Quantifying and Stratifying the Spatial Patterns of Residential Clusters: A Socioeconomic and Geographic Comparisons of Metropolitan Areas

by

Myoung-Gu Kang

Master of Science in Urban Engineering
Seoul National University, Seoul, Korea, 1995

Bachelor of Science in Urban Engineering
Seoul National University, Seoul, Korea, 1993

Submitted to the Department of Urban Studies and Planning
in Partial Fulfillment of the Requirements for the Degree of

Master in City Planning
at the
Massachusetts Institute of Technology

June 2001

Signature of Author...
Department of Urban Studies and Planning
May 17, 2001

Certified by...
Joseph Ferreira Jr.
Professor of Urban Planning and Operations Research
Thesis Supervisor

Accepted by...
Dennis Frenchman
Professor of Urban Studies and Planning
Chair, MCP Committee

Quantifying and Stratifying the Spatial Patterns of Residential Clusters: A Socioeconomic and Geographic Comparisons of Metropolitan Areas

by

Myoung-Gu Kang

Submitted to the Department of Urban Studies and Planning
on May 17, 2001 in Partial Fulfillment of the Requirements
for the Degree of Master in City Planning

ABSTRACT

What activities are located where and why? This is a fundamental question in urban studies, which should be answered before planning. In urban economics, fine mathematical models have been developed and have provided an internally consistent economic framework for answering these questions. However, as cities grow and get larger, the spatial structure of cities has changed and become increasingly complicated. The emergence of sub-centers and socioeconomically distinct clusters within the metropolitan areas weaken the assumptions on which the urban economic models are based. Also, there are increasing needs of the people and space interaction models at the micro level, or neighborhood level. This thesis revisits this fundamental issue in a different way. In particular, it investigates the spatial patterns of residents within the metropolitan areas at the census blockgroup level, using Database Management Systems (DBMS) with Structured Query Language (SQL), Geographic Information Systems (GIS), and statistical methods including factor analysis and cluster analysis.

What socioeconomic factors make one type of neighborhood different from another within the metropolitan areas? This thesis finds four common socioeconomic factors; baseline factor, children factor, income factor, and age factor. Baseline factor shows that the major generic whites are, roughly speaking, more likely to be citizens, to speak only English at home, to drive to work, not to be poor, to own homes, and to live in the lower population density areas. These above variables move synchronously, so we can reduce them into one factor which we have abbreviated as the 'baseline factor'. Using the same factor analysis, we developed the four such socioeconomic factors above.

Then, I investigate where they are located? In all four metropolitan areas there are wedges of clustered neighborhoods with similar socioeconomic characteristics around the urban center. Each section contrasts with each other, for example, rich versus poor, or white versus non-white. Younger people formed their own wedges, too. Second, the downtown and subcenters, where the jobs are located, are more likely occupied by non-whites or low income individuals. So, they also appear along the circumferential highway corridor where the subcenters are located.

In addition to the common pattern over all metropolitan areas, each metropolitan area also has its own peculiar characteristics. In the Boston Metropolitan Area, the delineation of socioeconomically different neighborhoods coincides with town boundaries. That is, the characteristics of neighborhoods are discrete rather than continuous over the space. In the Chicago Metropolitan Area, the percentage of citizens is another key factor differentiating neighborhoods, and, hence, non-citizens occupy a separate cluster. The unique geography of the San Francisco Metropolitan Area creates two stark types of neighborhoods; affluent neighborhoods at the west of the bay along the ocean, and poor neighborhoods at the east of the bay, especially at the entering points of the bridges to the downtown. In Dallas Metropolitan Areas, the geographic contrast between rich and poor neighborhoods are clearer, i.e., the northern area is wealthier while the southern area is poorer.

In this thesis, I find the key socioeconomic factors characterizing the neighborhoods and the spatial pattern of residents. Also, I developed a different methodology to look at this issue. This study gives us a foundation for micro level urban simulation modeling by providing a systematic method of quantifying neighborhood characteristics in ways that can be incorporated into economic models. Furthermore, we can analyze the urban structure of diverse land uses over space and time simultaneously. This can make participatory planning far easier by supplying a clear picture of a city's profile, stimulating communications, and facilitating understanding among residents.

Thesis Supervisor: Joseph Ferreira, Jr.

Title: Professor of Urban Planning and Operations Research

ACKNOWLEDGMENTS

I deeply appreciate my thesis advisor, Professor Joseph Ferreira, Jr., for his invaluable guidance throughout the study.

I am also greatly indebted to Anne Kinsella Thompson, Mizuki Kawabata, Michael Shiffer, Qing Shen, Phillip Thompson, Thomas Grayson, and the entire Planing Support Systems Group (PSS) for helping me develop the ideas and providing a heartwarming atmosphere.

I want to say thanks to my friends and faculty in DUSP for helping me shape my ideas and providing emotional support.

Especially, I would like to thank my wife, Mihyun Kim, and my son, Minqu Simon Kang as well as my parents for their love and understanding.

Myoung-Gu Kang
May 17, 2001

Table of Contents

List of Figures

List of Tables

1

INTRODUCTION

We must understand cities as much as possible before planning. Urban economics sheds light on the urban spatial structure and gives us a better understanding about how cities work. Now, thanks to today's high technology and computation capacity – including Geographic Information Systems (GIS), Statistics, and Database Management Systems (DBMS) – and rich datasets like the U.S. Census, we can approach our cities with attention to spatial detail beyond that of traditional abstract and spatially aggregated urban economic models. This bottom-up way of research is essential to develop a micro level urban simulation model for participatory planning. It will help practitioners and residents understand their cities and planning, and hence will enrich the communication between them.

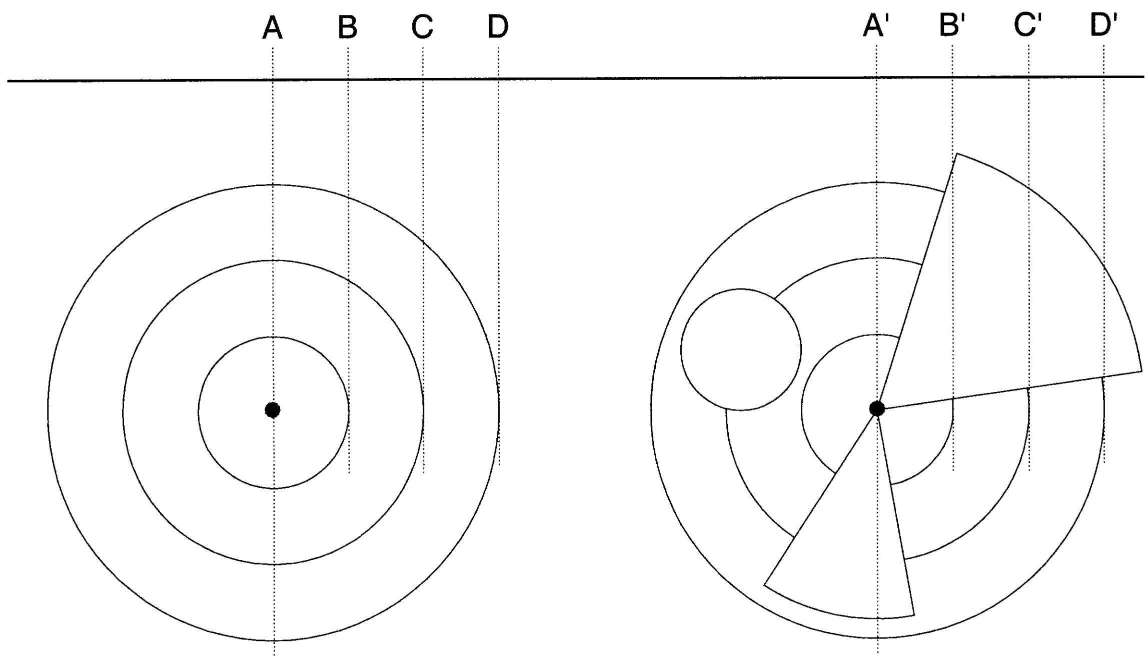
I hypothesize that there are socioeconomically distinct neighborhoods within metropolitan areas: similar neighborhoods generally tend to agglomerate geographically with locational preferences, and unlike the traditional urban economic models suggest, they are less likely to be located in circumferencial distribution.

As cities have grown, urban economists have developed theories to understand the internal structure of cities. With the urban economic models, we learn about where people live, where firms and retail stores are located, how housing and real estate markets work, how we can improve our cities, etc. The theories have given us a concise and clear picture of the internal structure of cities. The simple and powerful models successfully simplify the entangled interactions between people, and between people and spaces. However, most urban economic models are built based on three critical assumptions: one dimensional space, smooth transition over distance, and a predefined city center point.

1.1 Three Issues in Urban Economics with Space

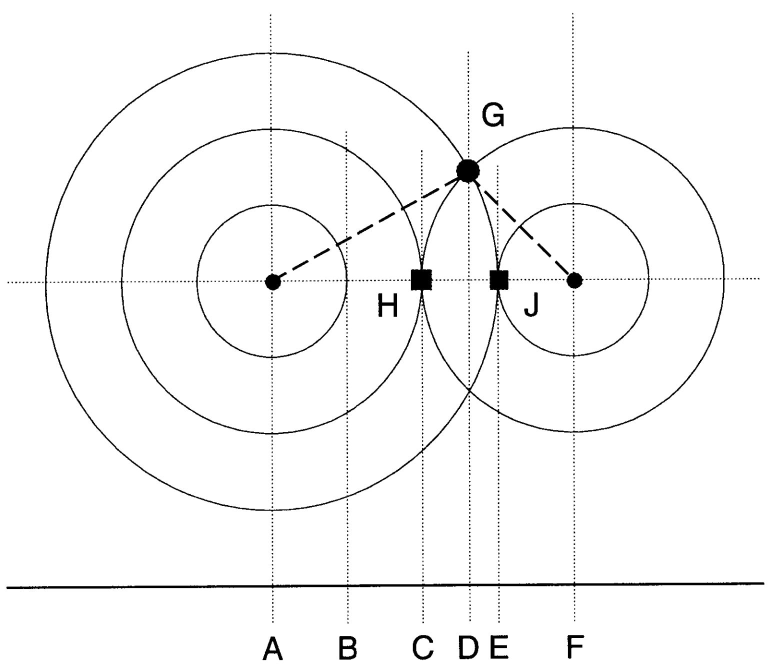
Urban economic models generally, Implicitly or explicitly, assume that space is a "line" and hence, interpret a "circular" city on two dimensional space. That is, the models only allow circular distribution, like a circle or donut, but don't allow "holes" or "wedges". However, we can easily find the "holes" and "wedges" in metropolitan areas. It is hard to explain the non-circular patterns with the traditional linear city models even though the models illustrate the overall structure of cities well with density and land value declining with destance from the ‘center.’ (Figure 1.1)

Figure 1.1 Linear city and two dimensional space



Furthermore, subcenters within metropolitan areas stand out these days, so understanding the role of the subcenters is becoming important for planning and transportation. Controlling these subcenters relates to whether urban growth is sprawl or sound development. We observed the fact that the sizes of cities have grown bigger and more people and industries tend to live away from the traditional city center. In many cases, however, the urban fringes have developed independently. Therefore, new developments are not consistent with other land usages, either new or old. Urban economists take the traditional view and try to explain it with linear city models. The two dimensional distance issues arise here again. In Figure 1.2, point D on the linear city cannot easily be intepreted as point G or even H or J. If we have more subcenters, like today's metropolitan areas, these issues become complicated.

Figure 1.2 Distance issue of linear city with a subcenter



Secondly, we assume that the transition over distance is smooth. For example, as the distance increases, rent goes down smoothly, population density goes down smoothly, and so on. In Figure 1.3, we can see the overall trend of the smoothly decreasing population density away from CBD. However, the flattening slope and the emergence of subcenters requires a different approach beyond traditional urban economics. Again, the subcenters become more important to explain the internal structure of cities.

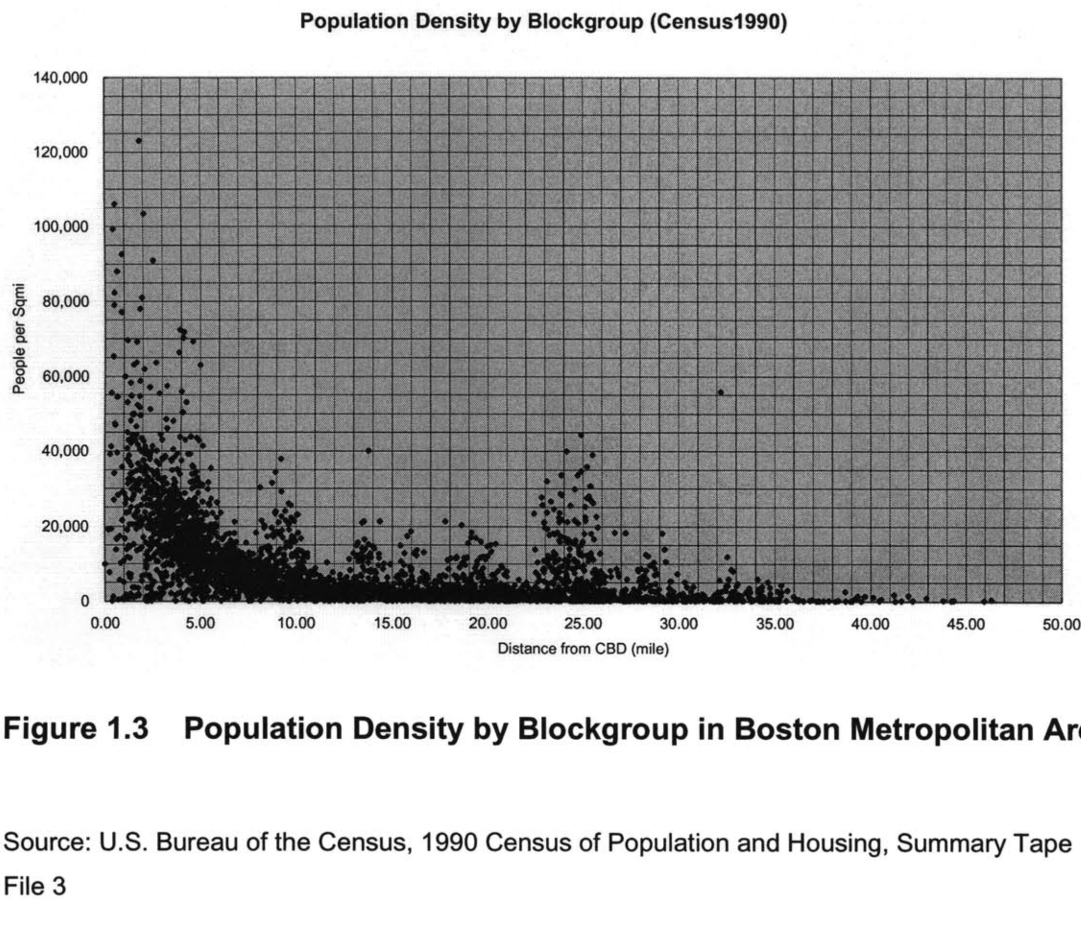


Figure 1.3 Population Density by Blockgroup in Boston Metropolitan Area

Source: U.S. Bureau of the Census, 1990 Census of Population and Housing, Summary Tape File 3

Thirdly, we have traditionally treated the center of a city as a fixed point. The assumption is closely related to the level of geographic aggregation. Considering the city of Boston merely as a point would definitely be a reasonable proxy when we do a national level of study, or beyond. However, for example, when we do research on the residential distribution within the Boston metropolitan area, the city of Boson is too big to be a point, or even an entity. Maps in the following chapters illustrate this point.

Urban economics has done an effective job in shedding light on the internal urban sturucture. Today's high technology and computation capacity – GIS, Statistics, and DBMS – and rich dataset like the U.S. Census, allow us to go further, beyond the current limit of urban economics. Therefore, I take a bottom-up approach to examine the spatial patterns of residential clusters utilizing modern technology.

1.2 Methodology

In this thesis, I am trying to measure the residential pattern of cities as it is on two dimensional space. First of all, I will classify the neighborhoods[1] by their

---

[1] In this thesis, I will use Census blockgroup as a unit of analysis, and call it neighborhood. Census blockgroup has quite good characteristics to be a homogeneous community and is the most detailed and richest data set I can use. (See Appendix for more detail information about Census Blockgroup)

socioeconomic characteristics. What are the effective factors which make them different from each other? How many different clusters do the metropolitan areas have? The combination of factor analysis and cluster analysis is a good tool to answer these questions. Secondly, I pose the following questions: Where are they located? Are similar neighborhoods located closely to each other? I will use GIS and draw maps of the residential distribution patterns for each metropolitan areas. All of the above procedures require intensive use of DBMS.

Conventionally, multiple linear regression analysis is used most frequently in quantitative modeling. However, the attendant problems of nonlinearities and interactions, multicollinearity, functional heterogeneity, and heteroscedasticity can severely degrade the accuracy of the estimates.

Some of these problems can be reduced by stratifying or clustering the data into more homogeneous subgoups, each of which is treated as separate independent data base for the purpose of regression modeling. The cluster analysis can be modified to accomplish such data base segmentations.

We can use factor analysis to adjust the variable interactions. The basic idea of factor analysis is finding few principal *latent factors* to explain complicated phenomena. This statistical method has been widely used in fields where factors cannot be manipulated, such as psychology.

In this research, I first do the factor analysis to extract the few principal *latent factors* out of the eighteen socioeconomic characteristics of residents, which differentiate the neighborhoods within metropolitan areas. I next perform a cluster analysis to stratify the neighborhoods into fewer groups of the socioeconomically similar neighborhoods with the factors extracted from factor analysis. Third, I input the result of statistical analysis into GIS, and draw a map of spatial patterns of residential clusters in order to answer whether or not similar neighborhoods are located near one another.

1.3 Study Areas and Data

In this cross-sectional analysis, I look into four metropolitan areas and compare their characteristics: Boston, Chicago, San Francisco, and Dallas. Each of these areas are selected to represent the four subregions of the Unites States: Northeast, Midwest, West, and South. I use 1990 Census STF 3A by the U.S. Census Bureau, Census Transportation Planning Package (CTPP) by Bureau of Transportation Statistics, and digital maps by Environmental Systems Research Institute, Inc. (ESRI).

I select eighteen socioeconomic variables from the Census to identify the characteristics of each neighborhood, including population density, racial composition (percent white), percent of children, percent of old, percent of students in secondary school or under, percent of citizenship, household size, language used at home, percent of bachelor's degree or higher, primary transportation mode to work, travel time to work, unemployment rate, labor force participation rate, household income, income per capita, source of income (percent of workers receiving wage and salary), percent of residents under the absolute poverty level, and home ownership rate.

