. The *z*-scores identify the clusters that are significantly different from the mean of the universe for a particular discriminant function, thus recognising the discriminant functions that best describe each

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

cluster.¹² Table 5 lists the *z*-scores of the relationship between each discriminant function and each cluster, evaluating the score at the 99 per cent, 95 per cent, and 90 per cent critical values. Discriminant functions 10–12 are not significantly associated with any of the clusters. Several of the clusters are significantly associated with the same discriminant function, but each is significantly associated with a unique combination of those discriminant functions.

One of the clusters serves as the ‘average’ cluster, in which the mean value of each discriminant function for that cluster is not statistically different from the mean value of those functions for the universe. In our case, the average cluster is the sixth, Heartland central cities. The Heartland cluster is not associated with any of the discriminant functions. This is also the cluster that had the largest number of overlaps with the other clusters in Table 3. Next, we use the discriminant functions and descriptive statistics to differentiate among the 13 types of central cities.

5. Interpreting the Clusters

There are two large groups of central cities in the US—one stressed, the other healthy. Granted, all central cities house the poor, many have higher unemployment rates than their suburban neighbours do, and nearly all face the full range of urban ills. Yet the data speak for themselves: About half of all central cities do not fit the stereotypical image of a central city. The problem lies with both the image itself and with the lack of alignment between the image and how central cities are defined.

The image for all central cities is based largely on the spatial concentration of social pathologies—outcomes; conversely, the Bureau of the Census’ definition is based on a fairly small population threshold, commuting patterns and the number of jobs contained in the municipality relative to the number of its employed residents. In other words, that definition is largely based on the economic functions of central cities. Using the term

‘central city’ as shorthand for the spatial concentration of poverty, racial isolation, unemployment, industrial abandonment and brownfields is an exercise in both specification error and mushy thinking, due to the way central cities are defined for statistical record-keeping purposes.

The two large groups of central cities are composed of 13 clusters. Five clusters form the stressed group and six the healthy group. We select the median central city from each cluster and display it on a cluster map (Figure 2) to provide a better feel for the types of places in each of the clusters. The second line in each cluster box contains the name of the median city (two are listed if the cluster has an even number of cities), and the bottom line contains the city that immediately precedes the median central city as well as that immediately following.¹³ The median central city for each cluster is determined by ranking all cities in the cluster by summary discriminant scores. These summary scores are calculated by weighing each city’s discriminant score on a particular discriminant function by the percentage of the variation that is explained by that function (see Table 4), then summing the weighted values for the city across all the discriminant functions.

Norwalk, Connecticut, and Columbia, South Carolina, are the median central cities for the universe of 508 central cities; they are bracketed by Fort Wayne, Indiana, and Elgin, Illinois. The median places among the group of stressed central cities are Akron, Ohio, and New Britain, Connecticut. These municipalities are bracketed by Anniston, Alabama, and Santa Ana, California. The median places among the healthy group of central cities are Lawton, Oklahoma, and Bossier, Louisiana. Their immediate neighbours on our lists are Lubbock, Texas, and Arlington, Virginia. If your immediate reaction is that some of these places are not ‘really’ central cities—if Anniston, Fort Wayne or Lawton do not fit your image of what a central city is—we have made our point. In the remainder of this article, we discuss the characteristics of the clusters.

We use five tables (Tables 5–9) to identify

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

Table 5. Z-scores of the canonical discriminant functions evaluated at the cluster means (two-tailed test)

Cluster Group	Discriminant Function											
	Low human capital (1)	Decline (2)	High poverty (3)	Metro inter dependence (4)	Industrial city (5)	Tight labour market (6)	Small employment base (7)	High unemployment rate (8)	9	Low percentage African-American (10)	11	Large percentage public service employment (12)
Extreme distress 1	4.15***	3.52***	-0.69	0.02	-2.96***	-2.29**	0.16	3.24***	1.02	-0.15	-0.10	0.11
Distress 2	1.26	1.88**	0.65	-0.16	-1.71**	1.59	0.16	-0.88	0.45	-0.63	-0.01	-0.06
Larger manufacturing 3	2.16**	0.49	-0.26	0.98	-0.01	0.39	1.52	-0.16	-0.77	0.51	-0.07	0.11
Small, less success, manufacturing 4	1.81*	0.87	-0.21	0.25	0.81	-0.92	-0.55	-0.35	-0.06	0.10	0.34	-0.26
Small, more success, manufacturing 5	0.74	-1.01	-1.44	0.29	1.97**	-0.13	0.19	0.17	0.13	-0.70	-0.42	0.06
Heartland 6	0.07	0.53	0.88	-0.90	0.44	-0.10	-0.66	-0.20	0.12	0.25	-0.07	0.33
Administrative 7	-1.30	-0.88	1.21	-1.86*	0.30	0.67	0.37	0.79	-0.55	-0.14	0.19	-0.16
Border Texas 8	-1.47	-2.97***	8.03***	6.69***	0.79	-0.72	0.02	0.70	0.53	-0.47	0.10	0.04
Smaller Sunbelt 9	-0.52	-2.81***	-0.45	0.08	-1.96**	-0.97	-0.48	-0.34	-0.47	0.04	-0.31	-0.07
Larger Sunbelt 10	-1.25	-3.31***	-2.10**	0.61	-0.48	0.54	0.45	0.06	1.02	0.03	0.91	0.25
Edge 11	-1.28	0.71	-1.72*	2.21**	0.15	3.51***	-1.74*	1.13	-0.11	0.65	-0.29	-0.19
State universities 12	-4.00***	1.05	0.07	-0.37	0.34	-0.66	0.93	-0.30	1.33	0.54	-0.39	-0.26
Research 13	-8.99***	5.91***	-1.85*	1.99**	-0.58	-1.55	-0.04	-0.08	-1.66*	-0.71	0.40	0.26
1	2.16	0.49	-0.26	0.39	-0.01	0.39	1.52	-0.16	-0.77	0.51	-0.07	0.11
2	-1.28	0.71	-1.72	2.21	0.15	3.51	-1.74	1.13	-0.11	0.65	-0.29	-0.19
3	0.74	-1.01	-1.44	0.29	1.97	-0.13	0.19	0.17	0.13	-0.70	-0.42	0.06
4	1.26	1.88	0.65	-0.16	-1.71	1.59	0.16	-0.88	0.45	-0.63	-0.01	-0.06
5	0.07	0.53	0.88	-0.90	0.44	-0.10	-0.66	-0.20	0.12	0.25	-0.07	0.33
6	-1.30	-0.88	1.21	-1.86	0.30	0.67	0.37	0.79	-0.55	-0.14	0.19	-0.16
7	1.81	0.87	-0.21	0.25	0.81	-0.92	-0.55	-0.35	-0.06	0.10	0.34	-0.26
8	-0.52	-2.81	-0.45	0.08	-1.96	-0.97	-0.48	-0.34	-0.47	0.04	-0.31	-0.07
9	4.15	3.52	-0.69	0.02	-2.96	-2.29	0.16	3.24	1.02	-0.15	-0.10	0.11
10	-8.99	5.91	-1.85	1.99	-0.58	-1.55	-0.04	-0.08	-1.66	-0.71	0.40	0.26
11	-4.00	1.05	0.07	-0.37	0.34	-0.66	0.93	-0.30	1.33	0.54	-0.39	-0.26
12	-1.47	-2.97	8.03	6.69	0.79	-0.72	0.02	0.70	0.53	-0.47	0.10	0.04
13	-1.25	-3.31	-2.10	0.61	-0.48	0.54	0.45	0.06	1.02	0.03	0.91	0.25

***is at the 99.0 per cent confidence interval (2.57 critical value)

**is at the 95.0 per cent confidence interval (1.96 critical value)

*is at the 90.0 per cent confidence interval (1.65 critical value)

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

the important distinguishing characteristics of each cluster. The discriminant factors that are statistically associated with each cluster, listed in Table 5, are key to identifying the differences among the clusters. In Table 6, we calculate a version of the location quotient, which we call the regional concentration ratio (RCR), to determine whether any of the clusters of central cities are predominantly located in particular regions of the nation. The RCR is calculated by taking the percentage distribution of clusters in each Census division and dividing it by the percentage distribution of clusters in the nation as a whole. If the RCR of a cluster in a division is equal to 1.0, its presence in the division is proportionate to that cluster's representation in the nation; if the RCR is less than 1.0, the cluster is not as prominent in the region as it is nationally; if the RCR exceeds 1.0, it is overrepresented in the region. The more specialised a region is in a particular cluster, the higher the RCR. In Table 6, we mark three levels of 'cluster specialisation': between one and one-and-a-half times the national proportion, between one-and-a-half and twice the national proportion, and at least twice the national proportion.