Table 1.1 Brief Summary of the Four Metropolitan Areas

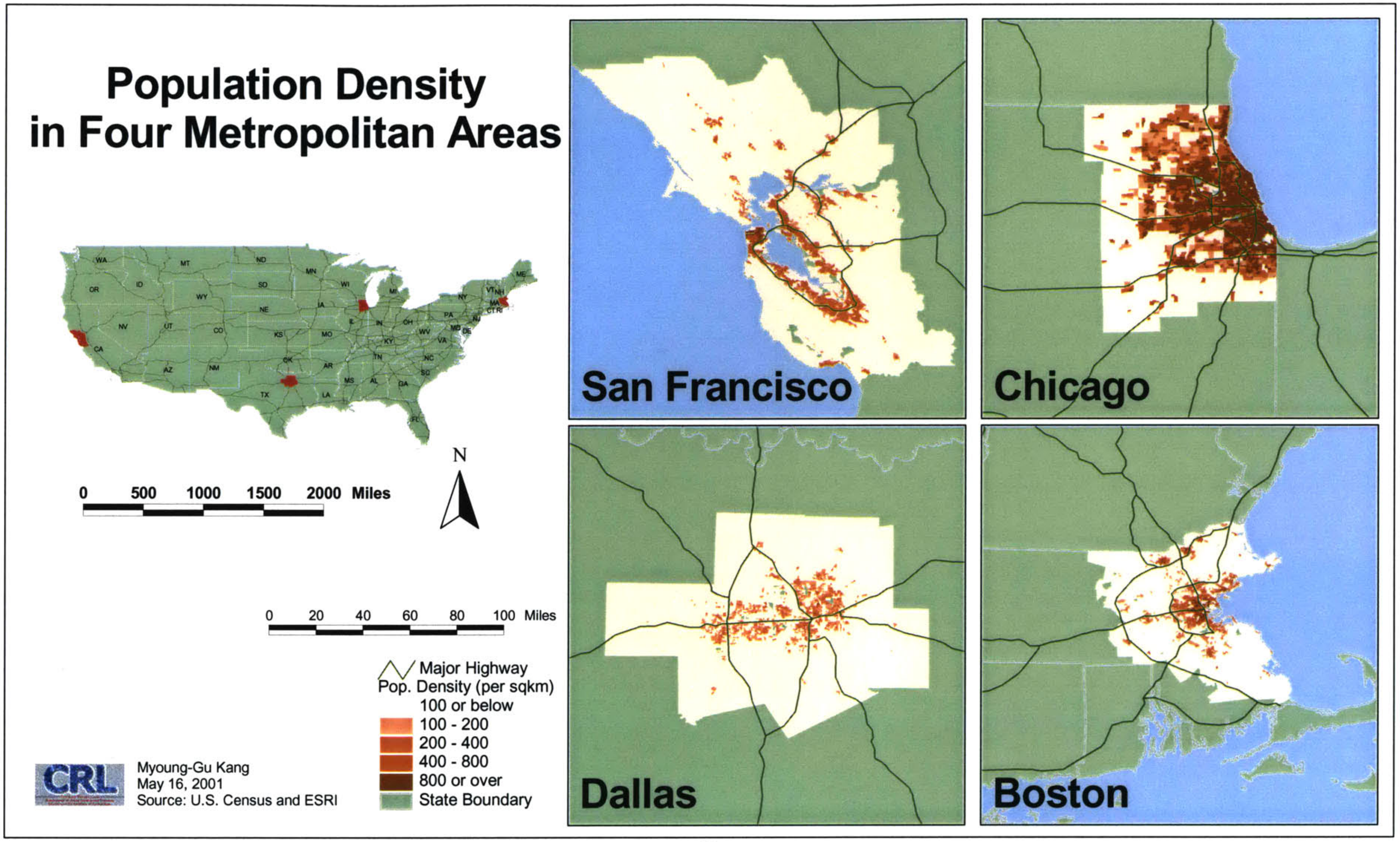
	Number of Blockgoups	Area of Land (Km^2)	Total Population	Income per Capita
Boston, MA *	3,419	6,564.9	3,867,738	18,690
Chicago, IL	6,222	11,363.4	7,326,291	16,736
San Francisco, CA	4,679	18,749.3	6,230,376	19,664
Dallas, TX	3,510	17,968.1	3,884,004	15,904

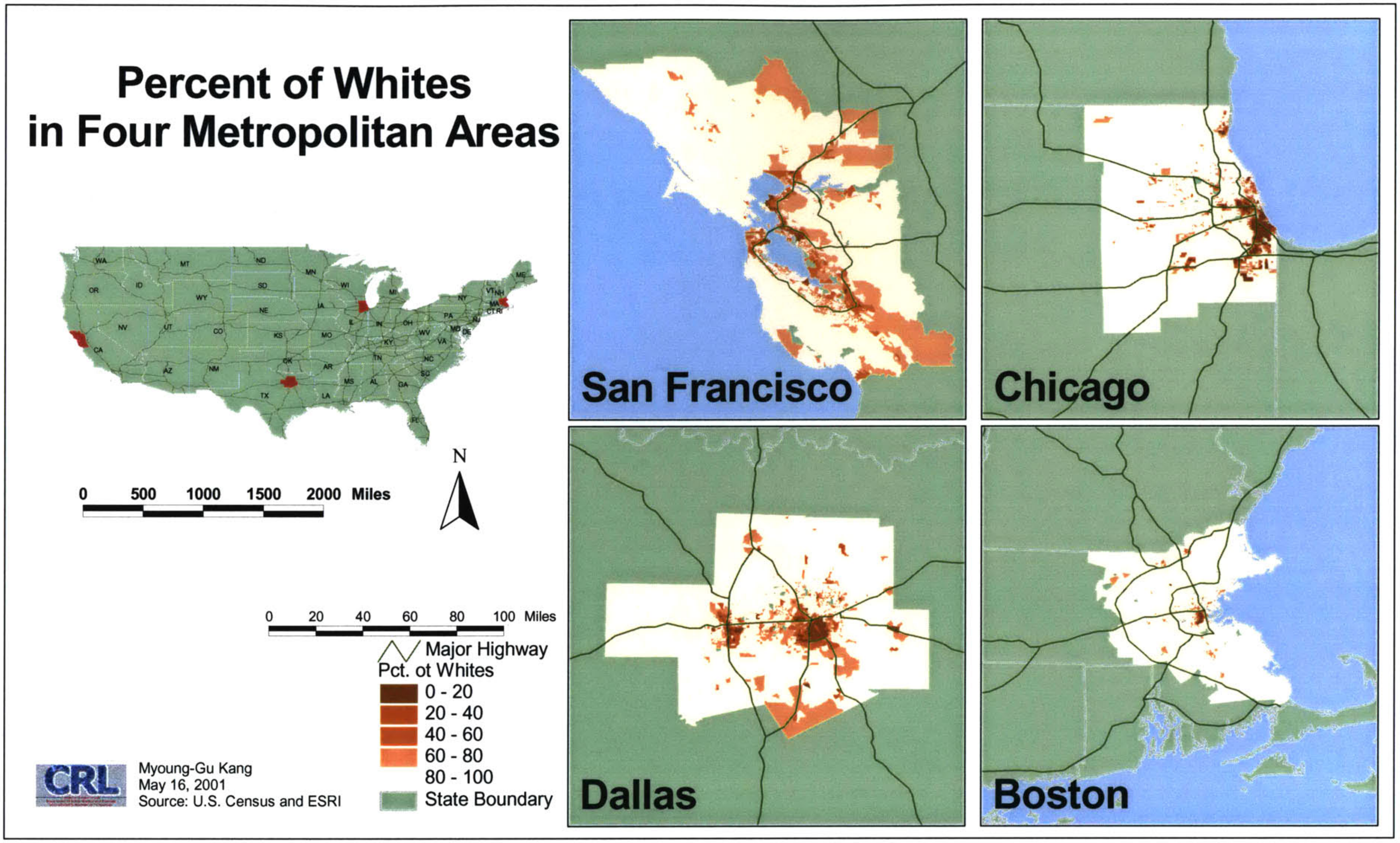
† The Census blockgroups having no population, no household, and no workers have been excluded. (See Appendix for more detail information.)

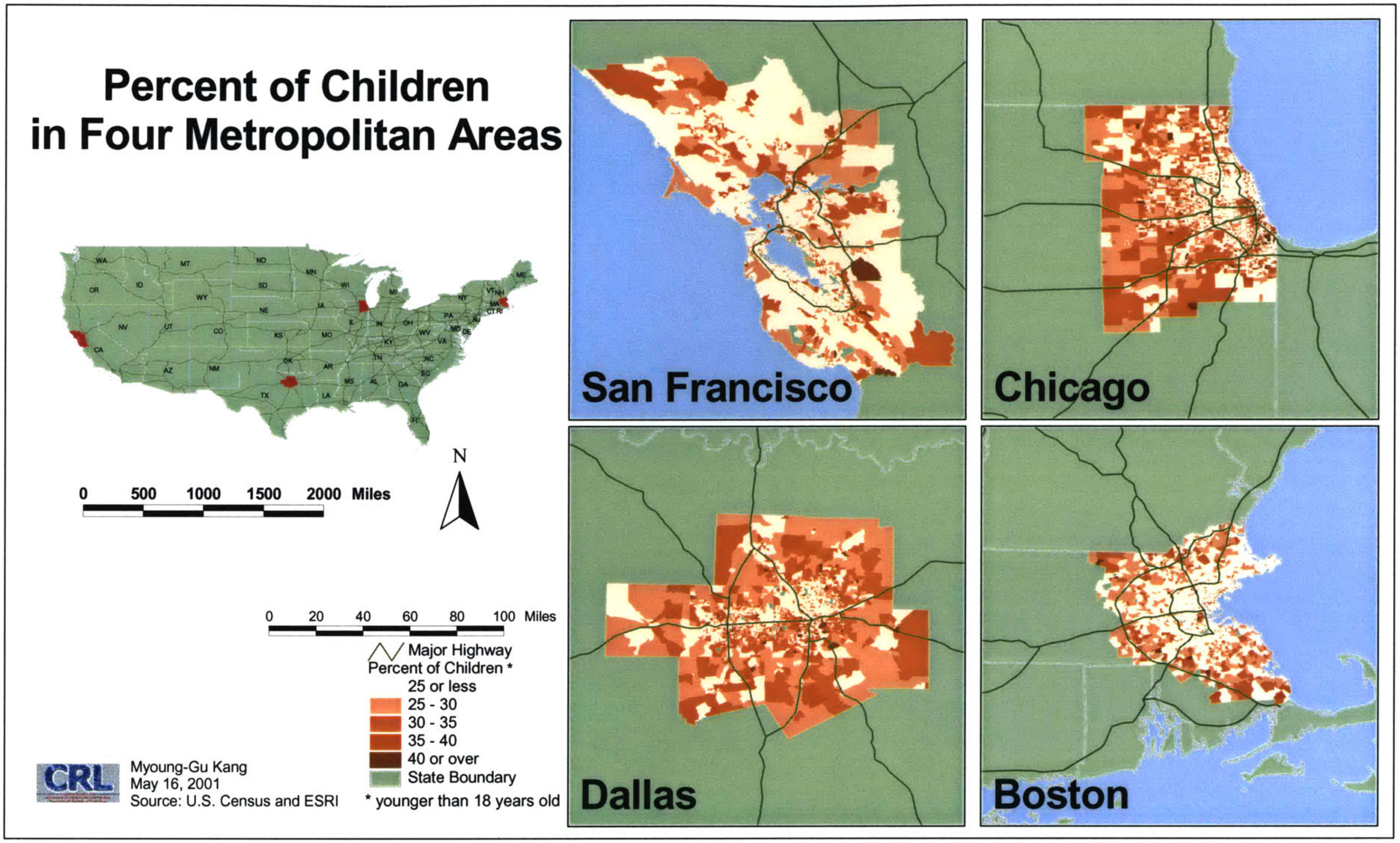
* The error of 1990 Census STF3A of Boston has been fixed.

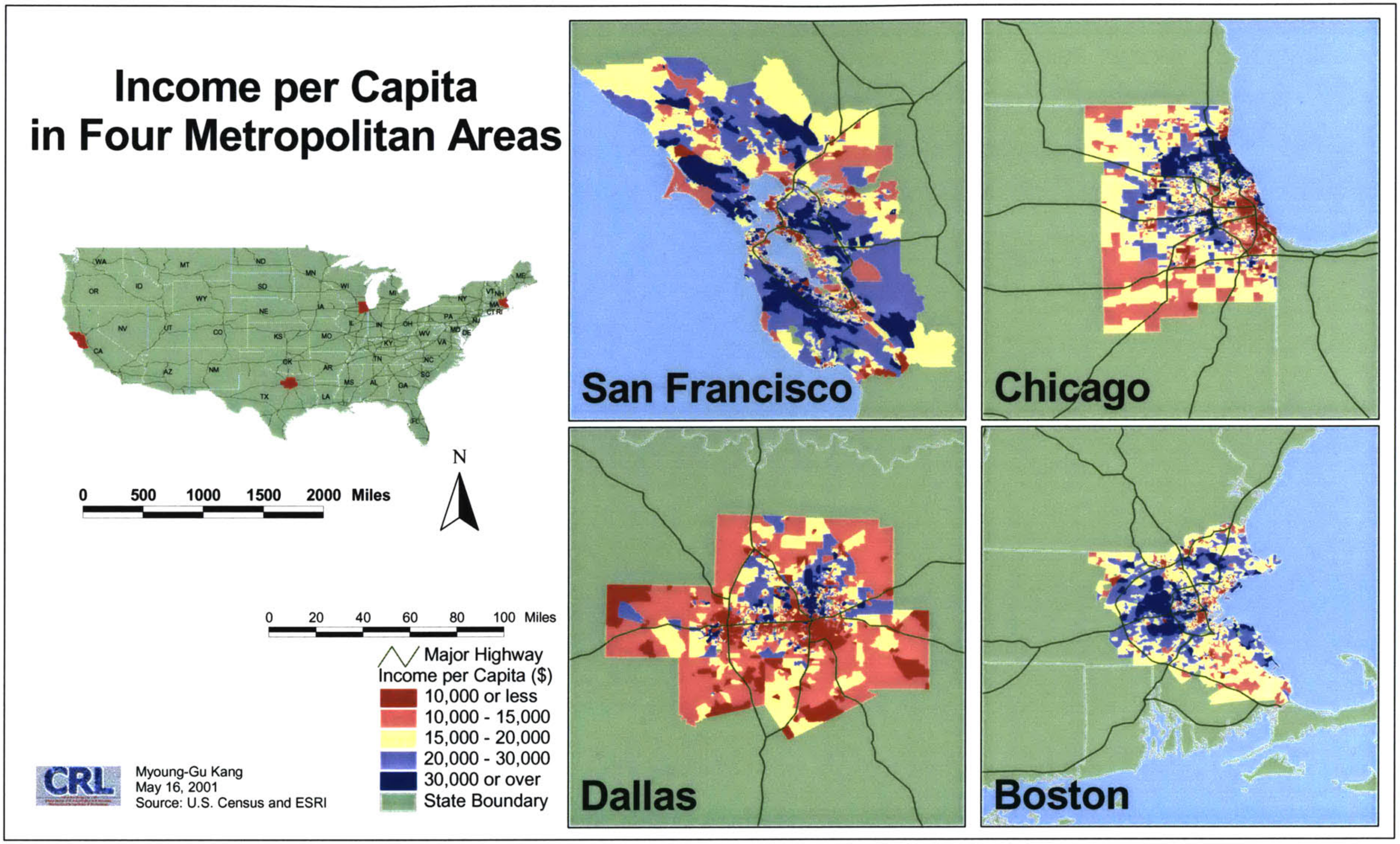
e 1.2 Variables of Socioeconomic Characteristics of Residents

Variable Name	Variable Definition	Related Census Table	Relate Census
POP_DEN	population density	P1	STF3(
WH_PCT	percent white	P8	STF3(
KID_PCT	percent of kids younger than 18	P13	STF3(
OLD_PCT	percent of seniors older than 64	P13	STF3(
CIT_PCT	percent citizen	P37	STF3(
HHS_PCT	percent of households having less than three members	P16	STF3(
ENG_PCT	percent of people who speak English at home	P31	STF3(
STU_PCT	percent of the elementary and secondary students	P54	STF31
CAR_PCT	percent of workers commuting by drive-alone or car pool	P49	STF3(
TIME_AVG	average travel time to work in minutes for whose workers who do not work at home	P51	STF3(
HIED_PCT	percent of adults (25 years and over) with college or higher degree	P57	STF31
LAB_PCT	percent of labor force	P70	STF3(
UEMP_PCT	percent of unemployed workers	P70	STF31
HHI_MED	median household income	P80A	STF31
WAGE_PCT	percent of people with wage or salary income	P90	STF32
INC_PC	income per capita	P114A	STF32
POV_PCT	percent of people below poverty level	P121	STF32
OWN_PCT	percent of home ownership	H19	STF32


Population Density
in Four Metropolitan Areas
0 500 1000 1500 2000 Miles
N
0 20 40 60 80 100 Miles
Major Highway
Pop. Density (per sqkm)
100 or below
100 - 200
200 - 400
400 - 800
800 or over
State Boundary
CRL
Myoung-Gu Kang
May 16, 2001
Source: U.S. Census and ESRI
San Francisco
Chicago
Dallas
Boston


Percent of Whites
in Four Metropolitan Areas
0 500 1000 1500 2000 Miles
N
0 20 40 60 80 100 Miles
Major Highway
Pct. ot Whites
0 - 20
20 - 40
40 - 60
60 - 80
80 - 100
State Boundary
CRL
Myoung-Gu Kang
May 16, 2001
Source: U.S. Census and ESRI
San Francisco
Chicago
Dallas
Boston


Percent of Children
in Four Metropolitan Areas
0 500 1000 1500 2000 Miles
N
0 20 40 60 80 100 Miles
Major Highway
Percent of Children *
25 or less
25 - 30
30 - 35
35 - 40
40 or over
State Boundary
* younger than 18 years old
CRL
Myoung-Gu Kang
May 16, 2001
Source: U.S. Census and ESRI
San Francisco
Chicago
Dallas
Boston


Income per Capita
in Four Metropolitan Areas
0 500 1000 1500 2000 Miles
N
0 20 40 60 80 100 Miles
Major Highway
Income per Capita ($)
10,000 or less
10,000 - 15,000
15,000 - 20,000
20,000 - 30,000
30,000 or over
State Boundary
CRL
Myoung-Gu Kang
May 16, 2001
Source: U.S. Census and ESRI
San Francisco
Chicago
Dallas
Boston

2

QUANTIFYING and STRATIFYING RESIDENTIAL CLUSTERS in BOSTON METROPOLITAN AREA

There are 3.9 million people in the Boston Metropolitan Area as of 1990[1]. The average population density is 589 persons per square kilometer (1,525 persons per square mile). The white population is 3.4 million, which is 88 percent of total population. That is, roughly speaking, nine out of ten people in the Boston Metropolitan Area are white. This is higher than the United States average of 80 percent. 22 percent of the population is under 18 years old, which is less than the United States average of 26 percent. 13 percent of the population is 65 years and over. The adults having bachelor's degree or higher degree out of people aged 25 years and over is 31 percent, which is far greater than the U.S. average of 20 percent.

---

[1] Throughout this thesis, all the numbers are as of 1990 according to the 1990 Census Summary Tape File 3A if there is no other remark.

Table 2.1 Comparison of U.S. and the Four Metropolitan Areas

	U.S. *	Boston	Chicago	San Francisco	Dallas
Population	248,709,873	3,867,738	7,326,297	6,230,376	3,884,004
Land Area	-	6,564.9	11,363.4	18,749.3	17,968.1
Pop. Den.	-	589.2	644.7	332.3	216.2
Pct. White	80.3	88.4	70.9	69.5	75.3
Pct. Kids	25.6	21.8	26.0	23.1	27.2
Pct. Old	12.6	12.7	11.3	11.1	8.0
Pct. LF.	65.3	69.3	68.5	69.6	73.0
Pct. Unemp.	6.2	6.4	6.7	5.1	5.7
Pct. Car	86.5	81.3	80.0	84.2	94.7
Pct. Hi. Ed.	20.3	30.9	24.5	30.9	25.9
Med. HH Inc.	30,056.0	42,167.4	38,500.7	44,119.5	35,925.5
Inc. per Cap.	14,420.0	18,690.8	16,736.0	19,663.6	15,903.5
Pct. Pov.	13.1	8.4	11.3	8.6	11.7
Pct. Own.	64.2	65.2	67.4	60.8	61.9
CPI **	134.6	140.9	135.1	136.7	131.4
Adj. HH Inc.		40,282.0	38,358.2	43,441.7	36,800.4
Adj. Inc pc		17,855.1	16,674.1	19,361.5	16,290.8
V ***		270.0	180.0	165.0	360.0
Radius (km)		52.8	85.1	114.1	75.6
Radius (mi)		32.8	73.1	70.9	47.0

† (1) Population Density, (2) Percent of White, (3) Percent of Kids, (4) Percent of Old, (5) Percent of Labor Force, (6) Unemployment Rate, (7) Percent of Drive Alone or Carpool, (8) Percent of Bachelor's degree or Higher, (9) Median Household Income, (10) Income per capita, (11) Percent of People Under Absolute Poverty Level, and (12) Percent of Home Ownership.

* Source: U.S. Bureau of the Census, 1990 Census of Population and Housing, Summary Tape File 1. The study areas are based on the selected blockgroups which come from U.S. Bureau of the Census, 1990 Census of Population and Housing, Summary Tape File 3

** January 1991 CPI; North Eastern Urban, Chicago CMSA Urban, San Francisco CMSA Urban, and South Urban for all Items. Base period 1982-84 = 100

*** Here, I calculate a radius as a simple measure of geographic size. V is the angle of available wedge, i.e., Area = [(V / 360) * p * R^2]. So, R = [(360 / V) * (Area / p)]$^{1/2}$.

The median household income of the Boston Metropolitan Area was much higher than the U.S. average. The nominal median household income of Boston is 42,000 dollars, which is 40 percent higher than that of U.S. Even though I adjust the nominal income with the Consumer Price Index (CPI), Boston households earn 34 percent more than the U.S. average. The percent of people who are under absolute poverty threshold was 8.4 percent, which is also far less than the United States average of 13.1 percent. The percent of people having bachelor's degree or a higher degree is much higher than the U.S average, and about one out of three adults 25 years or older living in the Boston Metropolitan Area, had college or higher level of education.

2.1 What Latent Factors Differentiate Neighborhoods?

Factor analysis is a statistical approach that can be used to analyze interrelationships among a large number of variables and to explain these variables in terms of their common underlying dimensions (factors). Factor analysis is a way of condensing the information contained in a number of original variables into a smaller set of dimensions (factors) with a minimum loss of information.

Factor analysis is especially useful in social science, where there are no obvious fundamental variables as in physical science, and also no way of performing laboratory experiments to keep selected variables constant. Thus, we can start with what may be a rather arbitrary selection of characteristics and reduce them to a formally fundamental set of factors. This concept is based on the principles of parsimony. Or, in many cases, we interprete the newly sorted variables as a fundamental underlying, or latent, forces dictating the observable social phenomena.

For example, one can summarize the correlation between two variables in a scatterplot. A regression line can then be fitted that represents the "best" summary of the linear relationship between the variables. If we could define a variable that would approximate the regression line in such a plot, then that variable would capture most of the "essence" of the two items. Subjects' single scores on that new factor, represented by the regression line, could then be used in future data analyses to represent that essence of the two items. In a sense we have reduced the two variables to one factor. Note that the new factor is actually a linear combination of the two variables.

The example described above, combining two correlated variables into one factor, illustrates the basic idea of factor analysis, or of principal components analysis. If we extend the two-variable example to multiple variables, then the computations become more involved, but the basic principle of expressing two or more variables by a single factor remains the same.

After we have found the line on which the variance is maximal, there remains some variability around this line. In principal components analysis, after the first factor has been extracted, that is, after the first line has been drawn through the data, we continue and define another line that maximizes the remaining variability, and so on. In this manner, consecutive factors are extracted. Because each consecutive factor is defined to maximize the variability that is not captured by the preceding factor, consecutive factors are independent of each other. Put another way, consecutive factors are uncorrelated or orthogonal to each other.

Note that as we extract consecutive factors, they account for less and less variability. The decision of when to stop extracting factors basically depends on when there is only very little "random" variability left. The nature of this decision is arbitrary; however, various guidelines have been developed. First, we can retain only factors with eigenvalues greater than 1. In essence this is like saying that, unless a factor extracts at least as much as the equivalent of one original variable, we drop it. This criterion was proposed by Kaiser (1960), and is probably the one most widely used. Second, A graphical method is the scree test first proposed by Cattell (1966). We can plot the eigenvalues in ascending order in a simple line plot. Cattell suggests to find the place where the smooth decrease of eigenvalues appears to level off to the right of the plot. To the right of this point, presumably, one finds only "factorial scree" -- "scree" is the geological term referring to the debris which collects on the lower part of a rocky slope. Both criteria have been studied in detail (Browne, 1968; Cattell & Jaspers, 1967; Hakstian, Rogers, & Cattell, 1982; Linn, 1968; Tucker, Koopman & Linn, 1969). Using this general technique, the first method (Kaiser criterion) sometimes retains too many factors, while the second technique (scree test) sometimes retains too few; however, both do quite well under normal conditions, that is, when there are relatively few factors and many cases. In practice, an additional important aspect is the extent to which a solution is interpretable.

The extraction of principal components amounts to a variance maximizing (varimax) rotation of the original variable space. We could rotate the axes in any direction without changing the relative locations of the points to each other; however, the actual coordinates of the points, that is, the factor loadings would of course change. For example, in a scatterplot we can think of the regression line as the original X axis, rotated so that it approximates the regression line. This type of rotation is called variance maximizing because the criterion for (goal of) the rotation is to maximize the variance (variability) of the "new" variable (factor), while minimizing the variance around the new variable

Using the above set of factor analysis, I standardize the description of neighborhoods, which in turn should allow for comparisons between different neighborhoods on a common basis. We can see if the fundamental factors are the same for each neighborhood. In addition, we can calculate factors which are independent, and can be used as basic variables for another model, such as cluster analysis and multiple regression analysis.

In the Boston Metropolitan Area, four principal components of the 18 variables explain 71.4 percent of total variation. In other words, we reduced the variables by 22.2 percent (18 to 4), but they can still explain 71.4 percent of variations across all blockgroups. As we shall see, examination of the four principal components suggests that they focus on baseline, children, income, and age in order of importance.

Figure 2.1 Scree Plot (Boston)

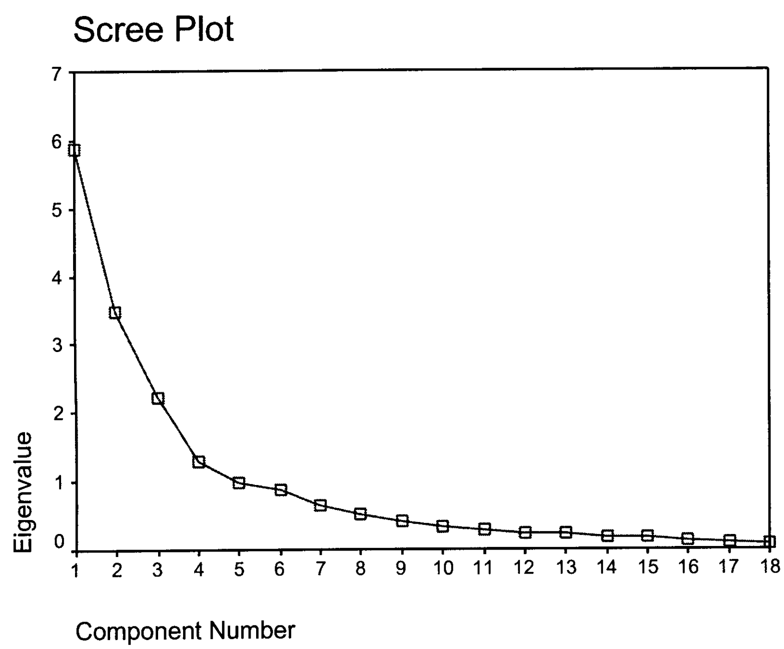


Table 2.2 Percent of variance explained by factors (Boston)

	Initial Eigenvalues			Rotation Sums of Squared Loadings		
Comp.	Total	% of Var.	Cumul. %	Total	% of Var.	Cumul. %
1	5.871	32.618	32.618	4.681	26.006	26.006
2	3.486	19.364	51.982	3.216	17.868	43.874
3	2.203	12.241	64.224	2.642	14.678	58.552
4	1.293	7.184	71.408	2.314	12.855	71.408

* Extraction Method: Principal Component Analysis.

If all the original variables were perfectly independent ideally, the eigen values of each variables would be one and each variable would explain 5.5 percent of the total variation. We would then need all of the variables to explain the differences among neighborhoods. In actuality, however, many observable characteristics are correlated with each other. A single cause can make them move synchronously, or they move to the same direction by chance. Because they vary simultaneously, we can reduce the number of variables according to principles of parsimony.

In Table 2.3, each coefficient in each cell in the component matrix represent the load, in both magnitude and direction. The square of each coefficient shows the percent of loading of each variables on the extracted principal components, i.e., -0.731 means Component 1 is loaded by 53.4 percent of the variation of the population density (POP_DEN). The negative sign indicates a negative relationship between the original variables and the component. That is, neighborhoods having higher scores of Component 1 are less dense areas.

In addition, the composition of components shows the relationship among original variables. For example, we can say that white people tend to live in lower density neighborhoods because a higher score of Component 1 arise in blockgroups with lower density (POP_DEN -0.731) and higher percentage of white people (WH_PCT 0.779) at the same time.

Table 2.3 Rotated Component Matrix (Boston)

	Component 1 *Baseline Factor*	Component 2 *Children Factor*	Component 3 *Income Factor*	Component 4 *Age Factor*
POP_DEN	-.731	-.315	-.082	.106
WH_PCT	.779	-.309	.110	.066
KID_PCT	-.060	.904	-.173	.095
OLD_PCT	.218	-.355	-.067	-.788
CIT_PCT	.786	-.058	.008	.075
HHS_PCT	-.171	-.813	-.002	-.269
ENG_PCT	.768	-.191	.203	.091
STU_PCT	-.052	.894	-.060	.057
CAR_PCT	.753	.419	-.070	-.000
TIME_AVG	-.044	.282	.326	-.015
HIED_PCT	.027	-.259	.839	.184
LAB_PCT	.199	-.013	.102	.856
UEMP_PCT	-.406	.245	-.396	-.169
HHI_MED	.454	.211	.754	.213
WAGE_PCT	.238	.112	.251	.818
INC_PC	.269	-.161	.846	.033
POV_PCT	-.743	.140	-.319	-.242
OWN_PCT	.767	.307	.343	.095

* Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 6 iterations.

** The shaded cells represent main loads to each component. They are the largest load out of each row and not less than 0.6, i.e. at least 36 percent of the variation of each variable ($0.6^2 = 0.36$).

**** The shaded variables (leftmost column) tend to evenly spread across the components instead of focusing on a specific component.

The most important factor of the Boston Metropolitan Area is Component 1, called 'Baseline Factor'. In Boston, this Baseline Factor consists of the seven variables out of eighteen variables.[2] The variables move at the same time, i.e., higher score of Baseline Factor means; higher percent of white, lower population density, higher percent of citizenship, higher percent of english speaking households at home, higher percent of driving to work, lower percent of the absolutely poor population, and higher percent of home ownership in the neighborhood. The compositions of principal components are a little different from each Metropolitan Areas, as you will see in the following chaper.

By itself, this factor explains 26 percent of the total variation among neighborhoods across all 18 variables. High value of this factor apply to neighborhoods of not poor white citizens speaking only English at home, owning a home, living in low density area, and driving to work. Low values arise in neighborhoods of poor non-white foreigners speaking a language other than English at home, not owning a home, living in high density area, and not driving to work.

---

[2] The composition of each factor may subtly vary across metropolitan areas, but I use the same name for similar factor.

The Children Factor is the second most important factor in the Boston Metropolitan Area. Component 2 shows a similar patterns of variation in the percent of children, the percent of students attending secondary school or below, and the percent of households having three and more members. So, Component 2 represents children and family characteristics which are independent from other factors (geometrically perpendicular). It explains 18 percent of overall variation. For example, if a neighborhood has a higher children factor, it tends to have more kids, more students, and more household members.

Third, education and income vary among neighborhoods independently from other socioeconomic variables. The coefficients suggests, if more people have bachelor's or higher degree, the people in the neighborhood tend to have higher household income and personal income per capita. As you notice, both higher education and income go in same direction, i.e., higher educated people have higher income. It explains 15 percent of the total variations.

The last important delineating factor of the Boston Metropolitan Area is what we call 'age' consisting of percentage of old population and percentage of labor force participation. The Component 4 is bigger if a neighborhood has less old population, more workers, and more people whose income source is wage or salary. It explains 13 percent.

In total, all the above four principal factors explain 71.6 percent of variations among neighborhood characteristics in the Boston Metropolitan Area in 1990.

Interestingly, the average travel time to work and unemployment rate are not important component of our socioeconomic factors. Both of them are dispersed in many components, so they don't form independent components. This means that the variation of both of them among neighborhoods are not significantly different compared to other socioeconomic variables. The correlation coefficients between average travel time to work (or unemployment rate) and other variables are also smaller that other coefficients. In other words, the covariance between average travel time and percentage of people driving to work, for example, is weaker than the covariance between percentage of whites and percentage of people driving to work.

2.3 Stratification of Neighborhoods

In order to classify the neighborhoods, I use cluster analysis based on the four independent factors extracted by the factor analysis in the previous section. Cluster analysis is a multivariate analysis technique that seeks to organize information about cases or variables so that relatively homogeneous groups, or "clusters," can be formed. The clusters should be internally homogenous and externally heterogeneous. In other words, members in a cluster are similar to one another in the same group, and members are not like members of other groups.

The joining or tree clustering method uses the dissimilarities or distances between objects when forming the clusters. These distances can be based on a single dimension or multiple dimensions. For example, if we were to cluster fast foods, we could take into account the number of calories they contain, their price, subjective ratings of taste, etc. The most straightforward way of computing distances between objects in a multi-dimensional space is to compute Euclidean distances. If we had a two- or three-dimensional space this measure is the actual geometric distance between objects in the space (i.e., as if measured with a ruler). We can use actual real distances, or some other derived measure of distance that is more meaningful to the researcher.

In this thesis I use Ward's method. This method uses an analysis of variance approach to evaluate the distances between clusters. In short, this method attempts to minimize the Sum of Squares (SS) of any two (hypothetical) clusters that can be formed at each step. (Ward, 1963). In general, this method is regarded as very efficient.

Cluster analysis presents the problem of how many clusters to keep. Yet, no widely accepted statistical method to determine the number of clusters has developed. However, we can use R^2 to decide a reasonal number of groups. For example, we can classify all neighborhoods into two groups, then we can measure R^2 taking the variance between the groups (explaned portion by the clustering) divided by the total variance among all neighborhoods. We can again calculate R^2 with three groups, and so on. R^2 would be 0 (zero) with no classification because there is no explanation. If we use the same number of groups as many as the number of neighborhoods, R^2.would eventually be one because we treat each blockgroup as each particular group.

As we use more number of groups, R^2 increases, but the increment of R^2, generally speaking, decreases. So, there is a trade-off: number of groups vs. explanation power. In other words, we need to decide the smaller number of groups with a minimum loss of explanation power. In this thesis, I decided number of clusters when R^2 is between 0.7 and 0.8, which means the classification explain 70 or 80 percent of the total variation.

Table 2.4 Types of Neighborhoods (Boston)

Type	Factor	Minimum	Maximum	Mean	Std. Deviation
A	Baseline Factor	-1.65725	1.77389	.5391403	0.4317911
N=1,881	Children Factor	-3.24035	2.78211	-0.039556	0.7371879
(55 %)	Income Factor	-3.36684	1.35712	-0.4846772	0.5036457
Generic White	Age Factor	-5.28612	2.86687	0.0510127	0.8919396
B	Baseline Factor	-4.23790	0.82914	-.9561663	0.8502623
N=527	Children Factor	-3.73016	1.27724	-.9994230	0.9772320
(15.4 %)	Income Factor	-1.21592	3.56374	0.4200407	0.7253910
Young Labor	Age Factor	-1.31324	3.19455	0.8050371	0.6173182
C	Baseline Factor	-5.23018	0.76252	-1.3982398	1.1513438
N=454	Children Factor	-0.69366	3.84356	1.2546407	0.7641717
(13.3 %)	Income Factor	-2.15750	2.38877	-0.2111196	0.8621008
Non-white Low Income	Age Factor	-3.62040	2.44578	-0.2428194	0.9576004
D	Baseline Factor	-2.41803	1.11860	0.2240604	0.5051806
N=556	Children Factor	-2.85245	2.28831	0.0566430	0.8303874
(16.3 %)	Income Factor	-1.06391	5.95639	1.4139652	1.0904758
High Income	Age Factor	-6.57086	1.09179	-0.7373552	1.0654339

* The shaded cells are selected by author to highlight the characteristics of each cluster.

In the Boston Metropolitan Area, I find there are four distinct types of neighborhoods. Type A is a typical white neighborhood having a slightly higher score of the Baseline factor. That is, major whites in Boston are generally citizens living in lower density neighborhoods and driving to their workplaces. This type of neighborhoods look like the typical suburban neighborhoods of the Boston Metropolitan Area. 55 percent of the neighborhoods are in this category.

Another group of neighborhoods (Type B) has a non-white younger labor force whose family size is small. 15.4 percent of neighborhoods are in this category. The third type of residents are non-white with many children (13.3 percent). The last unique type of neighborhoods are highly educated people whose income is high.