We arrayed the Census divisions in Table 6 so that, as you read from left to right, the divisions generally move from west to east. The healthiest clusters are located at the bottom of the table, the more distressed clusters of central cities at the top. There is a regional pattern to the clusters. As you move from west to east and from south to north, the conditions of the central cities deteriorate. The exception to this general pattern is found in the East South Central and South Atlantic Census divisions; both these Divisions contain a broad mix of central cities—including some Distressed central cities. Our interpretation of this national pattern is that there is an association between the condition of central cities and their economic age—the time when they experienced their greatest growth. There are a few older central cities in the South.

We then list the median values for 15 of the variables for the universe, as well as for

each of the clusters, in Table 7, along with the regional distribution of the clusters. Although we know that each cluster differs from the universe, based on the discriminant factors, we do not know whether all the variables in each cluster differ from the central tendency of the universe—or, more important, whether a critical policy or descriptive variable (such as the poverty status of the population or the size of the municipality) in a specific cluster is different from that of the universe. Table 8 contains the results of pseudo *t*-tests, testing the null hypothesis that the median value of each variable listed in Table 7, for each cluster, is the same as for the universe of central cities.¹⁴ We also wanted to know whether the medians of the variables listed in Table 7 for each cluster were significantly different from each other. To determine this, we took advantage of the order that exists among the clusters to examine the null hypothesis that the median for each variable in each cluster is not statistically different from the median of the same variable in its adjoining cluster—again, using the pseudo *t*-test.¹⁵ If these medians are statistically different, the median of the cluster and of clusters that they do not adjoin, by transitivity, will also be statistically different.

Table 9 contains the results of these pseudo *t*-tests. In two instances, the proper order among the clusters was not obvious. This is the case for the Manufacturing central cities set of clusters, where we tested for differences between all clusters in this set: Larger manufacturing central cities (cluster 3); Smaller, less successful manufacturing cities (cluster 4); and Smaller, more successful central cities (cluster 5). We also tested for differences between the Smaller, less successful manufacturing central cities (cluster 4) and the Heartland central cities (cluster 6), as well as between the cluster of Smaller, more successful central cities (cluster 5) and the Heartland cluster (cluster 6).

Each of the clusters can now be evaluated, based on the data developed above. Because of its policy importance we spend more time

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

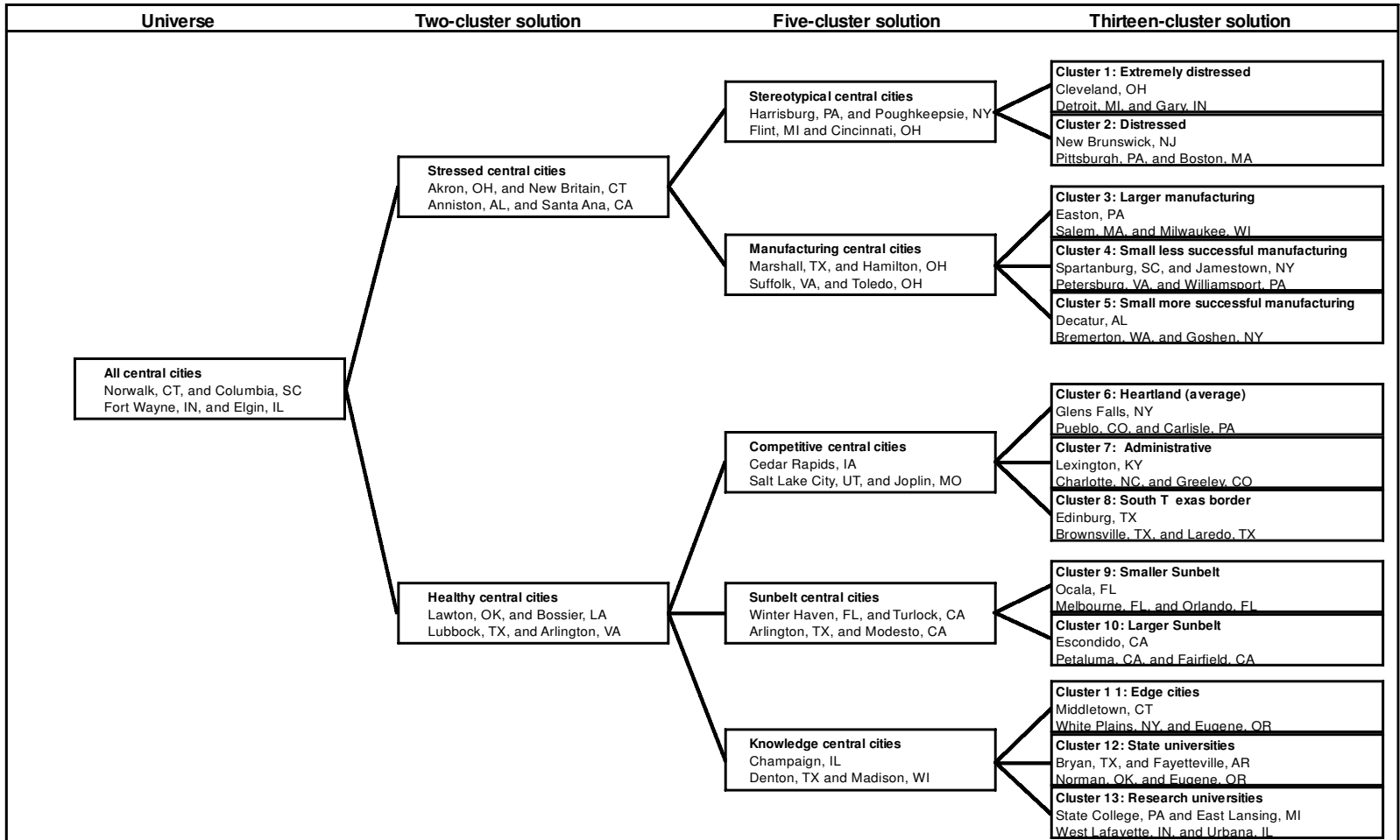


Figure 2. Median central city in each cluster, by discriminant score. *Notes:* The second row of each cell contains the median central city; two are listed if the cell has an even number of cities; the bottom row contains the cities that precede and follow the median cities.

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

Table 6. Regional Concentration of the Clusters

Cluster	Pacific	Mountain	West North Central	West South Central	East South Central	South Atlantic	East North Central	Middle Atlantic	New England
Extreme distress 1	0.0	0.0	1.0*	0.0	0.0	0.0	3.9***	1.8**	0.0
Distress 2	0.7	0.0	0.8	0.4	1.1*	1.2*	0.4	3.5***	1.5**
Larger manufacturing 3	1.1*	0.0	1.0*	0.0	0.0	0.5	0.7	2.5***	4.7***
Small, less success, manufacturing 4	0.1	0.0	0.2	0.9	2.3***	0.7	2.2***	1.7**	0.0
Small, more success, manufacturing 5	0.8	0.7	0.0	0.8	0.9	1.3*	1.4*	0.4	2.6***
Heartland 6	0.3	0.7	1.8**	1.3**	1.9**	1.3*	1.0*	0.5	0.2
Administrative 7	0.6	3.4***	2.7***	2.2***	1.0*	0.5	0.4	0.0	0.0
Border Texas 8	0.0	0.0	0.0	7.9***	0.0	0.0	0.0	0.0	0.0
Smaller Sunbelt 9	2.9***	1.0*	0.0	0.6	0.3	2.5***	0.0	0.0	0.0
Larger Sunbelt 10	3.2***	2.9***	1.4*	0.6	0.0	1.1*	0.0	0.0	0.0
Edge 11	2.2***	0.0	0.0	0.5	0.9	0.7	0.6	1.2*	3.7***
State universities 12	1.5**	2.4***	2.3***	1.1*	0.0	0.4	1.2*	0.0	0.5
Research 13	0.7	1.8**	0.0	0.8	0.0	0.6	2.7***	1.0*	0.0

Notes: Regional concentration ratio (RCR) is defined as the percentage distribution of clusters within a region divided by the percentage distribution in the nation. The value 1.0 indicates that the cluster's representation in the region is proportionate with that cluster's representation in the nation. A value in excess of 1.0 indicates that the cluster is overrepresented, and a value less than 1 indicates that it is underrepresented.