2.4 Where are the Clusters?

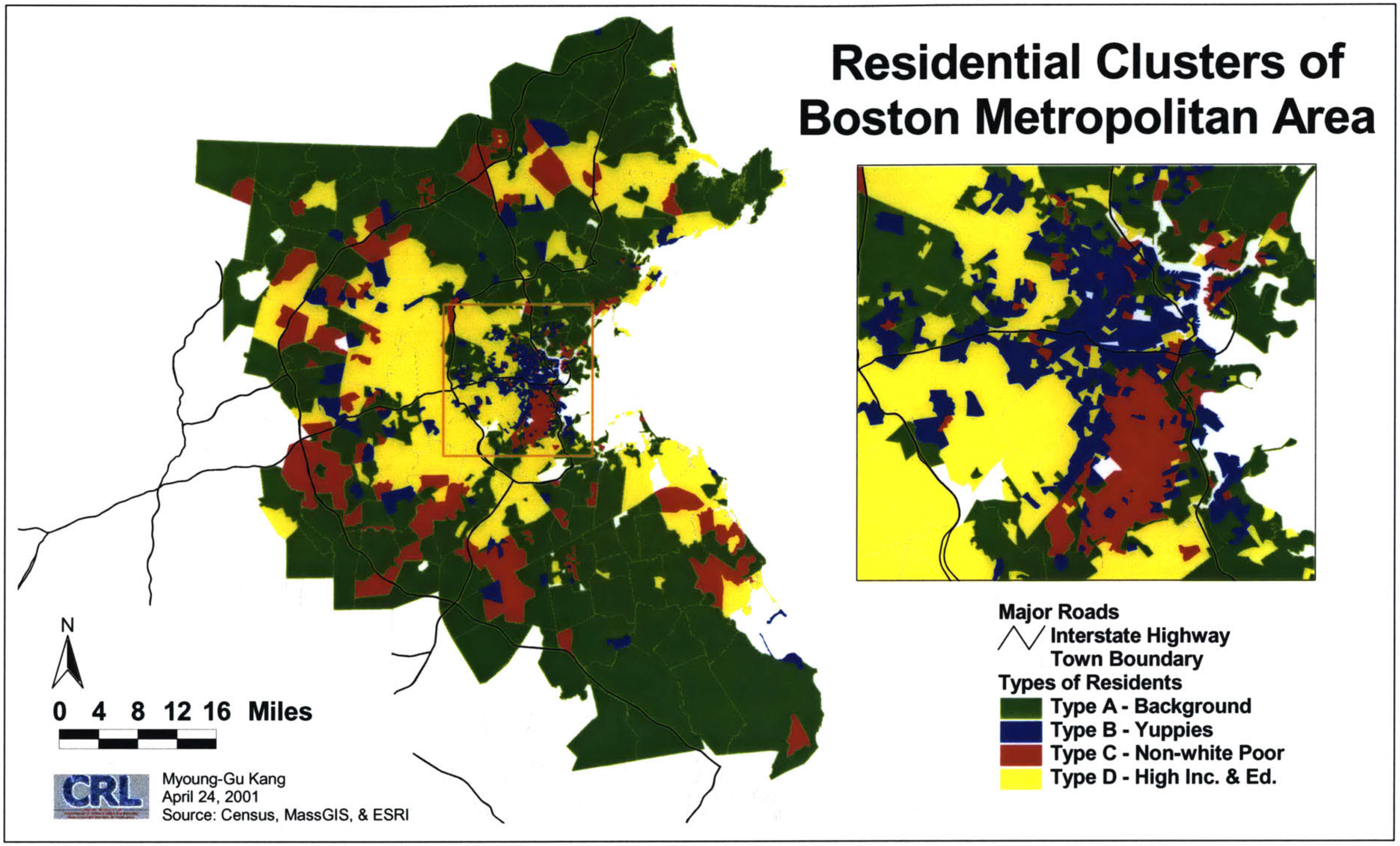
The socioeconomic clustering reveals the locational preferences of residents on the map. First of all, as indicated, socioeconomically similar neighborhoods are geographically close to each other. (Figure 2.2) That is, they tend to agglomerate.

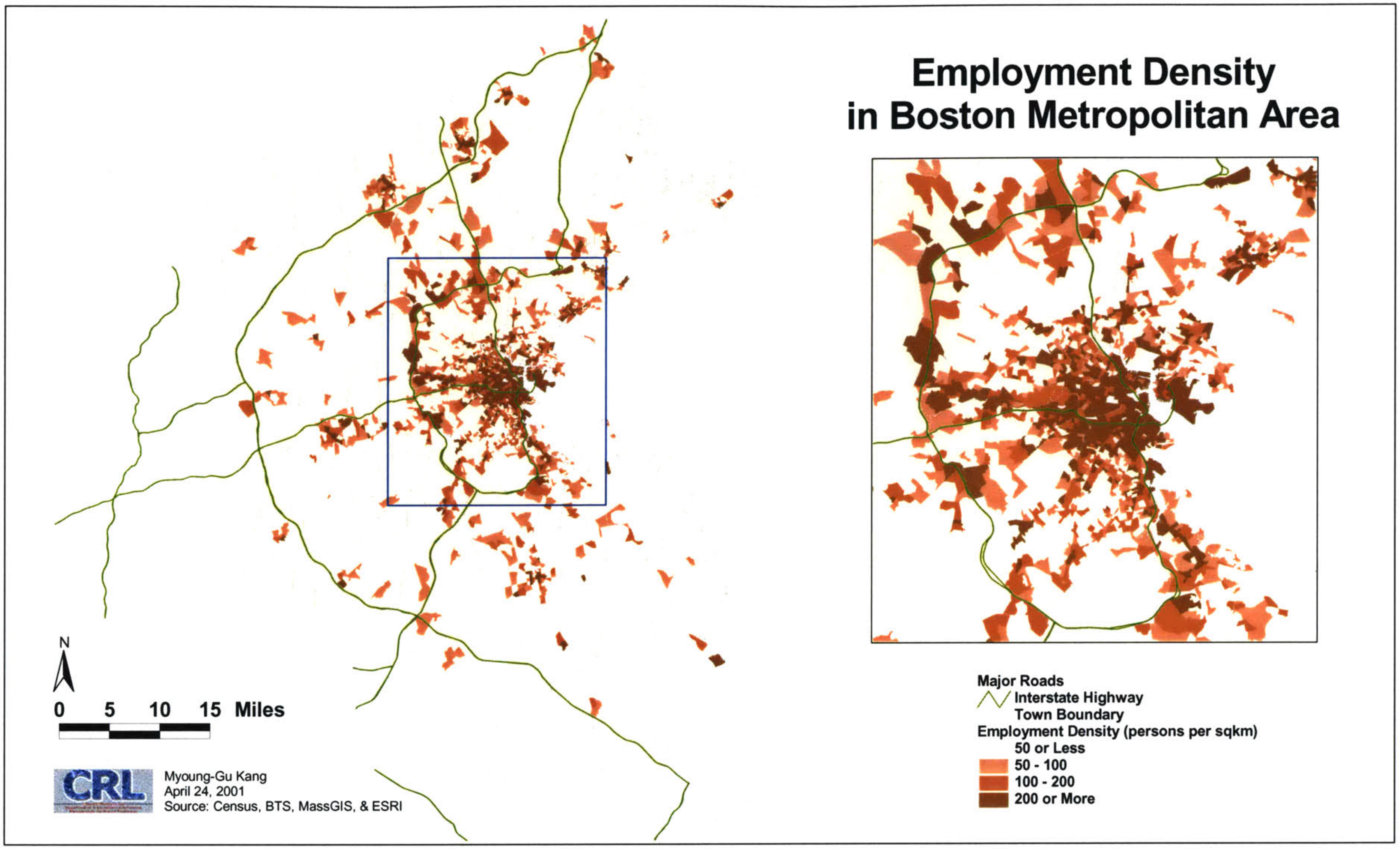
Typical white citizens (type A) live in surburban area of the Boston Metropolitan Area. This type of neighborhood comes the majority of the Boston Metropolitan Area; 55 percent of neighborhoods. These are the generic neighborhoods in Boston. Generally, young workers tend to have less children and live near the city, including nothern Boston, Cambridge, Somerville, and the western area along the Interstate Highway 90 which has many jobs. Non-whites with many children live in downtown Boston, Lowell, Lawrence, or subcenters around Interstate Highway 495. Highly educated people earning higher income mainly live in the western suburban area of Boston.

Interestingly, the edges of the neighborhood agglomeration match the town boundaries. Considering the fact that the city/town governments have the authority to decide the school quality, property tax rate, zoning, etc., the relationship between neighborhood characteristics and municipal government would be another topic for further research.

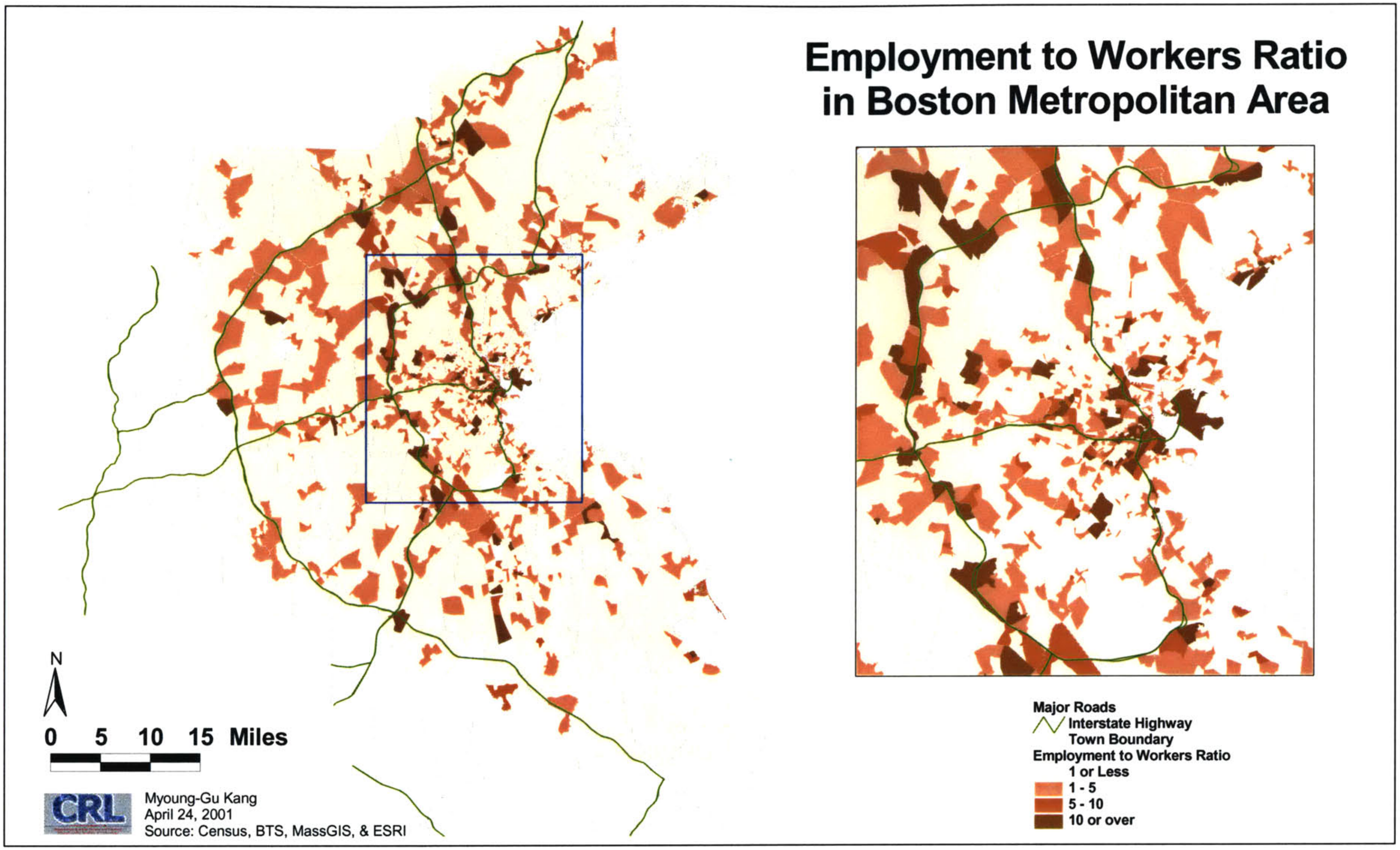
Non-whites with many children live in the downtowns of cities, but their average time to work is not less than that of other areas. In a simple monocentric city model of urban economics, downtown residents should have more benefit of transportation than suburban residents. It is another possible topic that should be examined in terms of location choice in the future.[3]

[3] Note that the geographic unit of this analysis is a Census blockgroup, especially when you see maps in this paper. Each blockgroup can be assumed equivalent for the purpose of this thesis in terms of population or socioeconomic characteristics. However, the bigger shading on the map may not represent more people or more importance because the physical sizes of blockgroups vary. The blockgroups in the suburban area are generally larger.


Residential Clusters of
Boston Metropolitan Area
N
0 4 8 12 16 Miles
Myoung-Gu Kang
April 24, 2001
Source: Census, MassGIS, & ESRI
Major Roads
Interstate Highway
Town Boundary
Types of Residents
Type A - Background
Type B - Yuppies
Type C - Non-white Poor
Type D - High Inc. & Ed.



Employment Density
in Boston Metropolitan Area
N
0 5 10 15 Miles
CRL
Myoung-Gu Kang
April 24, 2001
Source: Census, BTS, MassGIS, & ESRI
Major Roads
Interstate Highway
Town Boundary
Employment Density (persons per sqkm)
50 or Less
50 - 100
100 - 200
200 or More


Employment to Workers Ratio
in Boston Metropolitan Area
N
0 5 10 15 Miles
CRL
Myoung-Gu Kang
April 24, 2001
Source: Census, BTS, MassGIS, & ESRI
Major Roads
Interstate Highway
Town Boundary
Employment to Workers Ratio
1 or Less
1 - 5
5 - 10
10 or over

3

CHICAGO, SAN FRANCISCO and DALLAS[1]

In this chapter, three more metropolitan areas are examined; Chicago (IL CMSA), San Francisco - Oakland - San Jose (CA CMSA), and Dallas - Fort Worth (TX CMSA). What are the factors which delineate the neighborhoods in each metropolitan areas? Are the factors similar across the metropolitan areas? What are the common factors? How about the geographical distribution of each metropolitan area? Do they show identical preferences or different preferences across the metropolitan areas? I use the same methodology and basic criteria to look at these metropolitan areas as applied in the Boston Metropolitan Area.

---

[1] Refer to Appendix for all the detailed statistics.

3.1 Chicago Metropolitan Area

Chicago has about twice bigger land and population (11,363 km^2 and 7.3 million population) than Boston has (6,565 km^2 and 3.9 million population). The average population density is 644.7 persons per square kilometer, quite similar population density to that of Boston Metropolitan Area (589 persons per square kilometer). The percent of whites is 70.9 percent, which is less than the U.S. average (80.3 percent). The household income (38,501 dollars) and income per capita (16,736 dollars) are greater than those of the the United States average (30,056 dollars and 14,420 dollars, respectively).

Using the same methodology as I did on the Boston Metropolitan Area, I identified five key factors in Chicago Metropolitan Area: Baseline factor, children factor, income factor, age factor, and citizenship. Roughly speaking, we have a similar set of underlying factors in the Chicago Metropolitan Area as in the Boston Metropolitan Area, except the percent of the English speaking citizens.

Like the Boston Metropolitan Area, baseline factor is the most important factor of the Chicago Metropolitan Area. This component consists of; population density (-), percent of white (+), driving to work (+), unemployment rate (-), percent of absolutely poor people (-), and home ownership (+). That is, a higher score of this factor means; lower population density, higher percent of white, more driving to work, lower unemployment rate, lower percent of absolutely poor residents, and more home owners.

Second, children and household size relates to the location of residents in Chicago Metropolitan Area even though simple correlation coefficients show that the correlations between the percent of children and the other variables are not large.

The third important factor is education and income. These two variables move the same way, i.e., highly educated people generally get higher income. The fourth factor is the percent of younger labor force.

The fifth factor is percentage of non-citizen population. It is unique in the Chicago Metropolitan Area. As you will see, Dallas has the same non-citizen factor.

Table 3.1 Rotated Component Matrix (Chicago)

	Component 1 *Baseline Factor* (22.3 %)**	Component 2 *Children Factor* (17.6 %)	Component 3 *Income Factor* (16.0 %)	Component 4 *Age Factor* (12.5 %)	Component 5 *Citizen Factor* (12.0 %)
POP_DEN	-0.675	-0.072	-0.090	0.011	-0.434
WH_PCT	0.727	-0.296	0.329	0.092	-0.094
KID_PCT	-0.129	0.892	-0.187	0.091	-0.044
OLD_PCT	0.195	-0.506	0.002	-0.749	0.000
CIT_PCT	0.017	-0.004	0.025	-0.030	0.962
HHS_PCT	-0.085	-0.890	0.034	-0.140	0.051
ENG_PCT	0.076	-0.097	0.139	-0.012	0.948
STU_PCT	-0.128	0.894	-0.100	0.028	-0.005
CAR_PCT	0.870	0.113	-0.014	0.122	0.092
TIME_AVG	-0.517	0.294	0.105	-0.140	0.162
HIED_PCT	0.024	-0.210	0.847	0.225	0.061
LAB_PCT	0.265	-0.056	0.169	0.877	-0.035
UEMP_PCT	-0.655	0.299	-0.369	-0.232	0.131
HHI_MED	0.418	0.140	0.835	0.143	0.113
WAGE_PCT	0.335	0.135	0.263	0.799	-0.024
INC_PC	0.190	-0.206	0.882	0.052	0.088
POV_PCT	-0.737	0.247	-0.364	-0.272	-0.013
OWN_PCT	0.738	0.204	0.366	-0.053	0.226

* Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 8 iterations.

** Variance Explained by the component.

*** The shaded cells represent main loads to each component. They are the largest load out of each row and not less than 0.6, i.e. at least 36 percent of the variation of each variable ($0.6^2 = 0.36$).

**** The shaded variable (leftmost column) tends to evenly spread across the components instead of focusing on a specific component.

With these five factors, I do cluster analysis on Chicago Metropolitan Area. I find that threre are five clusters in Chicago Metropolitan Area: generic white neighborhoods as background (Type A), the young high-income neighborhoods with less children (Type B), non-white poor citizens' neighborhood (Type C), highly educated high-income neighborhoods (Type D), and non-citizen neighborhoods with many children (Type E).

In the Chicago Metropolitan Area, highly educated high income people formed their neighborhoods and are located at the northern part and around the West near the intersections of major highways.

Young labors with less children tend to agglomerate at the Nothern part along the lake side or at the suburban area.

Even though the land is physically much flatter than Boston Metropolitan Area, we see clear "wedges" at least in the inside of the inner interstate highway. Beyond the inner interstate highway, we see some mixture pattern of neighborhoods instead of continuous and homogeneous circumferential pattern – which is almost always assumed in the traditional urban economic model. Therefore, social factors affect people's choice of residential location at least as much as economic factors do. Furthermore, the combined effects – not a single effect – are necessary for us to understand the land use correctly, which is not easily captured in the multivariate regression analysis.

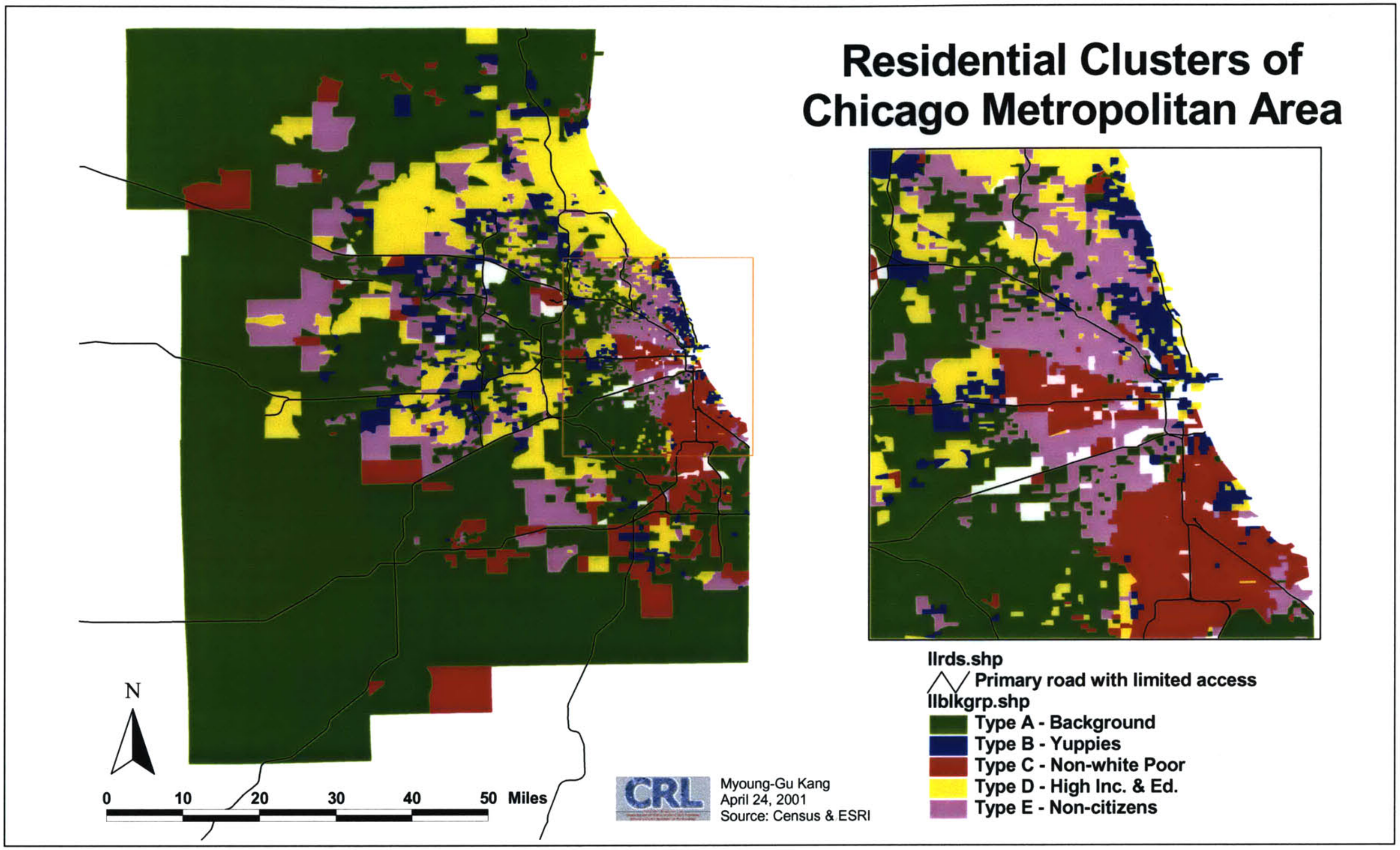
As you will see again, the socioeconomically similar neighborhoods generally tend to be located close to similar neighborhoods in the Chicago Metropolitan Area.

Notice that each factor alone is not enough to tell the characteristics of the types of neighborhoods. As you see in the cluster anlaysis, for example, a 'combination' of the percent of children and other variables define a neighborhood's characteristics.There are a group of neighborhoods in which non-citizens speaking a language other than English at home with more children live. These type of neighborhoods appear at the Northwest and Southwest wedges from downtown Chicago. They also appear along the circumferential highway corridor.

Table 3.2 Types of Neighborhoods (Chicago)

Type	Factor	Minimum	Maximum	Mean	Std. Deviation
A	Baseline Factor	-1.572	2.378	0.732	0.351
N = 2,479	Children Factor	-4.840	2.982	-0.198	0.748
(39.8 %)	Income Factor	-2.269	1.654	-0.391	0.464
Background	Age Factor	-7.226	3.047	-0.070	1.009
	Citizen Factor	-1.981	1.512	0.204	0.517
B	Baseline Factor	-3.740	0.647	-0.801	0.959
N = 530	Children Factor	-3.974	0.634	-1.356	0.942
(8.5 %)	Income Factor	-1.365	3.018	0.399	0.755
Young Labor	Age Factor	-0.382	3.531	1.313	0.652
	Citizen Factor	-2.442	1.609	0.087	0.649
C	Baseline Factor	-4.822	1.589	-1.189	0.976
N = 1,177	Children Factor	-2.889	3.488	0.610	0.878
(18.9 %)	Income Factor	-1.954	1.313	-0.526	0.480
Non-white	Age Factor	-6.310	2.088	-0.350	1.008
Low Income	Citizen Factor	-1.628	2.286	0.885	0.391
D	Baseline Factor	-3.412	1.066	0.096	0.583
N = 810	Children Factor	-3.361	3.367	-0.057	1.039
(13.0 %)	Income Factor	-0.545	7.760	1.739	1.304
High Incomers	Age Factor	-4.055	1.924	-0.320	0.765
	Citizen Factor	-3.006	1.921	0.129	0.494
E	Baseline Factor	-3.959	1.470	-0.055	0.752
N = 1,226	Children Factor	-1.970	4.374	0.439	0.825
(19.7 %)	Income Factor	-1.623	3.980	-0.027	0.640
Non-citizens	Age Factor	-2.586	2.850	0.120	0.721
	Citizen Factor	-4.637	0.810	-1.386	1.129

* The shaded cells are selected by author to highlight the characteristics of each cluster.


Residential Clusters of Chicago Metropolitan Area
llrds.shp
Primary road with limited access
llblkgrp.shp
Type A - Background
Type B - Yuppies
Type C - Non-white Poor
Type D - High Inc. & Ed.
Type E - Non-citizens
N
0 10 20 30 40 50 Miles
CRL
Myoung-Gu Kang
April 24, 2001
Source: Census & ESRI

3.2 San Francisco Metropolitan Area

San Francisco has a land area three times larger (18,749 km^2) than Boston (6,565 km^2), but about one and a half times more population (6.2 millions) than Boston (3.9 millions). The average population density, therefore, is 332 persons per square kilometer, i.e., about a half of Boston Metropolitan Area (589 persons per square kilometer). The average percent of white people is 69.5 percent, which is less than the U.S. average, 80.3 percent. The household income (44,119 dollars) and income per capita (19,663 dollars) are greater than those of the the United States average (30,056 dollars and 14,420 dollars, respectively).

Using the same factor analysis approach as for Boston and Chicago, I identify five key factors in the San Francisco metropolitan area which can be interpreted, in order of importance, as primarily related to income, children, baseline, age, and the average travel time to work. Four of these are the same factors as the above metropolitan areas, and one factor is new which is average travel time to work.

Table 3.3 Rotated Component Matrix (San Francisco)

	Component 1 *Baseline Factor* (16.2 %)**	Component 2 *Children Factor* (18.9 %)	Component 3 *Income Factor* (21.4 %)	Component 4 *Age Factor* (13.7 %)	Component 5 *Time Factor* (6.5 %)
POP_DEN	-0.620	-0.378	-0.229	0.061	0.221
WH_PCT	0.606	-0.197	0.470	0.086	-0.260
KID_PCT	-0.048	0.866	-0.224	0.194	0.150
OLD_PCT	0.060	-0.314	0.128	-0.852	-0.162
CIT_PCT	0.928	-0.133	0.086	-0.022	0.103
HHS_PCT	0.235	-0.848	-0.051	-0.240	-0.040
ENG_PCT	0.890	-0.224	0.208	-0.041	0.113
STU_PCT	-0.071	0.843	-0.133	0.145	0.201
CAR_PCT	0.488	0.557	0.281	0.049	-0.375
TIME_AVG	0.025	0.186	0.055	-0.004	0.724
HIED_PCT	0.090	-0.397	0.711	0.156	0.227
LAB_PCT	0.053	-0.011	0.176	0.890	-0.068
UEMP_PCT	-0.129	0.148	-0.622	-0.130	0.298
HHI_MED	0.095	0.189	0.886	0.137	0.108
WAGE_PCT	-0.054	0.215	0.227	0.849	-0.099
INC_PC	0.179	-0.263	0.804	-0.007	0.154
POV_PCT	-0.232	-0.042	-0.706	-0.117	0.302
OWN_PCT	0.277	0.453	0.673	-0.176	0.045

* Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 8 iterations.

** Variance Explained by the component.

*** The shaded cells represent main loads to each component. They are the largest load out of each row and not less than 0.6, i.e. at least 36 percent of the variation of each variable (0.6^2 = 0.36).

**** The shaded variable (leftmost column) tends to evenly spread across the components instead of focusing on a specific component.

The most important factor of the San Francisco Metropolitan Area is the income component. This component consists of; percent of highly educated people (+), median household income (+), income per capita (+), and percent of absolutely poor people (-). Higher score of this factor means; higher percent of highly educated people, higher median household income, higher income per capita, and lower percent of absolutely poor people. This component alone explain the 21 percent of total variations among neighborhoods in San Fracisco Metropolitan Area.

Second, children and household size is another distinct factor amongh neighborhoods in San Francisco Metropolitan Area. Simple correlation coefficients show very little correlation between the percent of children and the other variables. However, this factor alone explains 19 percent of the total.

The third important factor is baseline. The fourth factor is the percentage of younger labor force. Finally, the average travel time to work is a unique factor in the San Francisco.

Cluster analysis shows that there are five clusters in San Francisco Metropolitan Area: High income neighborhood, young labor force with no children, low income people spending more time to get to work, white citizens, and non-white people. As you see on the map, the neighborhoods in a same group generally tend to agglomerate. The high income residents or young workers live at the left side of the bay along the ocean. Non-white people or poor residents live at the right side of the bay along the water. Beyond them, we can find wealthy neighborhood again, then relatively low income people appear next.

In the San Francisco Metropolitan Area, generic white neighborhoods (Type A) tend to be located in low density suburban area. Yuppies (Type B), who are young labor force with no children, appear at the Northwestern and Southwestern of downtown and left bottom side of inner bay.

Interestingly, the entering points at the East side of the bay, of the bridges to downtown, are occupied by low income neighborhoods. This could be examined in a later study. It could have been a very far fringe of the San Francisco Metropolitan Area created by the geography, "water body". The affluent residents expanded outward along the ocean side. The bridge suddenly created a new land near downtown, then low income people went there because of the access. The rich people can go a liitle further than them, then reside at the East of the bay.

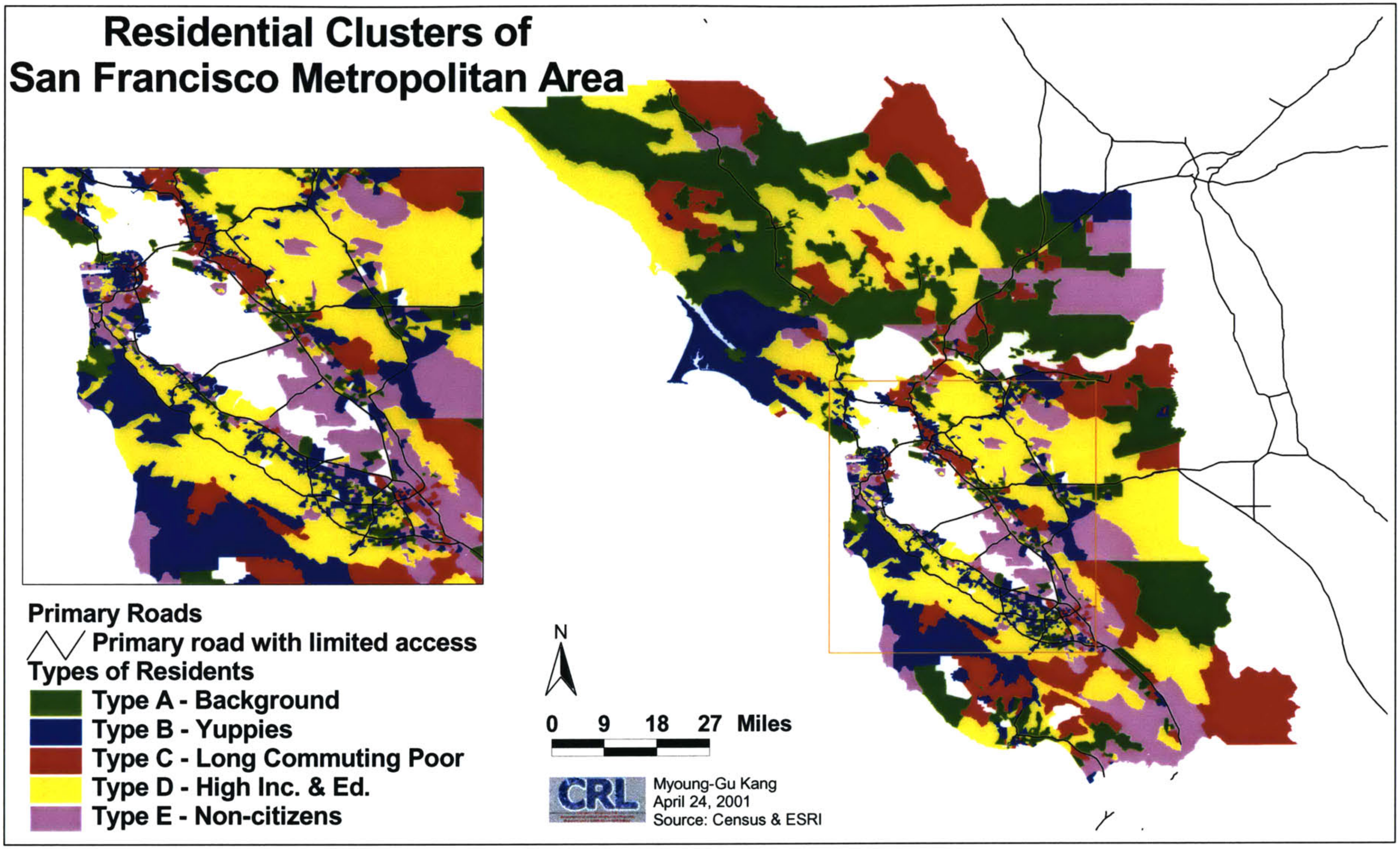
We can observe the long commuting low income neighborhood clusters (Type E) at the right side of the bay in San Fracisco Metropolitan Area. These group of neighborhoods are low income people, however, spend more time to get to workplace than other groups. They tend to be located either on the outskirt of the metropolitan area or on the right side of the bay along the interstate highway corridors.

Table 3.4 Types of Neighborhoods (San Francisco)

Type	Factor	Minimum	Maximum	Mean	Std. Deviation
A	Baseline Factor	-1.043	1.905	0.752	0.403
N = 916	Children Factor	-2.080	3.323	0.232	0.585
(19.6 %)	Income Factor	-2.537	0.888	-0.462	0.526
Background	Age Factor	-2.855	2.538	0.156	0.692
	Time Factor	-4.292	1.211	-0.505	0.624
B	Baseline Factor	-3.771	1.955	-0.157	0.910
N = 1,000	Children Factor	-3.962	1.720	-1.053	0.998
(21.4 %)	Income Factor	-2.338	2.667	-0.094	0.688
Yuppies	Age Factor	-2.622	2.818	0.792	0.609
	Time Factor	-2.805	3.499	0.074	0.954
C	Baseline Factor	-3.137	2.437	0.026	0.971
N = 695	Children Factor	-3.606	2.997	0.560	0.842
(14.9 %)	Income Factor	-4.170	1.071	-1.087	0.930
Long Commut	Age Factor	-6.507	2.873	-0.347	0.904
Poor	Time Factor	-0.934	5.263	1.063	0.915
D	Baseline Factor	-3.840	1.669	0.316	0.592
N = 1,199	Children Factor	-3.086	1.612	-0.149	0.633
(25.6 %)	Income Factor	-1.166	4.525	0.929	0.853
High Incomers	Age Factor	-6.957	1.731	-0.678	1.125
	Time Factor	-4.017	3.348	-0.126	0.945
E	Baseline Factor	-5.654	0.823	-1.072	1.058
N = 866	Children Factor	-2.285	3.429	0.728	0.786
(18.5 %)	Income Factor	-2.052	3.622	0.183	0.670
Non-whites	Age Factor	-6.314	2.007	0.137	0.747
	Time Factor	-4.328	4.166	-0.230	0.878

* The shaded cells are selected by author to highlight the characteristics of each cluster.

Generally speaking, the young workers and low income people tend to live closer to the highway corridor than other residents do. The residential location of San Francisco Metropolitan Area seems more likey a mixture of the parrallel development by the sea and the circumferencial development from the downtown thanks to the bridges.


Residential Clusters of
San Francisco Metropolitan Area
Primary Roads
Primary road with limited access
Types of Residents
Type A - Background
Type B - Yuppies
Type C - Long Commuting Poor
Type D - High Inc. & Ed.
Type E - Non-citizens
N
0 9 18 27 Miles
CRL
Myoung-Gu Kang
April 24, 2001
Source: Census & ESRI

3.3 Dallas Metropolitan Area

Dallas has about three times larger land area (17,968 km^2) than Boston (6,565 km^2), but the same number of population (3.9 million) as Boston (3.9 million). The average population density, hence, is 216 persons per square kilometers, i.e., approximately one third of Boston Metropolitan Area (589 persons per square kilometer). The average percent of whites is 75.3 percent, which is a little bit less than the U.S. average, 80.3 percent. The percent of eldery people (65 years and over) was 8 percent, far less than U.S. average, (12.6 percent). The percent of labor force was 73 percent higher than the U.S. average (65.3 percent). The median household income (35,926 dollars) and income per capita (15,904 dollars) are greater than those of the the United States average (30,056 dollars and 14,420 dollars, respectively).

As shown in Table 3.5, I find five key factors in Dallas metropolitan area: Children, baseline, income, age, and citizenship in order of importance. These variables are consistent with other metropolitan areas.

Table 3.5 Rotated Component Matrix (Dallas)

	Component 1 *Baseline Factor* (16.1 %)**	Component 2 *Children Factor* (18.5 %)	Component 3 *Income Factor* (16.0 %)	Component 4 *Age Factor* (14.3 %)	Component 5 *Citizen Factor* (13.2 %)
POP_DEN	-0.276	-0.238	-0.026	0.276	-0.535
WH_PCT	0.696	-0.184	0.334	0.050	0.209
KID_PCT	-0.095	0.890	-0.197	0.098	-0.116
OLD_PCT	0.067	-0.309	0.000	-0.863	0.099
CIT_PCT	0.153	-0.112	0.120	-0.074	0.904
HHS_PCT	-0.021	-0.919	0.037	-0.109	0.060
ENG_PCT	0.139	-0.221	0.250	-0.017	0.866
STU_PCT	-0.086	0.898	-0.068	0.026	-0.003
CAR_PCT	0.814	0.085	-0.021	0.138	0.116
TIME_AVG	-0.108	0.432	-0.192	0.191	0.429
HIED_PCT	0.172	-0.287	0.822	0.180	0.070
LAB_PCT	0.248	-0.100	0.078	0.883	-0.037
UEMP_PCT	-0.721	0.139	-0.313	-0.182	0.015
HHI_MED	0.336	0.185	0.872	0.087	0.153
WAGE_PCT	0.302	0.110	0.139	0.848	-0.051
INC_PC	0.188	-0.203	0.884	-0.031	0.101
POV_PCT	-0.735	0.093	-0.387	-0.194	-0.237
OWN_PCT	0.461	0.494	0.347	-0.243	0.388

* Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 6 iterations.

** Variance Explained by the component.

*** The shaded cells represent main loads to each component. They are the largest load out of each row and not less than 0.6, i.e. at least 36 percent of the variation of each variable (0.6^2 = 0.36).

**** The shaded variables (leftmost column) tend to evenly spread across the components instead of focusing on a specific component.

The most important factor of the Dalllas Metropolitan Area is the children and household size component. Again, simple correlation coefficients show there is no correlation between the percent of children and the other variables. Nevertheless, this component alone explain the 19 percent of total variation.