***RCR > = 2.0.

**1.5 > = RCR < 2.0.

*1.0 > = RCR < 1.5.

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

Table 7. Median characteristics of the clusters

Variable	Stressed central cities						Healthy central cities								
	Stereotypical central cities		Manufacturing central cities				Competitive central cities			Sunbelt central cities		Knowledge central cities			
	All Central Cities	Extreme Cluster 1	Distressed Cluster 2	Larger Cluster 3	Small, less Successful cluster 4	Small, more Successful cluster 5	Heartland Cluster 6	Administrative Cluster 7	Border TX Cluster 8	Smaller Cluster 9	Larger Cluster 10	Edge Cluster 11	State university Cluster 12	Research Cluster 13	
Number of central cities in the cluster	508	11	43	47	74	49	83	59	7	53	25	17	30	10	
Percentage distribution of central cities among the clusters	100.0	2	8	9	15	10	16	12	1	10	5	3	6	2	
<i>Demographic Variables</i>															
Central city population	63 518	95 732	265 852	99 922	35 509	46 380	70 580	175 781	48 735	59 646	86 335	55 900	56 461	48 443	
MSA population	383 545	1 824 321	981 747	1 162 093	244 149	315 121	239 971	348 428	383 545	370 712	1 332 053	1 238 816	177 572	199 182	
Percentage change central city population 1980-90	3.9	-14.7	-4.0	1.4	-4.0	8.8	-1.9	10.9	21.1	24.7	29.1	2.2	12.9	11.1	
Percentage change MSA population 1980-90	8.2	-3.6	5.8	6.7	0.2	10.6	4.6	10.8	35.4	30.8	34.2	3.9	17.3	10.1	
Central city population as percentage of MSA population	21.0	17.4	19.1	9.6	15.7	15.6	35.6	60.8	18.7	15.4	11.1	4.4	34.8	26.2	
Percentage of central city population African-American	10.7	56.3	31.5	11.4	12.9	7.6	17.2	7.6	0.3	12.8	2.8	10.7	2.5	5.2	
<i>Income and Employment Variables</i>															
Central city to MSA income ratios	87.5	58.3	69.3	79.4	81.1	93.4	88.6	96.6	104.7	94.0	104.6	102.8	90.2	80.3	
Unemployment rate	7.10	16.7	9.0	8.4	8.5	5.9	7.3	6.1	12.2	7.1	4.8	5.1	5.8	5.6	
Poverty rate	9.6	18.1	12.3	9.4	10.9	6.2	10.6	8.2	22.4	8.2	4.4	4.6	11.5	15.5	
Employees to residents ratio	1.4	1.6	1.7	1.2	1.4	1.2	1.5	1.3	1.4	1.3	1.0	1.5	1.3	1.6	
Percentage out-commuter rate	36.7	50.4	33.9	53.4	42.4	46.2	23.3	15.6	31.3	42.7	55.6	54.1	31.9	31.1	
Percentage symbolic analysts	41.6	28.7	39.7	35.2	35.4	38.2	42.2	45.4	40.7	41.8	46.2	55.4	50.3	61.2	
Percentage manufacturing jobs	15.9	20.7	13.6	24.7	23.3	27.0	14.4	12.2	10.2	10.0	14.9	16.7	10.0	6.2	
<i>Educational attainment of adult population</i>															
Percentage with some high school	15.0	25.8	18.7	18.7	19.0	14.8	15.2	12.0	13.9	14.5	10.2	11.6	7.9	3.1	
Percentage terminal bachelor's degree	12.3	5.7	12.0	9.0	7.5	11.4	12.5	16.9	8.4	11.7	16.6	20.7	20.7	29.5	
<i>Percentage distribution of clusters by census region</i>															
Pacific	14	0	9	15	1	10	4	8	0	40	44	29	20	10	
Mountain	6	0	0	0	0	4	4	19	0	6	16	0	13	10	
West North Central	9	9	7	9	1	0	16	24	0	0	12	0	20	0	
West South Central	13	0	5	0	11	10	17	27	100	8	8	6	13	10	
East South Central	6	0	7	0	15	6	12	7	0	2	0	6	0	0	
South Atlantic	18	0	21	9	12	22	23	8	0	45	20	12	7	10	
East North Central	19	73	7	13	42	27	19	7	0	0	0	12	23	50	
Middle Atlantic	10	18	35	26	18	4	5	0	0	0	0	12	0	10	
New England	6	0	9	30	0	16	1	0	0	0	0	24	3	0	

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

Table 8. Results of pseudo *t*-tests of the difference between medians for the universe of central cities and each cluster

Variables	Stressed central cities					Healthy central cities							
	Stereotypical central cities		Manufacturing central cities			Competitive central cities			Sunbelt central cities		Knowledge central cities		
	Extreme 1	Distressed 2	Larger 3	Smaller, less successful 4	Smaller, more successful 5	Heartland 6	Administrative 7	Border TX 8	Smaller 9	Larger 10	Edge 11	State university 12	Research 13
<i>Demographic variables</i>													
Central city population	-1.61	-14.22**	-3.52**	3.71**	1.84	-0.96	-8.10**	0.60	0.42	-1.73	0.48	0.59	0.74
MSA population	-7.23**	-6.02**	-7.57**	2.02*	0.80	2.22*	0.43	0.00	0.15	-7.36**	-4.94**	1.93	1.00
Percentage change central city population 1980-90	4.59**	3.81**	1.28	4.98**	-2.48*	3.89**	-3.86**	-3.37**	-10.56**	-9.12**	0.54	-3.62**	-1.66
Percentage change MSA population 1980-90	2.90**	1.12	0.75	5.08**	-1.23	2.32*	-1.44	-5.37**	-11.62**	-9.49**	1.32	-3.64**	-0.44
Central city population as percentage of MSA population	0.53	0.55	3.45**	2.00*	1.68	-5.55**	-12.82**	0.27	1.82	2.19*	3.05**	-3.29**	-0.73
Percentage of central city population African-American	-8.94**	-7.58**	-0.31	-1.09	1.25	-3.13**	1.40	1.65	-0.92	2.36*	-0.02	2.69**	1.05
<i>Income and employment variables</i>													
Central city to MSA income ratios	6.73**	8.27**	3.83**	3.73**	-2.81**	-0.70	-4.86**	-3.17**	-3.21**	-5.86**	-4.34**	-1.02	1.58
Unemployment rate	-13.27**	-5.27**	-3.65**	-4.70**	3.34**	-0.96	3.05**	-5.64**	0.03	4.79**	3.44**	3.02**	2.00*
Poverty rate	-7.64**	-4.74**	0.40	-2.93**	6.33**	-2.37*	2.82**	-9.26**	2.74**	7.08**	5.63**	-2.79**	-5.01**
Employees to residents ratio	-2.12*	-4.35**	3.24**	-0.38	2.53*	-3.52**	1.81	0.00	0.99	4.39**	-1.46	0.45	-1.55
Percentage out-commuter rate	-2.26*	0.90	-5.59**	-2.40*	-3.19**	5.98**	7.98**	0.71	-2.14*	-4.68**	-3.55**	1.28	0.88
Percentage symbolic analysts	5.97**	1.79	5.91**	7.39**	3.33**	-0.73	-3.92**	0.36	-0.18	-3.13**	-7.86**	-6.50**	-8.55**
Percentage manufacturing jobs	-1.91	1.83	-7.06**	-7.47**	-9.10**	1.57	3.41**	1.83	5.16**	0.63	-0.40	3.93**	3.72**
<i>Educational attainment of adult population</i>													
Percentage with some high school	-7.45**	-4.97**	-5.25**	-7.02**	0.27	-0.35	4.61**	0.62	0.73	4.90**	2.86**	8.08**	7.79**
Percentage terminal bachelor's degree	3.97**	0.35	4.07**	7.33**	1.08	-0.35	-6.24**	1.87	0.77	-3.82**	-6.16**	-8.22**	-9.77**

*exceeds the 95 per cent critical value of the *t* distribution; **exceeds the 99 per cent critical value.

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

Table 9. Results of pseudo *t*-tests of the difference between medians for the indicated pairs of clusters

	Pseudo <i>t</i> -test between the following pairs of clusters													
	1 and 2	2 and 3	3 and 4	3 and 5	4 and 5	4 and 6	5 and 6	6 and 7	7 and 8	8 and 9	9 and 10	10 and 11	11 and 12	12 and 13
Degrees of freedom ($N1 + N2 - 2$)	52	88	119	94	121	155	130	140	64	58	76	40	45	38
<i>Demographic variables</i>														
Central city population	- 2.35*	4.47**	5.59**	3.75**	- 2.90**	- 6.65**	- 3.68**	- 3.76**	1.33	- 0.46	- 1.83	1.80	- 0.05	0.87
MSA population	1.91	- 0.73	5.90**	4.34**	- 1.19	0.12	1.51	- 1.55	- 0.15	0.07	- 5.02**	0.17	2.55*	- 0.43
Percentage change central city population 1980-90	- 4.99**	- 4.05**	4.53**	- 5.10**	- 9.64**	- 1.83	7.86**	- 9.76**	- 3.22**	- 0.62	- 1.23	7.62**	- 5.34**	0.64
Percentage change MSA population 1980-90	- 3.42**	- 0.44	4.92**	- 1.99*	- 7.52**	- 3.46**	3.35**	- 3.69**	- 6.63**	0.84	- 1.01	8.41**	- 4.02**	1.58
Central city population as percentage of MSA population	- 0.43	3.89**	- 2.56*	- 2.49*	0.04	- 7.13**	- 6.25**	- 6.86**	4.86**	0.72	1.61	2.57*	- 6.55**	1.26
Percentage of central city population African-American	3.02**	5.01**	- 0.53	1.42	1.86	- 1.33	- 2.63**	3.03**	2.10*	- 2.45*	3.63**	- 3.51**	4.30**	- 2.30*
<i>Income and employment variables</i>														
Central city to MSA income ratios	- 4.36**	- 6.29**	- 1.01	- 6.46**	- 6.17**	- 5.02**	2.52*	- 5.60**	- 2.49*	2.36*	- 3.70**	0.37	3.54**	3.31**
Unemployment rate	12.82**	1.94	- 0.24	7.47**	6.65**	3.49**	- 4.93**	4.60**	- 9.70**	66.47**	5.54**	- 1.02	- 2.47*	0.44
Poverty rate	6.14**	5.17**	- 3.05**	5.88**	9.56**	0.73	- 9.45**	5.86**	- 15.46**	15.63**	7.88**	- 0.37	- 10.40**	- 3.16**
Employees to residents ratio	- 0.13	6.50**	- 3.02**	- 0.51	2.08*	- 2.31*	- 4.53**	5.10**	- 1.00	0.34	3.14**	- 5.05**	1.88	- 2.24*
Percentage out-commuter rate	3.65**	- 6.64**	4.93**	2.29*	- 1.49	11.88**	9.89**	4.92**	- 3.75**	- 2.01*	- 4.27**	0.47	5.00**	0.14
Percentage symbolic analysts	- 6.69**	3.56**	- 0.13	- 2.44*	- 3.42**	- 10.14**	- 4.77**	- 4.28**	2.95**	- 0.61	- 3.44**	- 5.22**	3.23**	- 5.26**
Percentage manufacturing jobs	4.88**	- 8.58**	1.08	- 1.50	- 2.91**	9.58**	10.88**	2.77**	1.46	0.15	- 6.03**	- 1.22	4.40**	2.75**
<i>Educational attainment of adult population</i>														
Percentage with some high school	7.15**	- 0.09	- 0.45	6.02**	9.05**	10.70**	- 0.77	6.87**	- 1.58	- 0.60	6.46**	- 1.28	4.38**	9.26**
Percentage terminal bachelor's degree	- 5.09**	4.08**	3.23**	- 4.14**	- 8.83**	- 12.00**	- 2.15*	- 8.85**	6.95**	- 3.25**	- 5.23**	- 2.19*	0.05	- 7.70**

*exceeds the 95 per cent critical values for the *t*-distribution; **exceeds the 99 per cent critical value.

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

discussing the stereotypical set of central cities than the other sets.

5.1 Stressed Central Cities

Stereotypical central cities: large and extremely troubled places. The stereotype of US central cities comes from the 54 central cities in the first two clusters of municipalities. We label the 11 cities in the first cluster Extremely distressed and the 43 in the second cluster as Distressed. Such central cities are large and come from large metropolitan areas, although these are not their most important distinguishing characteristics; more pertinent is the economic distress of their populations. These cities share significant positive values for the decline discriminant function and significant negative values for the industrial city function. Thus such central cities are characterised by population losses, both in the central city and the MSA—the central city being, on average, much poorer than the average of the MSA. Surprisingly, the percentage of their populations employed in manufacturing is statistically indistinguishable from that of the universe (Table 8), even though these places often have manufacturing heritages; most likely, those jobs left before the 1990 Census was conducted.

The 11 central cities in the Extremely distressed cluster differ from the Distressed cluster by their significantly positive association with the low human capital discriminant function, poor labour market conditions and high unemployment rates (see the fifth, sixth and eighth discriminant functions). These central cities have the highest median unemployment and poverty rates (Table 7); such rates are significantly higher than for the universe as a whole (Table 8) and for the cluster of Distressed central cities (Table 9). The cause of high levels of unemployment rests in part with the human capital characteristics of these Extremely distressed central cities; compared to the other clusters they have the highest proportion of their adult populations terminate their education as high-school dropouts. The positive association with the human capital discriminant

function means that adults in these municipalities have low levels of educational attainment and that a small proportion of their population hold symbolic analytical positions.

The Extremely distressed cluster includes the cities of Camden, Cleveland, East St Louis, Gary, Newark and Detroit. Regionally, this cluster is concentrated in the East North Central and Middle Atlantic Census divisions (Tables 6 and 7). Five of these 11 central cities are located in Michigan; besides Detroit, they consist of Benton Harbor, Pontiac, Flint and Saginaw. Common to all cities in the Extremely distressed cluster are a low-skilled labour force, an isolated African-American population and low employment levels of people in symbolic analytical occupations.

The 43 central cities in the Distressed cluster are not as poorly situated as the Extremely distressed cluster, but they are clearly troubled. Although central cities in the Distressed cluster are present in all Census divisions with the exception of the Mountain Division, the cluster is disproportionately located in Census divisions along the Eastern seaboard: 35 per cent of the cluster is located in the Middle Atlantic states and 21 per cent in the South Atlantic division. Included in this group are some of the nation's largest central cities: Atlanta, Baltimore, Boston, Chicago, Miami, New Orleans, New York, Philadelphia, Pittsburgh, San Francisco and Washington.¹⁶ These central cities are characteristically somewhat larger than their Extremely distressed brethren, but they are from similar-sized metropolitan areas (Table 9). However, the Distressed cluster does contain smaller places, such as Memphis, New Brunswick, and Richmond.

The central cities in the Distressed cluster have industrial economic bases, but they are balanced by higher levels of symbolic analytical positions than is typical of the other clusters in the group of Stressed central cities. They also have significantly larger proportions of residents with bachelor's degrees than do either the Extremely distressed cen-

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

tral cities or the other three clusters in the group of Stressed central cities. They have lower levels of poverty, and a higher proportion of their adult populations have education beyond the secondary level when compared to the central cities in the Extremely distressed cluster. Compared to all central cities, the Distressed cluster suffered the second-highest rate of population loss, the second lowest ratio of central city to metropolitan median family income and the fourth highest poverty rates. One difference between the Extremely distressed and Distressed clusters is that the MSAs that contain the Extremely Distressed clusters lost population at a much faster rate than did the universe; this was not true for the cluster of Distressed central cities. The set of stereotypical central cities has a much larger share of African-Americans in its population than does the universe of central cities.¹⁷

Manufacturing central cities. The three clusters in the set of Manufacturing central cities are distinguished from all the other clusters by their economic function: they all have large shares of their employment base in manufacturing industries. The percentage of manufacturing jobs for the median central city in each of these clusters is in excess of 20 per cent. What distinguishes these three clusters from each other is the social outcome of that specialisation. The cluster of Larger manufacturing cities acts as a bridge in the continuum of central cities between the clusters of Distressed central cities, sharing some social outcome characteristics of this cluster with the economic base—or function—of the smaller manufacturing cities. Allentown, Brockton, Dearborn and Los Angeles are examples of central cities in this cluster.