The second is the baseline factor. This component consists of; percent driving to work (+), unemployment rate (-), and percent absolutely poor people (-).This component alone explain the 16 percent of total variation among neighborhoods in the Dallas Metropolitan Area.

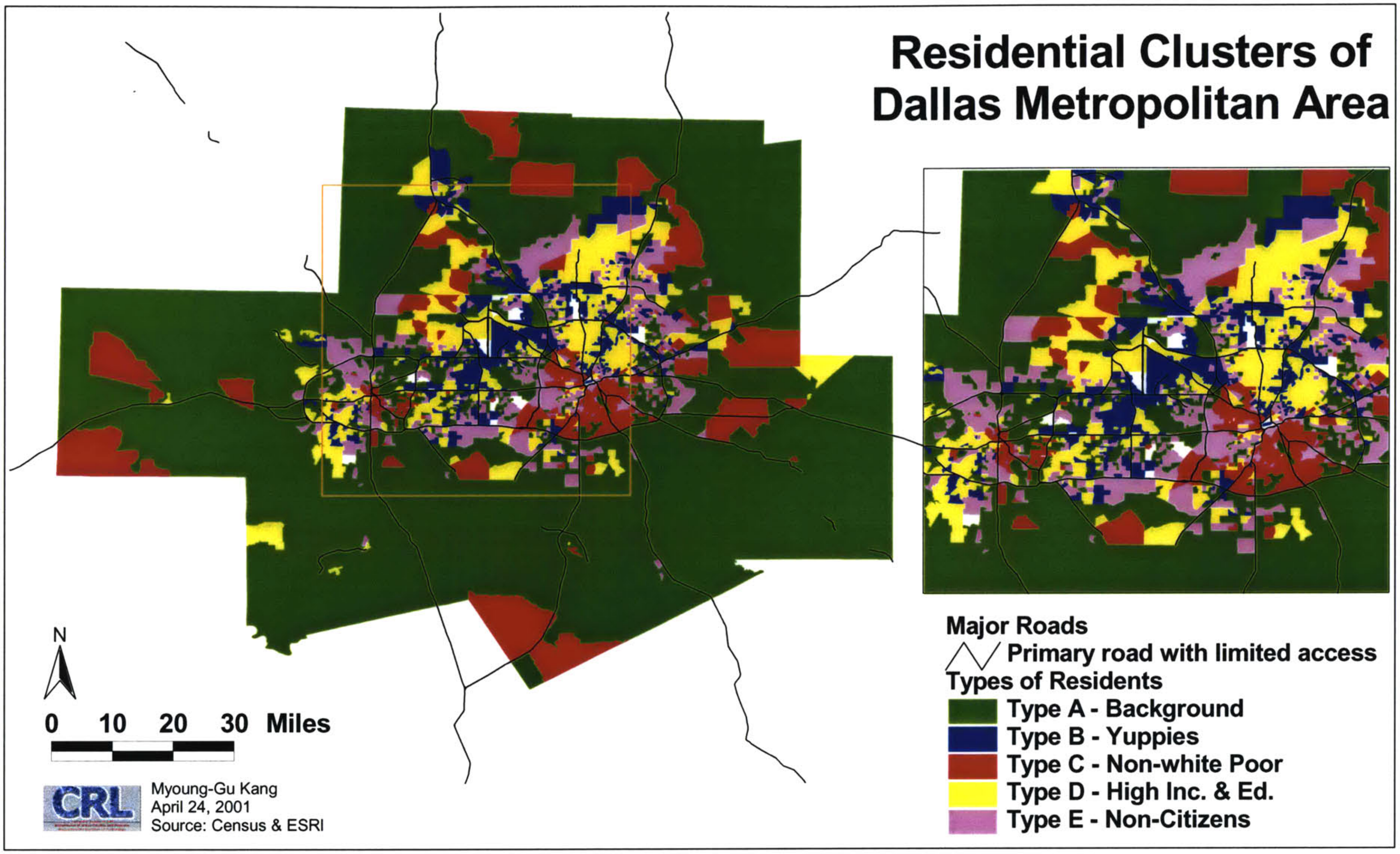
The third important factor is education and income. In the Dallas Metropolitan Area, there is a group of white citizen neighborhoods which are located in low density suburban area. The fourth factor is the percent of younger labor force. Finally, the percent citizens explains 13 percent of the variation in Dallas Metropolitan Area.

Through Cluster analysis, we can find five clusters in Dallas Metropolitan Area: general whites, non-citizens with children, young labor force with no children, high income residents, non-white citizens. As you see on the map, neighborhoods in the same group generally tend to agglomerate. The high income residents are found at the northern part of Dallas and non-white citizens live in the southern part of Dalls. The non-citizens seems to be filling the gap between them. Young workers without children tend to appreciate the access to the major transportation network; they live closer to major roads.

Table 3.6 Types of Neighborhoods (Dallas)

Type	Factor	Minimum	Maximum	Mean	Std. Deviation
A	Baseline Factor	-1.29033	1.87634	0.50753	0.425965
N = 1,126	Children Factor	-2.37854	2.25021	0.093043	0.625815
(32.1 %)	Income Factor	-2.26905	0.83423	-0.57341	0.395671
Background	Age Factor	-4.33009	2.16332	-0.33653	0.791451
	Citizen Factor	-1.71424	2.30495	0.420364	0.558369
B	Baseline Factor	-3.2353	1.33865	-0.19286	0.677868
N = 554	Children Factor	-3.22573	0.92134	-1.26434	0.96947
(15.8 %)	Income Factor	-2.05349	4.03954	0.043008	0.698138
Yuppies	Age Factor	-1.02361	2.86054	1.133238	0.592858
	Citizen Factor	-4.08883	1.41161	-0.14257	0.745804
C	Baseline Factor	-8.47196	0.52401	-1.87515	1.60156
N = 344	Children Factor	-2.53048	2.72893	0.379408	0.712503
(9.8 %)	Income Factor	-1.34349	1.24279	-0.33829	0.530167
Non-whites	Age Factor	-5.9533	1.7	-0.37062	0.951067
Poor	Citizen Factor	-0.63803	3.48027	0.904604	0.537501
D	Baseline Factor	-4.70221	1.46284	0.138577	0.610764
N = 652	Children Factor	-3.1742	2.1427	-0.22636	1.009677
(18.6 %)	Income Factor	-1.08903	7.92172	1.338267	1.261527
High Incomers	Age Factor	-7.18732	1.61761	-0.57155	1.05793
	Citizen Factor	-2.985	1.84889	0.11706	0.482845
E	Baseline Factor	-2.89482	1.27751	0.107992	0.680426
N = 834	Children Factor	-1.59022	2.7427	0.73471	0.582249
(23.8 %)	Income Factor	-1.41121	2.22043	-0.16109	0.630304
Non-citizens	Age Factor	-2.45766	2.47741	0.301272	0.68962
	Citizen Factor	-5.5286	1.59971	-0.93747	1.280647

* The shaded cells are selected by author to highlight the characteristics of each cluster.


Residential Clusters of Dallas Metropolitan Area
Major Roads
Primary road with limited access
Types of Residents
Type A - Background
Type B - Yuppies
Type C - Non-white Poor
Type D - High Inc. & Ed.
Type E - Non-Citizens
N
0 10 20 30 Miles
CRL
Myoung-Gu Kang
April 24, 2001
Source: Census & ESRI

4

Conclusion

In this thesis, I try to capture the socioeconomic topography of residents in metropolitan areas. As cities are evolving, the inner structures of metropolitan area are becoming more complex, different from the past, and creating new urban life. The booming emergence of subcenters, for example, changes people's lifestyle and also changes the people's location choice. Now, therefore, we need to look at the spatial structure of cities in more detailed, and review the theoretical models in order to capture the changing real world more correctly.

In addition, the dramatically increasing computation capacity – including GIS, Statistics, and Database Management Systems (DMBS) – and the available ample data set – including the U.S. Census information – make a more detailed examination of spatial pattern. Therefore, we can take a step forward beyond the existing urban models.

I adopt the factor analysis and the cluster analysis with four metropolitan areas including Boston, MA (Northeast), Chicago, IL (Midwest), San Francisco, CA (West), and Dallas, TX (South) out of each region. The factor analysis and the cluster analysis are sensitive to the initial data set. I use the 1990 Census blockgroup as a basic unit of my analysis because it is the finest exhaustive data set I have access to.

4.1 Summary of the Four Metropolitan Areas

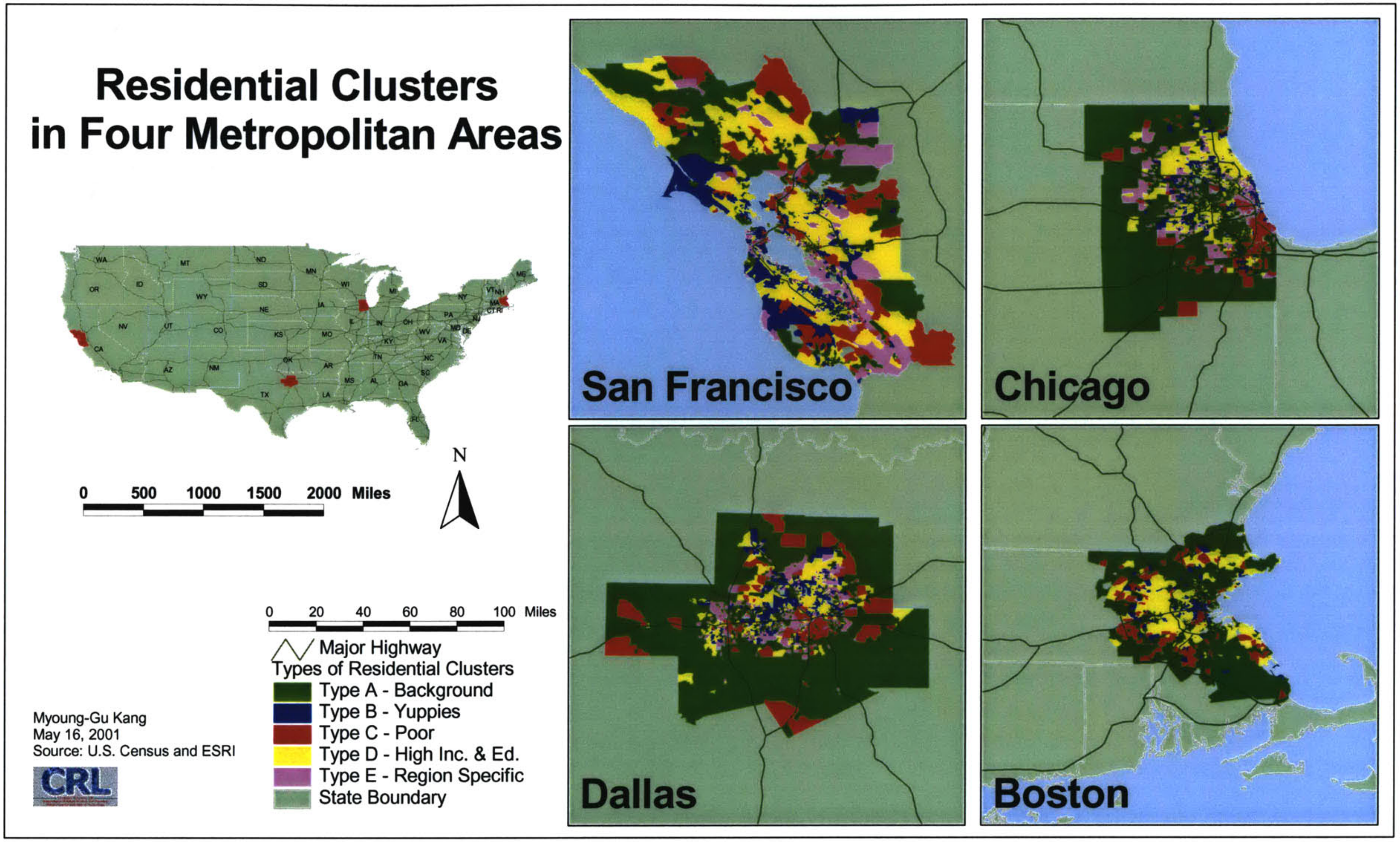
This research finds four common factors that account for above 70 percent of the variation in socioeconomic characteristics of local neighborhoods and that generate a spatial pattern with significant clustering: Baseline factor, children factor, income factor, and age factor. Average whites of the metropolitan areas live virtually any place which is not specialized yet by a certain type of residents. Non-whites, however, show the tendency of agglomeration. Number of children and large household size vary among neighborhoods and are generally not correlated with other variables. They are, however, related to location. The peculiar factor of San Francisco metropolitan area is the average travel time to work. In Chicago and Dallas, citizenship is an additional factor.

The variation of neighborhoods by income group clearly emerge. Generally, there are high-income neighborhoods in each metropolitan areas, which agglomerate at a preferable geographic location, not in downtown. Low income neighborhoods generally are near to downtown or job locations with high density. Incidentally, neighborhoods with larger number of low income households tend to have relatively more children than other neighborhoods.

Young workers, especially those with no children, tend to live basically near to their jobs in two ways. They tend to live physically near the employment and to the places where they can easily access transportation corridors. So, they are found near downtown or along the highways.

In addition, all the locational patterns do not follow the circumferential pattern in these four metropolitan areas. Near the downtown areas, up to about 15 miles away from center, we see vivid wedges of socio-economically different neighborhood clusters.

Travel time to work is not a factor differentiating residential clusters, except San Francisco. Interestingly, the downtown residents' travel time is not shorter than surburban residents' travel time. We might guess that walking to job or using public transportation to job may takes longer than driving in terms of time. In other words, poor people generally spend more time to commute even though they live in downtown.


Residential Clusters
in Four Metropolitan Areas
0 500 1000 1500 2000 Miles
N
0 20 40 60 80 100 Miles
Major Highway
Types of Residential Clusters
Type A - Background
Type B - Yuppies
Type C - Poor
Type D - High Inc. & Ed.
Type E - Region Specific
State Boundary
Myoung-Gu Kang
May 16, 2001
Source: U.S. Census and ESRI
CRL
San Francisco
Chicago
Dallas
Boston

By examining the geographical locations of the neighborhoods using GIS, we can see the same neighborhoods tend to be located together geographically as of 1990. Whether people choose their place close to the socioeconomically same group could be another longitudinal research topic. If it proves true, we can predict people's residential choice based on the existing residential neighborhood characteristics.

This alternative methodology is good for at least explanatory research for finding the key latent factors delineating neighborhoods in terms of socioeconomic characteristics. Then, we can stratify the neighborhoods based on the factors using cluster analysis. This methodology reduces the risk of the assumptions of the conventional regression analysis. We can also use the result of this analysis as an input to a further research including, for example, multivariate regression analysis.

4.2 Future Study

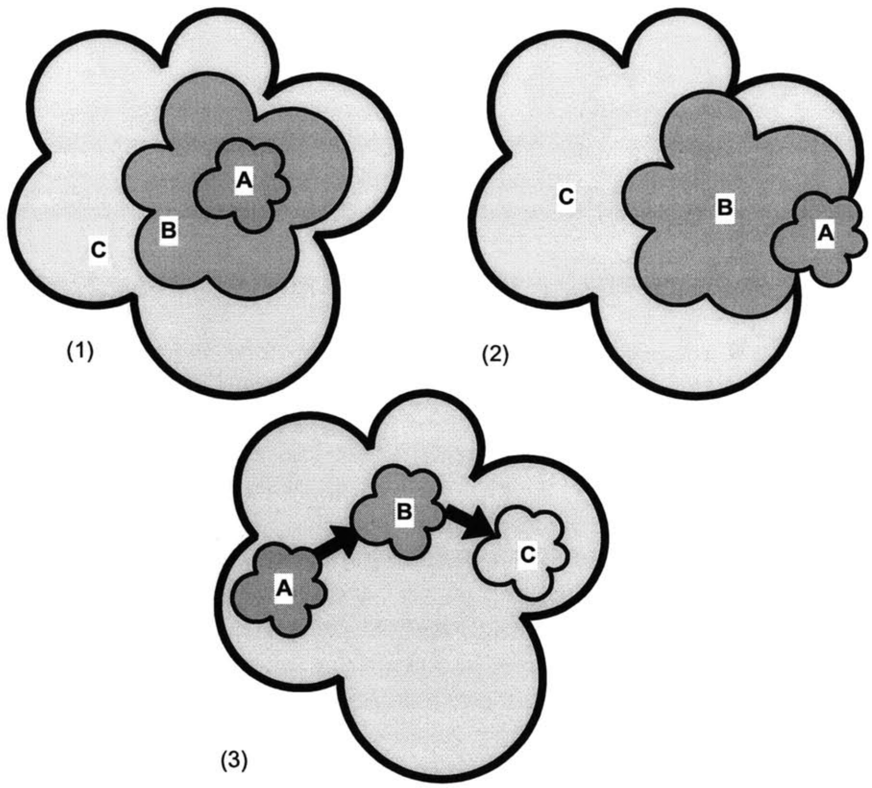
This kind of detailed research is critical for participatory planning as a consensus building process. Residents who are willing to participate in the planning need to know what the plans are, what the effect of the plans are, and so forth. The existing urban models are terse but too complicated and generalized for common residents to understand their community and plan. Furthermore, it is hard to answer the neighborhood level question.

For example, different neighborhoods requre different planning policy. A lot of non-white people live in the sothern part of the city of Chicago, and they spend more time to get their jobs than the rest metropolitan area, even though they live close to the Central Business District (CBD). Enhancing the inner-city transportation system could be more relevant in this area.

With panel data – time series as well as cross-sectional data – we can develop urban space-time simulation model for forecasting the future of the city: demographic pattern, land use, and so forth. The model also increases our capability of planning by exploring spatial patterns and correlationed factors that may suggest useful spatial dimension to begin modeling in theoretical models of urban spatial structure.

This model can be integrated with the Build-out scenarios and/or with the conventional economic models. We can then forecast the city through the time line, for one year, 3-5 years, and the full (eternal) equilibrium. Then, we can see the longitudinal pattern as well as cross sectional pattern at the same time.

Figure 4.2 Examples of people and land use change



(1) One type of residency (or land use) expands to outward (A -> B -> C).

(2) One type of residency (or land use) expands to outward and moves to West (A -> B -> C).

(3) One type of residency (or land use) moves to East keeping similar size (A -> B -> C).

Then, we can develop a space-time simulation model for planning; being able to work with the established economic models and helping people participate in the planning. New technology, in turn, make such simulation model possible.

APPENDIX A. GEOGRAPHIC AREAS REFERENCE[1]

Block Group (BG)

U.S. Census Bureau guidelines specify an ideal size for a BG of 400 housing units, with a minimum of 250, and a maximum of 550 housing units. The guidelines further required that BG boundaries follow clearly visible features, such as roads, rivers, and railroads.

A BG is a combination of census blocks[2] that is the finest grained subdivision of a census tract or block numbering area (BNA). (A county or its statistically equivalent entity contains either census tracts or BNAs; it can not contain both.)

---

[1] Source: U.S. Department of Commerce. 1994. Geographic Areas Reference Manual. http://www.census.gov/geo/www/garm.html

[2] Census blocks, the smallest geographic area for which the Bureau of the Census collects and tabulates decennial census data, are formed by streets, roads, railroads, streams and other bodies of water, other visible physical and cultural features, and the legal boundaries shown on Census Bureau maps.

Although most people intuitively think of census blocks as being rectangular or square, of about the same size, and occurring at regular intervals, as in many cities of the United States, census block configurations actually are quite different. Patterns, sizes, and shapes of census blocks vary within and between areas. Factors that influence the overall configuration of census blocks include topography, the size and spacing of water features, the land survey system, and the

A BG consists of all census blocks whose numbers begin with the same digit in a given census tract or BNA; for example, BG 3 includes all census blocks numbered in the 300s. The BG is the smallest geographic entity for which the decennial census tabulates and publishes sample data. It has now largely replaced the earlier enumeration district (ED) as a small-area geographic unit for purposes of data presentation.

Metropolitan Area (MA)[3]

The MA standards specify the step-by-step definition process by which the concept of a densely settled core area plus its suburbs becomes realized as individual MSAs, CMSAs, PMSAs, and NECMAs. Qualification of an MSA requires the presence of a city of 50,000 or more inhabitants, *or* a Census Bureau-defined UA (of at least 50,000 inhabitants) and a total population of at least 100,000 (75,000 in New England). The county or counties including the largest city in the core area of population become *central counties* of the MSA; so

---

extent, age, type, and density of urban and rural development.

[3] The collective term used for Federal metropolitan areas has varied over time, beginning with *standard metropolitan area (SMA)* in 1950, changing to *standard metropolitan statistical area (SMSA)* in 1959, to *metropolitan statistical area (MSA)* in 1983, and to *metropolitan area (MA)* in 1990.

does any adjacent county that has at least 50 percent of its population in the UA[4] surrounding the largest city. (In New England where all land is allocated to be in one or another town, the basic geographic unit for defining MSAs is the city or town rather than the county.)

Additional *outlying counties* are included in the MSA if they meet specified requirements of commuting to the central counties as well as other requirements of metropolitan character. The minimum level of commuting to central counties required to make a county eligible for consideration as an outlying county is 15 percent. In general, the lower the percentage of a county's resident workers commuting to the central counties, the more demanding the other requirements of metropolitan character the county must meet in order to qualify for inclusion. The measures of metropolitan character specified in the standards include required levels for the county's (1) population density; (2) percentage of population that is classified as urban; (3) percentage growth in population between the previous two decennial censuses; and (4) percentage of, or absolute number of, inhabitants within the UA that qualifies the MSA. Qualification of outlying cities and towns in New England is based on commuting and population density.

---

[4] Urbanized Areas (UAs): A UA is a continuously built-up area with a population of 50,000 or more. It comprises one or more places—*central place(s)*—and the adjacent densely settled surrounding area—*urban fringe*—consisting of other places and nonplace territory.

An area that meets the requirements for recognition as an MSA and also has a population of one million or more may be recognized as a CMSA if (1) separate component areas can be identified within the entire area by their meeting population and commuting criteria specified in the standards, and (2) local opinion indicates there is support for the component areas. If recognized, the component areas are designated PMSAs (and the entire area becomes a CMSA). If no PMSAs are recognized, the entire area is designated an MSA. (PMSAs, like the CMSAs that contain them, are composed of counties outside New England and cities and towns within New England.)

The collective term used for Federal metropolitan areas has varied over time, beginning with *standard metropolitan area (SMA)* in 1950, changing to *standard metropolitan statistical area (SMSA)* in 1959, to *metropolitan statistical area (MSA)* in 1983, and to *metropolitan area (MA)* in 1990.

Appendix B. A sample SPSS script for Boston

```
SET MXMEMORY 64000
   /MXCELLS 12000000
   /LENGTH 999999
   /WIDTH 132.

*******************************************.
GET TRANSLATE FILE = bos_sas.DBF
   /TYPE = DBF.

DESCRIPTIVES VARIABLES = pop_den, hh_den, wh_pct, kid_pct, old_pct, cit_pct, hhs_pct,
eng_pct, stu_pct, car_pct, time_avg, hied_pct, lab_pct, uemp_pct, hhi_med, wage_pct, inc_pc,
pov_pct, own_pct
   /STATISTICS = DEFAULT SKEWNESS.

CORRELATIONS VARIABLES = pop_den, hh_den, wh_pct, kid_pct, old_pct, cit_pct, hhs_pct,
eng_pct, stu_pct, car_pct, time_avg, hied_pct, lab_pct, uemp_pct, hhi_med, wage_pct, inc_pc,
pov_pct, own_pct.

FACTOR VARIABLES = pop_den, wh_pct, kid_pct, old_pct, cit_pct, hhs_pct, eng_pct, stu_pct,
car_pct, time_avg, hied_pct, lab_pct, uemp_pct, hhi_med, wage_pct, inc_pc, pov_pct, own_pct
   /METHOD = CORRELATION
   /PLOT = EIGEN ROTATION
   /CRITERIA = MINEIGEN(1.0) ITERATE(100)
   /EXTRACTION = PC
   /ROTATION = VARIMAX
   /SAVE = REG(ALL).

CLUSTER fac1_1, fac2_1, fac3_1, fac4_1
   /MEASURE = CORRELATION
   /METHOD = BAVERAGE
   /SAVE = CLUSTER(3, 6)
   /ID = bkg_key
   /PRINT = NONE
   /PLOT = NONE.

*************************************.
FREQUENCIES VARIABLES = clu4_1
   /PIECHART = PERCENT.

********************************************************.
TEMPORARY.
SELECT IF (clu4_1 = 1).
DESCRIPTIVES VARIABLES = fac1_1, fac2_1, fac3_1, fac4_1.

TEMPORARY.
SELECT IF (clu4_1 = 2).
DESCRIPTIVES VARIABLES = fac1_1, fac2_1, fac3_1, fac4_1.
```

```
TEMPORARY.
SELECT IF (clu4_1 = 3).
DESCRIPTIVES VARIABLES = fac1_1, fac2_1, fac3_1, fac4_1.

TEMPORARY.
SELECT IF (clu4_1 = 4).
DESCRIPTIVES VARIABLES = fac1_1, fac2_1, fac3_1, fac4_1.

*******************************************************.
SAVE OUTFILE = bos_1a.sav.

SAVE TRANSLATE
  /OUTFILE = bos_1a.dbf
  /TYPE = DB4
  /KEEP = ALL
  /REPLACE.
```

Appendix C. Statistical output for Boston

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std.	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
POP_DEN	3418	1.93424	50791.67	3679.884	4528.147	2.786	.042
HH_DEN	3418	.70335	28750.00	1474.740	2093.124	4.094	.042
WH_PCT	3418	.00000	100.00000	87.79765	21.16589	-2.716	.042
KID_PCT	3418	.00000	53.84615	21.45650	8.4642922	.041	.042
OLD_PCT	3418	.00000	74.27136	13.21695	8.1253392	1.728	.042
CIT_PCT	3418	.00000	100.00000	86.92303	11.73535	-2.003	.042
HHS_PCT	3418	.00000	100.00000	54.47186	15.71235	.168	.042
ENG_PCT	3418	.00000	100.00000	78.78759	14.19574	-2.030	.042
STU_PCT	3418	.00000	50.64935	15.74154	6.9567185	.189	.042
CAR_PCT	3418	.00000	100.00000	80.20602	19.31394	-1.486	.042
TIME_AVG	3418	1.00000	52.35294	24.07818	4.5301975	.437	.042
HIED_PCT	3418	.00000	100.00000	30.00180	19.35380	.850	.042
LAB_PCT	3418	3.44828	100.00000	69.14437	10.37315	-1.213	.042
UEMP_PCT	3418	.00000	86.04651	6.8782711	5.7527752	2.874	.042
HHI_MED	3418	4999.000	150001.0	42625.04	18024.40	1.252	.042
WAGE_PCT	3418	6.74157	100.00000	80.00046	11.46873	-1.343	.042
INC_PC	3418	2127.000	96975.00	18634.58	8349.946	2.334	.042
POV_PCT	3418	.00000	85.71429	8.6701534	10.53239	2.332	.042
OWN_PCT	3418	.00000	100.00000	63.96428	27.40745	-.516	.042
Valid N (listwise)	3418						

Correlations

Correlations

		POP_DEN	HH_DEN	WH_PCT	KID_PCT	OLD_PCT	CIT_PCT	HHS_PCT	ENG_PCT	STU_PCT	CAR_PCT	TIME_AVG	HIED_PCT	LAB_PCT	UEMP_PCT	HHI_MED	WAGE_PCT	INC_PC	POV_PCT	OWN_PCT
POP_DEN	Pearson Correlation	1.000	.942	-.408	-.172	-.092	-.460	.291	-.433	-.166	-.680	-.084	-.011	-.107	.212	-.423	-.171	-.214	.465	-.619
	Sig. (2-tailed)	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.530	.000	.000	.000	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
HH_DEN	Pearson Correlation	.942	1.000	-.287	-.271	-.026	-.370	.395	-.320	-.255	-.650	-.075	.066	-.060	.131	-.379	-.165	-.100	.394	-.569
	Sig. (2-tailed)	.000	.	.000	.000	.133	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
WH_PCT	Pearson Correlation	-.408	-.287	1.000	-.285	.209	.593	.090	.593	-.267	.450	-.124	.248	.221	-.422	.399	.197	.372	-.632	.515
	Sig. (2-tailed)	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
KID_PCT	Pearson Correlation	-.172	-.271	-.285	1.000	-.407	-.025	-.678	-.224	.893	.325	.134	-.309	.066	.244	.022	.071	-.268	.219	.124
	Sig. (2-tailed)	.000	.000	.000	.	.000	.139	.000	.000	.000	.000	.000	.000	.000	.000	.198	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
OLD_PCT	Pearson Correlation	-.092	-.026	.209	-.407	1.000	.025	.415	.095	-.368	.023	-.082	-.098	-.546	-.082	-.148	-.544	.067	-.108	-.001
	Sig. (2-tailed)	.000	.133	.000	.000	.	.152	.000	.000	.000	.180	.000	.000	.000	.000	.000	.000	.000	.000	.959
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
CIT_PCT	Pearson Correlation	-.460	-.370	.593	-.025	.025	1.000	-.057	.824	-.025	.442	.020	.129	.176	-.253	.353	.184	.263	-.472	.530
	Sig. (2-tailed)	.000	.000	.000	.139	.152	.	.001	.000	.151	.000	.242	.000	.000	.000	.000	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
HHS_PCT	Pearson Correlation	.291	.395	.090	-.678	.415	-.057	1.000	.032	-.645	-.415	-.130	.183	-.176	-.041	-.359	-.385	.121	.115	-.447
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.001	.	.062	.000	.000	.000	.000	.000	.018	.000	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
ENG_PCT	Pearson Correlation	-.433	-.320	.593	-.224	.095	.824	.032	1.000	-.144	.357	.103	.285	.224	-.378	.446	.250	.402	-.601	.573
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.062	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
STU_PCT	Pearson Correlation	-.166	-.255	-.267	.893	-.368	-.025	-.645	-.144	1.000	.286	.148	-.221	.037	.202	.095	.079	-.181	.178	.174
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.151	.000	.000	.	.000	.000	.000	.029	.000	.000	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
CAR_PCT	Pearson Correlation	-.680	-.650	.450	.325	.023	.442	-.415	.357	.286	1.000	.029	-.133	.182	-.178	.380	.215	.119	-.475	.658
	Sig. (2-tailed)	.000	.000	.000	.000	.180	.000	.000	.000	.000	.	.091	.000	.000	.000	.000	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
TIME_AVG	Pearson Correlation	-.084	-.075	-.124	.134	-.082	.020	-.130	.103	.148	.029	1.000	.070	.084	-.053	.134	.126	.078	-.068	.153
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.242	.000	.000	.000	.091	.	.000	.000	.002	.000	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
HIED_PCT	Pearson Correlation	-.011	.066	.248	-.309	-.098	.129	.183	.285	-.221	-.133	.070	1.000	.232	-.388	.599	.293	.723	-.316	.209
	Sig. (2-tailed)	.530	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
LAB_PCT	Pearson Correlation	-.107	-.060	.221	.066	-.546	.176	-.176	.224	.037	.182	.084	.232	1.000	-.231	.320	.678	.224	-.383	.259
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.029	.000	.000	.000	.	.000	.000	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
UEMP_PCT	Pearson Correlation	.212	.131	-.422	.244	-.082	-.253	-.041	-.378	.202	-.178	-.053	-.388	-.231	1.000	-.381	-.293	-.383	.519	-.373
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.018	.000	.000	.000	.002	.000	.000	.	.000	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
HHI_MED	Pearson Correlation	-.423	-.379	.399	.022	-.148	.353	-.359	.446	.095	.380	.134	.599	.320	-.381	1.000	.503	.746	-.585	.695
	Sig. (2-tailed)	.000	.000	.000	.198	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
WAGE_PCT	Pearson Correlation	-.171	-.165	.197	.071	-.544	.184	-.385	.250	.079	.215	.126	.293	.678	-.293	.503	1.000	.251	-.476	.387
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
INC_PC	Pearson Correlation	-.214	-.100	.372	-.268	.067	.263	.121	.402	-.181	.119	.078	.723	.224	-.383	.746	.251	1.000	-.448	.410
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
POV_PCT	Pearson Correlation	.465	.394	-.632	.219	-.108	-.472	.115	-.601	.178	-.475	-.068	-.316	-.383	.519	-.585	-.476	-.448	1.000	-.683
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418
OWN_PCT	Pearson Correlation	-.619	-.569	.515	.124	-.001	.530	-.447	.573	.174	.658	.153	.209	.259	-.373	.695	.387	.410	-.683	1.000
	Sig. (2-tailed)	.000	.000	.000	.000	.959	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.
	N	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418	3418

Factor Analysis

Communalities

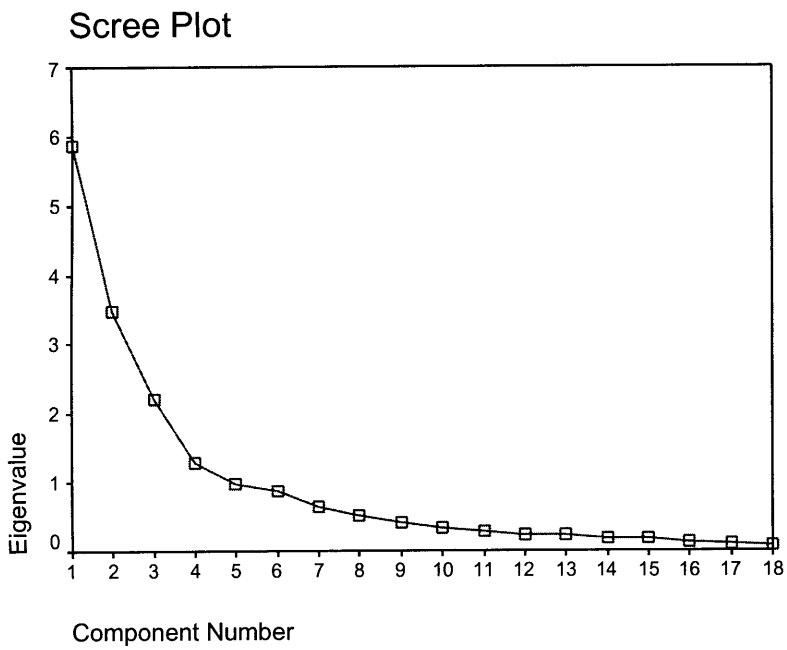
	Initial	Extraction
POP_DEN	1.000	.652
WH_PCT	1.000	.718
KID_PCT	1.000	.860
OLD_PCT	1.000	.794
CIT_PCT	1.000	.627
HHS_PCT	1.000	.762
ENG_PCT	1.000	.676
STU_PCT	1.000	.809
CAR_PCT	1.000	.747
TIME_AVG	1.000	.188
HIED_PCT	1.000	.806
LAB_PCT	1.000	.782
UEMP_PCT	1.000	.410
HHI_MED	1.000	.865
WAGE_PCT	1.000	.802
INC_PC	1.000	.815
POV_PCT	1.000	.732
OWN_PCT	1.000	.810

Extraction Method: Principal Component Analysis.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.871	32.618	32.618	5.871	32.618	32.618	4.681	26.006	26.006
2	3.486	19.364	51.982	3.486	19.364	51.982	3.216	17.868	43.874
3	2.203	12.241	64.224	2.203	12.241	64.224	2.642	14.678	58.552
4	1.293	7.184	71.408	1.293	7.184	71.408	2.314	12.855	71.408
5	.975	5.416	76.824						
6	.875	4.860	81.684						
7	.636	3.531	85.215						
8	.518	2.878	88.093						
9	.418	2.322	90.415						
10	.328	1.823	92.238						
11	.274	1.524	93.762						
12	.237	1.317	95.079						
13	.225	1.250	96.329						
14	.179	.996	97.324						
15	.168	.934	98.258						
16	.139	.773	99.031						
17	9.917E-02	.551	99.582						
18	7.524E-02	.418	100.000						

Extraction Method: Principal Component Analysis.



Component Matrix[a]

	Component			
	1	2	3	4
POP_DEN	-.627	-.222	.451	-7.27E-02
WH_PCT	.703	-.316	-.265	-.232
KID_PCT	-6.45E-02	.908	-9.98E-02	.143
OLD_PCT	-6.67E-02	-.569	-.615	.297
CIT_PCT	.674	-5.77E-02	-.344	-.225
HHS_PCT	-.257	-.830	1.293E-03	-8.33E-02
ENG_PCT	.751	-.218	-.212	-.139
STU_PCT	-1.56E-02	.863	-7.39E-02	.240
CAR_PCT	.609	.382	-.476	-6.63E-02
TIME_AVG	.126	.190	.136	.343
HIED_PCT	.464	-.368	.553	.386
LAB_PCT	.464	.211	.541	-.479
UEMP_PCT	-.564	.275	-.130	-8.33E-03
HHI_MED	.806	8.359E-02	.253	.379
WAGE_PCT	.561	.284	.552	-.318
INC_PC	.627	-.330	.314	.462
POV_PCT	-.832	.150	4.399E-02	.124
OWN_PCT	.840	.217	-.203	.127

Extraction Method: Principal Component Analysis.

a. 4 components extracted.