The two clusters of Smaller manufacturing central cities are distinguished from each other by the significantly lower poverty and unemployment rates in the More successful cluster, compared to the Less successful cluster. The Less successful cluster of small manufacturing central cities is dominated by cities in the East North Central Census div-

ision—primarily Ohio’s smaller central cities—but such central cities are also present in the Middle Atlantic and East South Central Census divisions. The More successful small manufacturing central cities are present in all Census divisions, with the exception of the West North Central Division. About one-quarter of the More successful manufacturing central cities is located in the East North Central Division; another quarter is located in the South Atlantic Division. What distinguishes the places in this cluster from other central cities in the group of Stressed central cities is the competitive position of the goods they produce.

5.2 *Healthy Central Cities*

The Healthy central city group is made up of three sets of clusters: Competitive central cities, Sunbelt central cities and Knowledge central cities. These three sets of central cities share structural differences with the Stressed group. None has a positive association with the low human capital discriminant function, and three share negative correlations with the decline discriminant function—meaning that these clusters are negatively correlated with declining populations for both the central cities and their MSAs and that a relatively high proportion of their residents are educated beyond the secondary level.

Competitive central cities. Nearly 30 per cent of all central cities in the US are assigned to the three clusters that form the set of Competitive central cities. We call the Heartland cluster the average cluster of central cities, because it is not significantly associated with any of the discriminant functions—meaning that there is no statistical difference between the value of the discriminant function for the cluster and for the universe. There is, however, a regional pattern in the location of these central cities. The second cluster is composed of 59 central cities that we call the Administrative cluster, due to their economic function and size. The third cluster is an anomalous collection of seven central cities

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

(four in the same MSA) on the south Texas border with Mexico; these MSAs are bisected by the US–Mexican border. They are rapidly growing but awash in poverty. All are idiosyncratic metropolitan areas.

The median central city of the 83 in the Heartland group is Glens Falls, New York; it is bracketed by Pueblo, Colorado, and Carlisle, Pennsylvania. The Heartland cluster is concentrated in a block of Census divisions that is roughly L-shaped, beginning in the West North Central Census Division and dropping into the southern tier of states, then moving east to reach the South Atlantic Division. This is a particularly difficult cluster of central cities to describe, because they are the average of the universe; none of the discriminant functions is significantly associated with this cluster. However, it is notable that, when the characteristics of the median central cities of the clusters are compared in Table 7, the Heartland cluster has the third-smallest median metropolitan population size; only the two university clusters are smaller.

The Administrative cluster is negatively associated with the metropolitan interdependence discriminant function—implying that it is negatively correlated with the percentage of its population that out-commutes and positively correlated with the percentage of the MSA population that lives in the central city. It has the highest proportion of central city residents of all the clusters. These central cities, therefore, tend to dominate their metropolitan areas, being significantly larger than the median of the universe, even though their MSAs are not significantly different in population size from those of the universe. A larger share of their workforce holds bachelor's degrees (they trail only the university city clusters in this regard). The central cities in the Administrative cluster are significantly different from the cities of the Heartland cluster in nearly all the variables listed in Table 9 (the only exception being the size of the MSA). We called this cluster of 59 central cities the Administrative cluster because these cities are either the political capitals of their states—such as Albuquerque, Austin,

Columbus (in Ohio), Little Rock, Raleigh and Springfield (in Illinois)—or the economic fulcrum of a region with significant employment in skilled service occupations; this is true of the median central city in this cluster, Lexington, Kentucky, as well as the two cities that bracket it, Charlotte and Greeley, Colorado. (The discriminant analysis indicates that Greeley should be classified in the State university cluster.)

The Heartland and Administrative clusters form the set of Competitive central cities, because they are at the competitive core of the US economy. They are rich in human capital and appear to specialise in advanced service production. They are representative of the nation geographically, with the exception of the north-east, and their economic bases do not depend on institutions that are difficult to replicate—such as major state research universities. These 142 central cities are at the competitive core of the economy.

The sunbelt clusters. There are two distinct clusters of central cities in the Sunbelt. The two most distinctive distinguishing characteristics between the two sets of Sunbelt central cities is the size of the MSA—the Larger cluster is from larger MSAs—and the fact that, although 60 per cent of the central cities in the Smaller cluster are the primary central city in the MSA, only 2 of the 25 central cities in the Larger MSA cluster are the primary central city. However, there is no statistically significant difference in the size of the central cities in these two clusters. Also, the central cities of the Larger Sunbelt cluster are mostly located in the Pacific, Mountain, West North Central and South Atlantic Census Divisions, whereas the Smaller Sunbelt cluster is dominated by smaller metropolitan areas in California and Florida. Upon inspection, it appears that central cities in the Larger Sunbelt cluster are in more urbanised, less isolated parts of the Sunbelt than are those in the Smaller Sunbelt cluster of MSAs. Typical of cities in the Larger Sunbelt cluster are Arlington (in the Fort Worth MSA) and Irving (in the Dallas MSA), as well as the median city in this

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

cluster, Escondido (in the San Diego MSA), and its brackets, Petaluma (in the Santa Rosa California MSA) and Fairfield (in Vallejo, California's MSA). All but five of the primary central cities in these MSAs are classified in the Healthy group of central cities.

The central cities in the Smaller Sunbelt cluster are predominantly located in the Pacific and South Atlantic Census Divisions. These central cities are the fastest growing central cities in the universe and are located in the fastest growing metropolitan areas. Both of these clusters are negatively correlated with the decline discriminant function; in other words, the populations of both the central cities and their MSAs are growing. The Larger Sunbelt cluster is negatively associated with the poverty discriminant function, and the Smaller Sunbelt cluster is negatively associated with the industrial city function.

The poverty rate for the Smaller Sunbelt cluster is significantly lower than for the universe of central cities, but it is twice as high as for the Larger Sunbelt cluster. The significant difference in poverty and unemployment rates between the Smaller and Larger Sunbelt clusters is most likely to be due to two factors: differences in the human capital of their adult residents and the fact that central cities in the Larger Sunbelt cluster are secondary central cities—allowing them to leave the poor to the care of their larger neighbours—whereas the central cities in the Smaller Sunbelt cluster are primary central cities.

The knowledge clusters. Three clusters of central cities form the set of clusters we call Knowledge central cities. The first cluster is a group of 11 Edge cities. This cluster is complemented by two clusters whose economies are dominated by large, prominent universities. The State university cluster is composed of 30 central cities, and the Research university cluster contains an additional 10 central cities. There is a major distinction between the Edge city cluster and the other two in this set. Even though there is

no statistical difference in the size of these three clusters of central cities, the central cities in the Edge city cluster tend to be secondary central cities and part of much larger metropolitan areas than are the central cities in the other two clusters. (The central cities in the two university clusters are most often dominant central cities within their metropolitan areas—or one of two equally dominant central cities, where both are members of one of the university clusters.)

Differences in the roles these three types of central cities play in their metropolitan areas are driven home when the median central cities in each cluster are displayed in Figure 2. The median Edge city is Middletown, Connecticut (in Hartford's MSA). It is bracketed by White Plains, a central city in New York's MSA, and Midland, Michigan, the third central city in Saginaw's MSA. The median central cities in the State University cluster are the cities of Bryan, Texas, and Fayetteville, Arkansas. They are bracketed by Norman, Oklahoma, and Eugene, Oregon. Of the 10 Research university central cities, only 3 are not the primary central city, or one of a pair of twin central cities, in their MSAs: Chapel Hill (in the Raleigh MSA) Davis (in Sacramento's MSA), and East Lansing (in Lansing's MSA).

The Edge city cluster shares many of the characteristics of the Larger Sunbelt central city cluster; together, they bridge the Sunbelt and Knowledge sets of clusters. The Edge city cluster differs from the cluster of central cities from Larger Sunbelt MSAs in that the Larger Sunbelt cluster is significantly negatively correlated with the decline discriminant function, whereas there is no statistically significant correlation between the Edge city cluster and that function. The Edge city cluster is significantly different from the Larger Sunbelt cluster in that Edge cities and their MSAs are slower growing, and statistically, a much higher proportion of Edge city residents are employed in symbolic analytical positions and have terminated their education with four-year college degrees.

The Edge city cluster has a significant positive association with the metropolitan in-

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

terdependency discriminant function (meaning that a high fraction of residents out-commute and that the population is a small percentage of MSA population) as well as the tight labour market discriminant function. The Edge city cluster is also negatively correlated with the high poverty and small employment base discriminant function. The Edge city cluster and the Research university cluster share significant negative correlations with the poverty discriminant function and positive correlations with the metropolitan interdependency discriminant function.

The two University clusters share significant negative correlations with the low human capital discriminant function, making them the polar opposite of the Extremely distressed and Manufacturing central city clusters. The Research university cluster of central cities also has a strong positive correlation with the decline discriminant function—meaning that they are not in decline—and a negative correlation with the high poverty discriminant function. It is clear that the strength of these central cities lies in their rich human capital base; this is evident not only in the discriminant functions but in the pseudo *t*-tests displayed in Table 8. Both University clusters have significantly greater shares of their adult population with bachelor's degrees than do any of the other clusters, as well as a larger fraction of their adult population in symbolic analytical positions. The central cities and the MSAs of both University clusters are growing faster than the universe as a whole and faster than the Edge city cluster, but there is no significant difference in the growth rates between the two University clusters.

What is interesting to note is that both these clusters have unemployment rates significantly lower than the median for the universe of central cities, but poverty rates significantly higher than the median for the universe; poverty rates for these two clusters are the highest of any clusters in the Healthy group of central cities. The poverty rate in the Research university cluster is significantly higher than for the State university cluster.

6. Conclusion

We have shown how central cities, as defined by the Census Bureau, can be classified into separate groups of similar cities. We accomplish this by combining existing methods to build meaningful categories of activities and to identify statistical differences among those categories. We first use hierarchical cluster analysis to construct categories of central cities. We then develop a method for identifying the cluster solutions that are candidates for the optimal solution; in this article, we use three 'candidate' solutions. Our next challenge is to discover the meaningful differences, if any, among the clusters of the lowest-order candidate solution. We employ discriminant analysis—first to test the internal validity to the cluster solution, using the cluster groupings as the prior, and then to differentiate among the clusters statistically. The latter use of the discriminant functions is the more important, because we use the functions as a basis for identifying variables that highlight differences among the clusters.

We order the cluster groupings, using information from the cluster analysis and two pieces of information from the discriminant analysis. We first group clusters, based on the way they were grouped by the cluster analysis as the clustering proceeded. Then we order the clusters within these groups by their discriminant scores from the discriminant analysis of the lowest-order candidate cluster solution. We confirm this ordering by comparing the group assignments from the cluster analysis and discriminant analysis as well as looking at the pattern of overlaps between the cluster group assignment and the suggested assignment from the discriminant analysis. We use pseudo *t*-tests of each variable to determine whether the median of that variable for a particular cluster is significantly different from the median for the universe of central cities. We also use pseudo *t*-tests to determine whether the median of a variable for a particular cluster is significantly different from the median of the same variable for its neighbouring cluster.

We illustrate how this methodology can be

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

used by employing a variety of variables in the cluster and discriminant analyses that measure city functions, social outcomes of interest to public policy and the spatial structure of the metropolitan areas. We find that the universe of municipalities labelled 'central cities' is not homogeneous. There are two distinct groups of central cities in the US—one stressed and the other healthy. In turn, these two large groups can be meaningfully disaggregated into 13 clusters of central cities.

The general method we have employed in this article can be used for any case in which developing taxonomies is important, and where the groupings—and the distinguishing characteristics of the groups—are not known *ex ante*. This is a frequent case in public policy research, where policy attempts to target different groups within populations. As an example, Hill and Brennan are working with a research team that is using this technique to identify industrial clusters in north-east Ohio. Wolman is using this research to provide contextual categories for his work on the voting behaviour of Congressional representatives on bills that are important components of urban public policy.

'Central city' is a statistical construct, based upon the function of municipalities as places of work and residence, whereas the popular and political image of central cities emphasises social outcomes that are prevalent in only 54 of America's 508 central cities. There are lessons in our work for the worlds of both urban research and public policy. The lesson for researchers and statistical modellers is about specification error. Simply grabbing data labelled 'central city' and using them as proxy variables for impoverished, socially challenged localities is a mistake. Doing so introduces a great deal of non-random spatial variation into the construct. Thought should be given to the actual universe of central cities that researchers are examining. One potentially productive approach might be to classify cities according to function and then examine whether a city's conditions, or the economic well-being

of its residents, differs across functional classifications.

The lesson for policy is similar. Central cities are a diverse group of places, both socially and functionally, and using the category 'central city' as a form of policy shorthand is wrongheaded. This work demonstrates that, at a minimum, the 54 central cities in the Stereotypical set should receive special consideration under federal urban policy for financial and developmental assistance; these are a very different group of large municipalities from other US central cities.

Notes

1. Cities have always been evocative, whether in politics or in literature. Three works, two of them anthologies, explore these images. Hauser and Schnore's classic 1965 collection of essays contains chapters that survey urbanism in the each of the social sciences. Hauser's overview chapter, Sjoberg's review of urban sociology and Hauser's essay on the urban-folk and urban-rural dichotomies are especially interesting. Lloyd Rodwin and Robert Hollister edited a collection in 1984 that was developed from a seminar they taught on the image of cities in the social sciences. This collection serves as a good follow-on to Hauser and Schnore. Finally, one should consider Kevin Lynch's classic book, *The Image of the City* (1960), and his reflection on that work (Lynch, 1984) published in Rodwin and Hollister.
2. By starting with the Bureau of the Census' definition of central cities, we will not be able to consider inner-ring suburbs with unusually large fractions of their populations living in poverty and small employment bases, such as East Cleveland, Ohio, or Highland Park, Michigan.
3. Coulton and her co-authors (Coulton *et al.*, 1996) use a similar method in their classification of 100 US metropolitan areas based on spatial concentrations of affluence and poverty in 1990. They used a non-hierarchical form of cluster analysis (the *K*-means algorithm) and found five distinct clusters, ranging from a group of 9 MSAs with high spatial concentrations of both affluence and poverty to a cluster of 28 MSAs that had low concentrations of affluence and poverty. They then used discriminant analysis to determine the socioeconomic differences among the clusters.

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

4. The method described here is known as the Ward method. A number of sources, such as Aldenderfer and Blashfield (1984), Hair *et al.* (1987), Everitt (1993) and Gittleman and Howell (1995) highlight the overall efficacy of this method compared to other clustering methods.
5. The hit ratio is calculated as part of the discriminant analysis. Because running separate discriminant analyses for each possible cluster solution is extremely time consuming, we conducted the discriminant analysis only for the three candidate solutions and listed the three hit ratios with the cluster maps displayed in Figure 1.
6. Because we were testing to see if the means were not statistically equivalent, wide tolerance levels in the *t*-test provide the strictest test.
7. SPSS's cluster analysis produces an *F*-test for the difference between each pair of clusters. The null hypothesis is that each pair is the same. The hypothesis was rejected for all pairs at the 0.01 level of significance.
8. The mean discriminant scores tracked with the median scores for all of the clusters. We report the medians in Table 3 to be consistent with the way we report other data in the article. The means of each group are given in parentheses in the following list.
 - All central cities (−0.040).
 - The two groups: Stressed (0.637) and Healthy (−0.503).
 - The five sets: Stereotypical (0.928), Manufacturing (0.544), Competitive (−0.081), Sunbelt (−1.120) and Knowledge (−0.760).
 - The 13 clusters: Extremely distressed (1.551), Distressed (0.769), Larger manufacturing (0.860), Smaller less successful manufacturing (0.689), Smaller more successful manufacturing (0.30), Heartland (0.157), Administrative (−0.503), Border Texas (0.656), Smaller Sunbelt (−1.080), Larger Sunbelt (−1.208), Edge (0.039), State university (−0.928) and Research university (−1.605).
9. Coulton *et al.* (1996) pp. 207–214 make a similar point concerning the clusters of MSAs grouped according to their degree of spatial concentrations of poverty and affluence.
10. We use the information on the overlaps to order the clusters from the 13-cluster solution to produce Table 3, giving us a second indication of the proper array of the clusters of central cities (the cluster map in Figure 1 is the first). We began by using the cluster hierarchy to array the cluster groups of central cities. First, we divided the central cities into the stressed and healthy categories, based on the 2-cluster solution. We then had to decide how to order the 5-cluster solution—the 2 sets in the stressed city category and the 3 in the healthy city group. We based this on the overlaps between the clusters in the 13-cluster solution. The first cluster of the 13-cluster solution, called Extremely distressed central cities, has no overlaps, and it anchors the left side of the table. This cluster also has the highest median discriminant score, which indicates that it is the most distressed. The right side of the table is anchored by the Research university cluster of central cities; that cluster has just one overlap—with the State university cluster—and it has the lowest median discriminant score, indicating that it is the healthiest cluster. We then located the cluster with the largest number of overlaps—the Heartland central city cluster—and placed it in the centre of the array. Once these three clusters were placed in Table 3, the order of the others was dictated by how the 13 clusters related to each other in the 5-cluster solution. There is one exception to this neat pattern of relationships, and this is the South (Border) Texas cluster. It is associated with the Heartland and Administrative clusters in the 5-cluster solution. However, this Texas cluster is an outlier in the analysis. The Border Texas cluster is made up of seven central cities where the Mexican border effectively bisects the metropolitan areas. These MSAs could be ignored without affecting the analysis but are reported for the sake of completeness.
11. This is equivalent to having a regression equation with a significant *F*-statistic and no significant *t*-statistics. This means that although the overall equation has significant but limited explanatory power, it cannot be determined whether any particular independent variable is closely associated with the dependent variable.
12. As we mentioned above, this is best thought of as being equivalent to a *t*-test of the relationship between independent and dependent variables in a regression equation.
13. Two lists of central cities are available, either from the authors or from the website of The Urban Center of the Levin College of Urban Affairs at Cleveland State University (www.urban.csuohio.edu). One contains a list of each of the clusters from the 13-cluster solution, with their constituent central cities. This table also shows the suggested cluster assignment from the discriminant analysis. The second list is an alphabetical list of

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

central cities, giving their cluster and discriminant group assignments from the 13-cluster solution.

14. The pseudo t -test is analogous to the t -test of the difference of two means (where the subscript i represents each sample). The median (M_i) is used instead of the mean, and the pseudo-standard deviation (PSD_i) is used instead of the standard deviation. The PSD_i is the inter-quartile range divided by 1.35. N_i is the number of cases in each sample. SQRT means square root. The formula we used is:

$$\frac{(M_1 - M_2) / \text{SQRT}[\frac{((N_1 - 1)PSD_1^2 + ((N_2 - 1)PSD_2^2))}{(N_1 + N_2 - 2)}]}{(1/N_1 + 1/N_2)^{1/2}}$$

This test is appropriate when the distributions are highly skewed.

15. Refer to Note 10, where we describe the way the clusters are ordered.
16. Some readers of earlier drafts of this article have examined the cluster placement of cities with which they are familiar and have disputed their placement—saying, for example, that Boston, New Brunswick, or Pittsburgh are not distressed cities and should not be labelled as such. This is an important point and we have several responses. First, the data are from the 1990 Census and it is possible, but not likely, that economic conditions have improved to the point that if more current data were available a new cluster analysis would change their assignment. Secondly, as we demonstrated in Wolman *et al.*, (1994), there is often a vast difference between perceptions, economic development marketing and the world as it is measured statistically. These central cities cluster together because their residents share common social outcomes and the cities share a common set of economic functions; not because of the architectural quality of their central cities. Thirdly, as we acknowledge when discussing differences in the placement of central cities in the cluster and discriminant analyses, all statistical processes contain errors, be it regression, cluster or discriminant analysis. Some may come from specification error and some from measurement error. However, in any typological assignment there is a third source of error, and it is probably the most important source. This is interpretive, or perceptual, error on the part of the analyst or reader. All typologies need to have face validity, but they must also be rooted on a firm and logical set of decision, or assignment rules, so that perceptions, or other sources of bias, do not corrupt the typology. Therefore, while it is important

that the outcomes have face validity, it is more important that the assignment rules have face validity.

17. There were seven central cities that the cluster analysis assigned to the Distressed cluster but that the discriminant analysis indicated belonged in other clusters. It is instructive to look at these places, because they show the multi-dimensional continuum these central cities share. The discriminant analysis indicated that Athens, Georgia, home of the University of Georgia, shares the characteristics of cities in the State university cluster; it placed Honolulu and Seattle with the Administrative central city cluster; it suggested that Providence has more in common with the Larger manufacturing cluster of central cities; and Dayton more in keeping with the Smaller less successful cluster of manufacturing central cities.

References

- ALDENDERFER, M. and BLASHFIELD, R. (1984) *Cluster Analysis*. Beverly Hills: Sage Publications.
- BERRY, B. (1996) Technology-sensitive urban typology, *Urban Geography*, 17, pp. 674–689.
- BERRY, B. and SMITH, K. (1972) *City Classification Handbook*. New York: Wiley-Interscience.
- BRADBURY, K., DOWNS, A. and SMALL, K. (1982) *Urban Decline and the Future of American Cities*. Washington, DC: Brookings Institution.
- COULTON, C., CHOW, J., WANG, E. and SU, M. (1996) Geographic concentration of affluence and poverty in 100 metropolitan areas, 1990, *Urban Affairs Review* 32(4), pp. 186–216.
- DAHMAN, D. and FITZSIMMONS, J. (1995) Introduction, in: D. DAHMAN and J. FITZSIMMONS (Ed.) *Metropolitan and Nonmetropolitan Areas: New Approaches to Geographical Definition*, pp. 1–8. Working Paper No. 12, US Bureau of the Census, Washington, DC.
- EVERITT, B. S. (1993) *Cluster Analysis*, 3rd Edn. New York: Halsted Press.
- FORESTALL, R. (1967) Economic classification of places over 10 000, 1960–1963, in: *The Municipal Yearbook 1967*, pp. 30–65. Chicago: International City Managers' Association.
- GITTLEMAN, M. and HOWELL, D. (1995) Changes in the structure and quality of jobs in the United States: Effects by race and gender, 1973–1990, *Industrial and Labor Relations Review*, 48, pp. 420–440.
- HAIR, J., ANDERSON, R. and TATHAM, R. (1987) *Multivariate Data Analysis*, 2nd Edn. New York: Macmillan Publishing.
- HARRIS, C. (1943) A functional classification of

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

- cities in the United States, *Geographical Review*, 33, pp. 86–99.
- HART, J. (1955) Functions and occupational structures of cities in the American south, *Annals of the Association of American Geographers*, 45, pp. 269–286.
- HAUSER, P. (1965a) Urbanization: an overview, in: P. HAUSER and L. SCHNORE (Eds) *The Study of Urbanization*, pp. 1–48. New York: John Wiley & Sons, Inc.
- HAUSER, P. (1965b) Observations on the folk–urban and urban–rural dichotomies as forms of western ethnocentrism, in: P. HAUSER and L. SCHNORE (Eds) *The Study of Urbanization*, pp. 503–518. New York: John Wiley & Sons, Inc.
- HAUSER, P. and SCHNORE, L. (Ed.) (1965) *The Study of Urbanization*. New York: John Wiley & Sons, Inc.
- HILL, E. and BIER, T. (1989) Economic restructuring: earnings, occupations and housing values in Cleveland, *Economic Development Quarterly*, 3, pp. 123–144.
- HILL, E. and WOLMAN, H. (1997a) Accounting for changes in income disparities between US central cities and their suburbs from 1980 to 1990, *Urban Studies*, 34, pp. 43–60.
- HILL, E. and WOLMAN, H. (1997b) City–suburban income disparities and metropolitan employment: can tightening labor markets reduce the gaps?, *Urban Affairs Review*, 32, pp. 558–582.
- JAMES, F. (1990) City needs and distress in the United States: 1970s to the mid 1980s, in M. KAPLAN and F. JAMES (Eds) *The Future of National Urban Policy*, pp. 13–31. Durham, NC: Duke University Press.
- JONES, V. and FORESTALL, R. (1963) Economic and social classification of metropolitan areas, in: *The Municipal Yearbook 1963*, pp. 31–44. Chicago: International City Managers' Association.
- KEELER, E. and RODGERS, W. (1973) *A Classification of Large American Urban Areas*. Santa Monica, CA: The Rand Corporation.
- KLECKA, W. R. (1980) *Discriminant Analysis*. Beverly Hills and London: Sage Publications.
- LADD, H. and YINGER, J. (1989) *America's Ailing Cities*. Baltimore: Johns Hopkins University Press.
- LYNCH, K. (1960) *The Image of the City*. Cambridge, MA: MIT Press.
- LYNCH, K. (1984) Reconsidering the image of the city, in: L. RODWIN and R. HOLLISTER (Eds) *Cities of the Mind*, pp. 151–162. New York: Plenum Press.
- NELSON, H. (1955) A service classification of American cities, *Economic Geography*, 31(3), pp. 189–210.
- NOYELLE, T. J. and STANBACK, T. M. (1983) *The Economic Transformation of American Cities*. Totowa, NJ: Rowan & Allanheld.
- OTTENSMANN, J. (1996) The new central cities: implications of the new definition of the metropolitan area, *Urban Affairs Review*, 31, pp. 681–691.
- RODWIN, L. and HOLLISTER, R. (Eds) (1984) *Cities of the Mind*. New York: Plenum Press.
- RUSK, D. (1993) *Cities Without Suburbs*. Washington, DC: Woodrow Wilson Center Press.
- SJOBERG, G. (1965) Theory and research in urban sociology, in: P. HAUSER and L. SCHNORE (Eds) *The Study of Urbanization*, pp. 157–190. New York: John Wiley & Sons, Inc.
- TEITZ, M. (1989) Neighborhood economics: local communities and regional labor markets, *Economic Development Quarterly*, 3, 111–122.
- US BUREAU OF THE CENSUS (1947) *County Data Book*. Washington, DC: US Government Printing Office.
- US BUREAU OF THE CENSUS (1952) *County and City Data Book, 1949*. Washington, DC: US Government Printing Office.
- US BUREAU OF THE CENSUS (1953) *County and City Data Book, 1952*. Washington, DC: US Government Printing Office.
- US BUREAU OF THE CENSUS (1957) *County and City Data Book, 1956*. Washington, DC: US Government Printing Office.
- US BUREAU OF THE CENSUS (1962) *County and City Data Book, 1962*. Washington, DC: US Government Printing Office.
- US BUREAU OF THE CENSUS (1973) *County and City Data Book, 1972*. Washington, DC: US Government Printing Office.
- US BUREAU OF THE CENSUS (1978) *County and City Data Book, 1977*. Washington, DC: US Government Printing Office.
- US BUREAU OF THE CENSUS (1991) *State and Metropolitan Area Data Book*. Washington, DC: US Government Printing Office.
- VARADY, D. and LIPMAN, B. (1994) What are renters really like? Results from a national survey, *Housing Policy Debate*, 5, pp. 491–531.
- WOLMAN, H., FORD, C. and HILL, E. (1994) Evaluating the success of urban success stories, *Urban Studies*, 31, pp. 835–850.

Appendix. Measuring Central Cities in the US

The operationalisation of the concept of 'central city' first appeared in US statistical sources in 1910, when the Bureau of the Census designated municipalities with more than 200 000 residents as the core of metropolitan districts (US Bureau of the Census, 1978, p. xvii). The definition of metropolitan districts was expanded in the 1940 Census to cover an area with "central city or cities having populations of 50 000 or more" and adjacent

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

ent places having population densities of 150 or more per square mile (US Bureau of the Census, 1947, p. v). In 1947, the Bureau of the Census used a county-based definition of a metropolitan area for the first time and reported these statistics in the *County Data Book*. At that time, a county was included in a metropolitan area if at least half its population was within the municipal limits of a metropolitan district in 1940. The operational definition of central cities evolved again in 1949, with the establishment of standard metropolitan areas (SMAs) by the Bureau of the Budget in preparation for the 1950 Census of Population. The criteria recognised that metropolitan areas may have more than one central city, stating that “Where two cities of 50 000 or over are within 20 miles of each other, they will ordinarily be included in the same area” (US Bureau of the Census, 1952, p. ii). (See Dahmann and Fitzsimmons (1995) for a list of dates of when the definition of metropolitan area changed in the US. The *County and City Data Book*, 1949 indicated that the Bureau of the Budget defined standard metropolitan areas in January 1949 (US Bureau of the Census, 1952, p. ii). The evolution of the concept of metropolitan areas is traced in US Bureau of the Census (1973, p. xxi; and 1978, pp. xvii–xviii). The *County and City Data Book*, 1949 noted that the

‘general concept adopted for the determination of a standard metropolitan area was that each area should represent an integrated economic unit with a large volume of daily travel and communication between central city and the outlying parts of the area’ (US Bureau of the Census, 1952, p. ii).

The *County and City Data Book*, 1952 was the first to define formally a central city, in keeping with its current usage. The Bureau of the Census recognised that not all cities with 50 000 population were necessarily central cities. The largest city in the central county was designated the primary central city, and any other municipality with at least 25 000 population that was at least one-third the size of the primary central city was also designated a central city—provided that no more than three central cities were designated for any standard metropolitan statistical area (SMSA). The title of the SMSA was to include all the central cities, with the exception of the New York-north-eastern New Jersey SMSA, where Jersey City and Newark were not made part of the SMSA name (US Bureau of the Census, 1953, p. XI).

The rules governing the title of a metropolitan area also regulated the maximum number of central cities recognised in any metropolitan area. Limiting the maximum number of central cities to

three remained a feature of US statistical policy until the revisions of 1983. In 1956, the naming rules were amended to name all cities of 250 000 or more as central cities, even if they were less than one-third the size of the largest city in the SMSA, provided that no more than three municipalities were defined as central cities of that metropolitan area (US Bureau of the Census, 1957, p. XI). These rules were only slightly altered in 1962. If the cities were less than 20 miles apart, measured border-to-border, and in different counties, they were considered central cities of the same SMSA unless “there is definite evidence that the two cities are not economically and socially integrated” (US Bureau of the Census, 1962, p. XI).

The operational definition of central cities changed drastically in 1983 (Ottensmann, 1996). Before that year, each metropolitan area was restricted to a maximum of three municipalities labelled as central cities. The largest municipality in an MSA or primary metropolitan statistical area (PMSA) is usually defined as a central city. Additional cities may be designated central cities in MSAs or PMSAs if they: have populations greater than, or equal to, 250 000 and contain at least 100 000 workers; or have at least 25 000 residents, the employment to employed resident ratio is greater than or equal to 0.75 (the ratio of the number of people employed in the municipality to the number of employed people living in the municipality) and less than 60 per cent of the employed residents out-commute; or are municipalities of 15 000 to 25 000 residents at least one-third as large as the largest central city, have an employment to employed resident ratio of at least 0.75, and less than 60 per cent of the employed residents out-commute. All municipalities that meet these criteria in a metropolitan area are classified as ‘central cities.’ (Several central cities do not strictly meet these criteria but are still labelled ‘central cities’, due to the US Congress passing legislation that created some metropolitan areas. The largest municipalities of these ‘instant’ MSAs automatically became ‘central cities’.) The historical limit of a maximum of three central cities within a metropolitan area, as reflected in the title of the metropolitan area, is gone. (The 1983 change in definition calls into question time-series data on aggregated central cities that bridge this change in definition without making adjustments to the data. Ottensmann (1996, p. 683) noted that, when this definition of central city came into effect in 1983, the number of central cities in the 1980 and 1990 Census of Population changed. The 1980 standard allowed a maximum of three municipalities to be labelled ‘central cities’. The largest city in a metropolitan area was always defined as a central city (with the exception of the Nassau–Suffolk metropolitan area in

WHAT IS A CENTRAL CITY IN THE UNITED STATES?

New York state), and up to two other municipalities could be labelled central cities if their population was at least one third that of the largest municipality. Ottensmann found that the new

definition resulted in 107 cities being defined as central cities in 1990 that were not part of the list in 1980, and that 21 municipalities lost their central city designations between 1980 and 1990.)