Rotated Component Matrix[a]

	Component			
	1	2	3	4
POP_DEN	-.731	-.315	-8.22E-02	.106
WH_PCT	.779	-.309	.110	6.615E-02
KID_PCT	-5.98E-02	.904	-.173	9.479E-02
OLD_PCT	.218	-.355	-6.69E-04	-.788
CIT_PCT	.786	-5.79E-02	7.917E-03	7.452E-02
HHS_PCT	-.171	-.813	-2.25E-03	-.269
ENG_PCT	.768	-.191	.203	9.114E-02
STU_PCT	-5.20E-02	.894	-5.96E-02	5.655E-02
CAR_PCT	.753	.419	-7.02E-02	-1.69E-04
TIME_AVG	-4.42E-02	.282	.326	-1.47E-02
HIED_PCT	2.698E-02	-.259	.839	.184
LAB_PCT	.199	-1.35E-02	.102	.856
UEMP_PCT	-.406	.245	-.396	-.169
HHI_MED	.454	.211	.754	.213
WAGE_PCT	.238	.112	.251	.818
INC_PC	.269	-.161	.846	3.257E-02
POV_PCT	-.743	.140	-.319	-.242
OWN_PCT	.767	.307	.343	9.465E-02

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

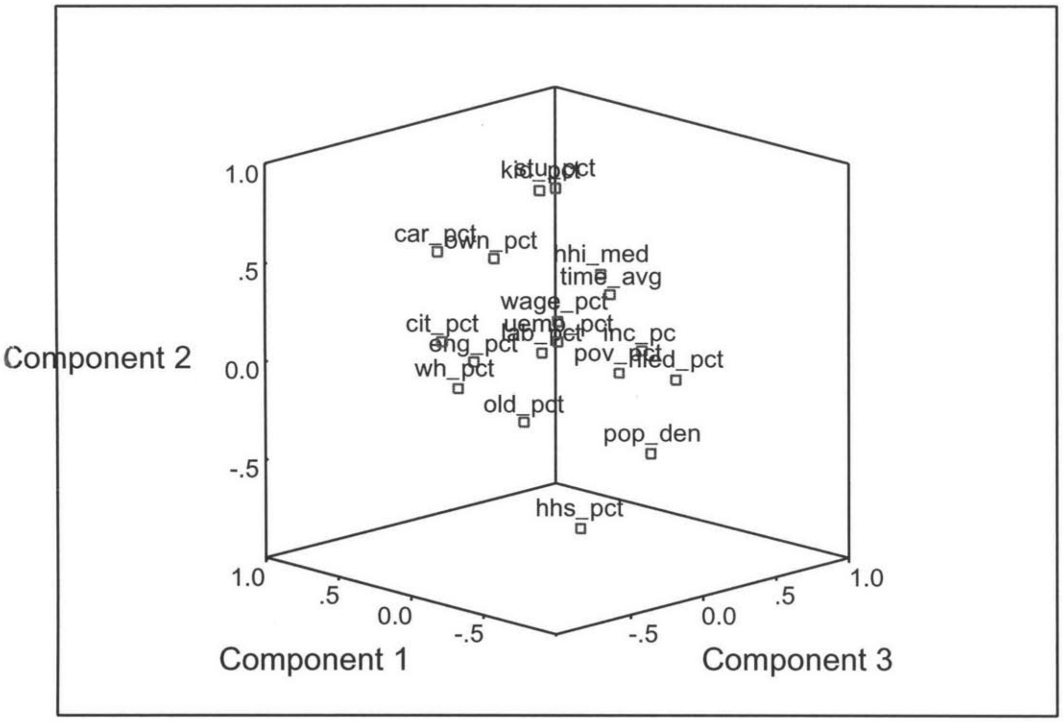
a. Rotation converged in 6 iterations.

Component Transformation Matrix

Component	1	2	3	4
1	.828	.044	.475	.293
2	-.031	.931	-.214	.294
3	-.517	-.124	.486	.693
4	-.212	.340	.702	-.589

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Component Plot in Rotated Space



Cluster

Case Processing Summary[a,b]

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
3418	100.0	0	.0	3418	100.0

a. Correlation between Vectors of Values used

b. Average Linkage (Between Groups)

Frequencies

Statistics

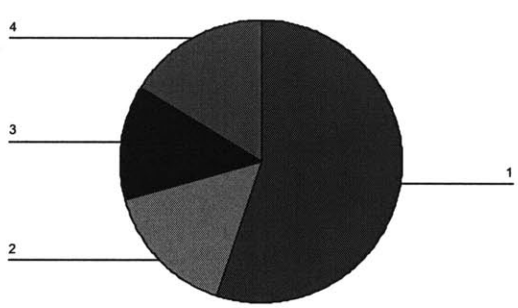
Average Linkage (Between Groups)

N	Valid	3418
	Missing	0

Average Linkage (Between Groups)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1881	55.0	55.0	55.0
	2	527	15.4	15.4	70.5
	3	454	13.3	13.3	83.7
	4	556	16.3	16.3	100.0
	Total	3418	100.0	100.0	

Average Linkage (Between Groups)



Type A

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	1881	-1.65725	1.77389	.5391403	.4317911
REGR factor score 2 for analysis 1	1881	-3.24035	2.78211	-4.0E-02	.7371879
REGR factor score 3 for analysis 1	1881	-3.36684	1.35712	-.4846772	.5036457
REGR factor score 4 for analysis 1	1881	-5.28612	2.86687	5.10E-02	.8919396
Valid N (listwise)	1881				

Type B

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	527	-4.23790	.82914	-.9561663	.8502623
REGR factor score 2 for analysis 1	527	-3.73016	1.27724	-.9994230	.9772320
REGR factor score 3 for analysis 1	527	-1.21592	3.56374	.4200407	.7253910
REGR factor score 4 for analysis 1	527	-1.31324	3.19455	.8050371	.6173182
Valid N (listwise)	527				

Type C

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	454	-5.23018	.76252	-1.39824	1.1513438
REGR factor score 2 for analysis 1	454	-.69366	3.84356	1.2546407	.7641717
REGR factor score 3 for analysis 1	454	-2.15750	2.38877	-.2111196	.8621008
REGR factor score 4 for analysis 1	454	-3.62040	2.44578	-.2428194	.9576004
Valid N (listwise)	454				

Type D

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	556	-2.41803	1.11860	.2240604	.5051806
REGR factor score 2 for analysis 1	556	-2.85245	2.28831	5.66E-02	.8303874
REGR factor score 3 for analysis 1	556	-1.06391	5.95639	1.4139652	1.0904758
REGR factor score 4 for analysis 1	556	-6.57086	1.09179	-.7373552	1.0654339
Valid N (listwise)	556				

Appendix D. Statistical output for Chicago

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std.	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
POP_DEN	6222	1.09250	44354.43	3923.842	3937.554	2.339	.031
HH_DEN	6222	.36416	29566.67	1437.002	1725.393	4.864	.031
WH_PCT	6222	.00000	100.00000	70.69488	35.59442	-1.093	.031
KID_PCT	6222	.00000	74.80916	25.58772	9.0230011	.045	.031
OLD_PCT	6222	.00000	100.00000	12.22726	8.8266194	2.008	.031
CIT_PCT	6222	.00000	100.00000	86.34224	14.88683	-1.798	.031
HHS_PCT	6222	.00000	100.00000	50.64885	16.80597	.197	.031
ENG_PCT	6222	.00000	100.00000	75.12777	18.49222	-1.641	.031
STU_PCT	6222	.00000	65.21739	18.96485	7.5222487	.299	.031
CAR_PCT	6222	.00000	100.00000	79.05242	17.99943	-1.324	.031
TIME_AVG	6222	1.00000	99.00000	28.75844	6.3645652	1.094	.031
HIED_PCT	6222	.00000	100.00000	22.30431	18.71267	1.167	.031
LAB_PCT	6222	1.97260	100.00000	67.48865	11.77196	-.847	.031
UEMP_PCT	6222	.00000	83.55263	7.8223303	9.0686227	2.564	.031
HHI_MED	6222	4999.000	150001.0	38900.75	19242.94	1.671	.031
WAGE_PCT	6222	7.29927	100.00000	80.22729	12.91634	-1.311	.031
INC_PC	6222	491.00000	127543.0	16608.44	10042.37	3.108	.031
POV_PCT	6222	.00000	100.00000	11.15343	15.12049	2.258	.031
OWN_PCT	6222	.00000	100.00000	69.20114	27.50460	-.791	.031
Valid N (listwise)	6222						

Correlations

Correlations

		POP_DEN	HH_DEN	WH_PCT	KID_PCT	OLD_PCT	CIT_PCT	HHS_PCT	ENG_PCT	STU_PCT	CAR_PCT	TIME_AVG	HIED_PCT	LAB_PCT	UEMP_PCT	HHI_MED	WAGE_PCT	INC_PC	POV_PCT	OWN_PCT
POP_DEN	Pearson Correlation	1.000	.907	-.420	.067	-.070	-.357	.062	-.399	.050	-.576	.246	-.148	-.173	.353	-.426	-.236	-.250	.444	-.547
	Sig. (2-tailed)	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
HH_DEN	Pearson Correlation	.907	1.000	-.258	-.149	.038	-.255	.287	-.256	-.147	-.534	.152	.007	-.094	.198	-.339	-.212	-.072	.296	-.499
	Sig. (2-tailed)	.000	.	.000	.000	.003	.000	.000	.000	.000	.000	.000	.570	.000	.000	.000	.000	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
WH_PCT	Pearson Correlation	-.420	-.258	1.000	-.341	.183	-.029	.224	.057	-.346	.565	-.466	.394	.352	-.702	.518	.312	.476	-.702	.517
	Sig. (2-tailed)	.000	.000	.	.000	.000	.020	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
KID_PCT	Pearson Correlation	.067	-.149	-.341	1.000	-.543	-.019	-.717	-.156	.888	.006	.199	-.267	-.026	.354	-.102	.044	-.354	.379	-.069
	Sig. (2-tailed)	.000	.000	.000	.	.000	.139	.000	.000	.000	.651	.000	.000	.044	.000	.000	.000	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
OLD_PCT	Pearson Correlation	-.070	.038	.183	-.543	1.000	.028	.508	.084	-.472	-.003	-.088	-.062	-.519	-.110	-.105	-.531	.075	-.118	.085
	Sig. (2-tailed)	.000	.003	.000	.000	.	.026	.000	.000	.000	.839	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
CIT_PCT	Pearson Correlation	-.357	-.255	-.029	-.019	.028	1.000	.049	.901	-.002	.092	.097	.081	-.064	.083	.129	-.047	.114	-.027	.228
	Sig. (2-tailed)	.000	.000	.020	.139	.026	.	.000	.000	.888	.000	.000	.000	.000	.000	.000	.000	.000	.034	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
HHS_PCT	Pearson Correlation	.062	.287	.224	-.717	.508	.049	1.000	.127	-.688	-.136	-.188	.213	-.072	-.200	-.164	-.302	.224	-.112	-.244
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
ENG_PCT	Pearson Correlation	-.399	-.256	.057	-.156	.084	.901	.127	1.000	-.080	.138	.058	.200	-.003	-.032	.236	.012	.233	-.141	.294
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.790	.012	.000	.337	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
STU_PCT	Pearson Correlation	.050	-.147	-.346	.888	-.472	-.002	-.688	-.080	1.000	-.015	.216	-.226	-.052	.334	-.035	.034	-.278	.349	-.015
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.888	.000	.000	.	.233	.000	.000	.000	.000	.006	.008	.000	.000	.228
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
CAR_PCT	Pearson Correlation	-.576	-.534	.565	.006	-.003	.092	-.136	.138	-.015	1.000	-.382	.039	.328	-.509	.404	.371	.195	-.603	.622
	Sig. (2-tailed)	.000	.000	.000	.651	.839	.000	.000	.000	.233	.	.000	.002	.000	.000	.000	.000	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
TIME_AVG	Pearson Correlation	.246	.152	-.466	.199	-.088	.097	-.188	.058	.216	-.382	1.000	-.135	-.196	.359	-.136	-.147	-.156	.309	-.122
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
HIED_PCT	Pearson Correlation	-.148	.007	.394	-.267	-.062	.081	.213	.200	-.226	.039	-.135	1.000	.339	-.432	.682	.338	.753	-.395	.242
	Sig. (2-tailed)	.000	.570	.000	.000	.000	.000	.000	.000	.000	.002	.000	.	.000	.000	.000	.000	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
LAB_PCT	Pearson Correlation	-.173	-.094	.352	-.026	-.519	-.064	-.072	-.003	-.052	.328	-.196	.339	1.000	-.427	.360	.767	.262	-.523	.218
	Sig. (2-tailed)	.000	.000	.000	.044	.000	.000	.000	.790	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
UEMP_PCT	Pearson Correlation	.353	.198	-.702	.354	-.110	.083	-.200	-.032	.334	-.509	.359	-.432	-.427	1.000	-.518	-.455	-.465	.754	-.495
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.012	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
HHI_MED	Pearson Correlation	-.426	-.339	.518	-.102	-.105	.129	-.164	.236	-.035	.404	-.136	.682	.360	-.518	1.000	.496	.820	-.598	.633
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.006	.000	.000	.000	.000	.000	.	.000	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
WAGE_PCT	Pearson Correlation	-.236	-.212	.312	.044	-.531	-.047	-.302	.012	.034	.371	-.147	.338	.767	-.455	.496	1.000	.281	-.553	.359
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.337	.008	.000	.000	.000	.000	.000	.000	.	.000	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
INC_PC	Pearson Correlation	-.250	-.072	.476	-.354	.075	.114	.224	.233	-.278	.195	-.156	.753	.262	-.465	.820	.281	1.000	-.480	.379
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
POV_PCT	Pearson Correlation	.444	.296	-.702	.379	-.118	-.027	-.112	-.141	.349	-.603	.309	-.395	-.523	.754	-.598	-.553	-.480	1.000	-.675
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.034	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222
OWN_PCT	Pearson Correlation	-.547	-.499	.517	-.069	.085	.228	-.244	.294	-.015	.622	-.122	.242	.218	-.495	.633	.359	.379	-.675	1.000
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.228	.000	.000	.000	.000	.000	.000	.000	.000	.000	.
	N	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222	6222

Factor Analysis

Communalities

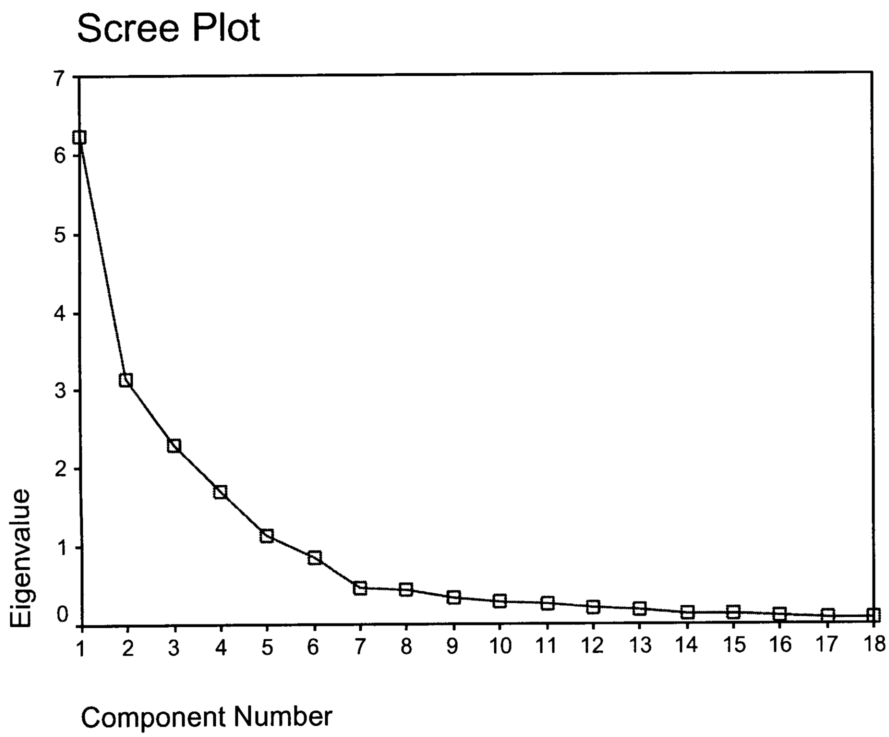
	Initial	Extraction
POP_DEN	1.000	.657
WH_PCT	1.000	.742
KID_PCT	1.000	.857
OLD_PCT	1.000	.855
CIT_PCT	1.000	.928
HHS_PCT	1.000	.823
ENG_PCT	1.000	.934
STU_PCT	1.000	.826
CAR_PCT	1.000	.794
TIME_AVG	1.000	.411
HIED_PCT	1.000	.816
LAB_PCT	1.000	.873
UEMP_PCT	1.000	.726
HHI_MED	1.000	.925
WAGE_PCT	1.000	.838
INC_PC	1.000	.868
POV_PCT	1.000	.811
OWN_PCT	1.000	.775

Extraction Method: Principal Component Analysis.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.028	33.490	33.490	6.028	33.490	33.490	4.005	22.251	22.251
2	3.296	18.313	51.804	3.296	18.313	51.804	3.163	17.573	39.824
3	2.229	12.381	64.185	2.229	12.381	64.185	2.888	16.042	55.866
4	1.710	9.499	73.684	1.710	9.499	73.684	2.246	12.477	68.343
5	1.195	6.639	80.323	1.195	6.639	80.323	2.156	11.980	80.323
6	.852	4.732	85.055						
7	.461	2.560	87.615						
8	.433	2.405	90.020						
9	.330	1.831	91.851						
10	.278	1.547	93.397						
11	.255	1.417	94.814						
12	.217	1.205	96.020						
13	.197	1.093	97.112						
14	.148	.819	97.932						
15	.131	.728	98.660						
16	.108	.603	99.262						
17	6.761E-02	.376	99.638						
18	6.517E-02	.362	100.000						

Extraction Method: Principal Component Analysis.



Component Matrix[a]

	Component				
	1	2	3	4	5
POP_DEN	-.578	-.138	-.429	.344	4.637E-02
WH_PCT	.801	-.122	-.155	-.238	7.501E-02
KID_PCT	-.412	.801	9.750E-02	-9.73E-02	.164
OLD_PCT	6.587E-02	-.775	.174	-.329	.333
CIT_PCT	.152	-1.76E-02	.887	.143	-.311
HHS_PCT	.159	-.851	-8.08E-02	.131	-.224
ENG_PCT	.285	-8.18E-02	.853	.189	-.287
STU_PCT	-.376	.774	.158	-5.85E-02	.240
CAR_PCT	.641	.260	7.940E-02	-.554	-4.79E-02
TIME_AVG	-.411	.128	.248	.365	.176
HIED_PCT	.613	-5.84E-02	-4.80E-02	.624	.213
LAB_PCT	.541	.406	-.350	.220	-.495
UEMP_PCT	-.805	6.494E-02	.241	.123	-1.17E-02
HHI_MED	.785	.252	.106	.295	.383
WAGE_PCT	.573	.544	-.271	.188	-.324
INC_PC	.708	-.118	4.358E-02	.479	.347
POV_PCT	-.879	-9.22E-03	.112	.151	4.744E-02
OWN_PCT	.701	.245	.289	-.272	.255

Extraction Method: Principal Component Analysis.

a. 5 components extracted.

Rotated Component Matrix[a]

	Component				
	1	2	3	4	5
POP_DEN	-.675	-7.22E-02	-9.03E-02	1.092E-02	-.434
WH_PCT	.727	-.296	.329	9.151E-02	-9.42E-02
KID_PCT	-.129	.892	-.187	9.127E-02	-4.38E-02
OLD_PCT	.195	-.506	1.535E-03	-.749	-2.87E-04
CIT_PCT	1.703E-02	-3.60E-03	2.528E-02	-2.96E-02	.962
HHS_PCT	-8.45E-02	-.890	3.404E-02	-.140	5.112E-02
ENG_PCT	7.556E-02	-9.71E-02	.139	-1.22E-02	.948
STU_PCT	-.128	.894	-9.99E-02	2.827E-02	-5.07E-03
CAR_PCT	.870	.113	-1.40E-02	.122	9.170E-02
TIME_AVG	-.517	.294	.105	-.140	.162
HIED_PCT	2.396E-02	-.210	.847	.225	6.121E-02
LAB_PCT	.265	-5.59E-02	.169	.877	-3.48E-02
UEMP_PCT	-.655	.299	-.369	-.232	.131
HHI_MED	.418	.140	.835	.143	.113
WAGE_PCT	.335	.135	.263	.799	-2.39E-02
INC_PC	.190	-.206	.882	5.191E-02	8.756E-02
POV_PCT	-.737	.247	-.364	-.272	-1.30E-02
OWN_PCT	.738	.204	.366	-5.31E-02	.226

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

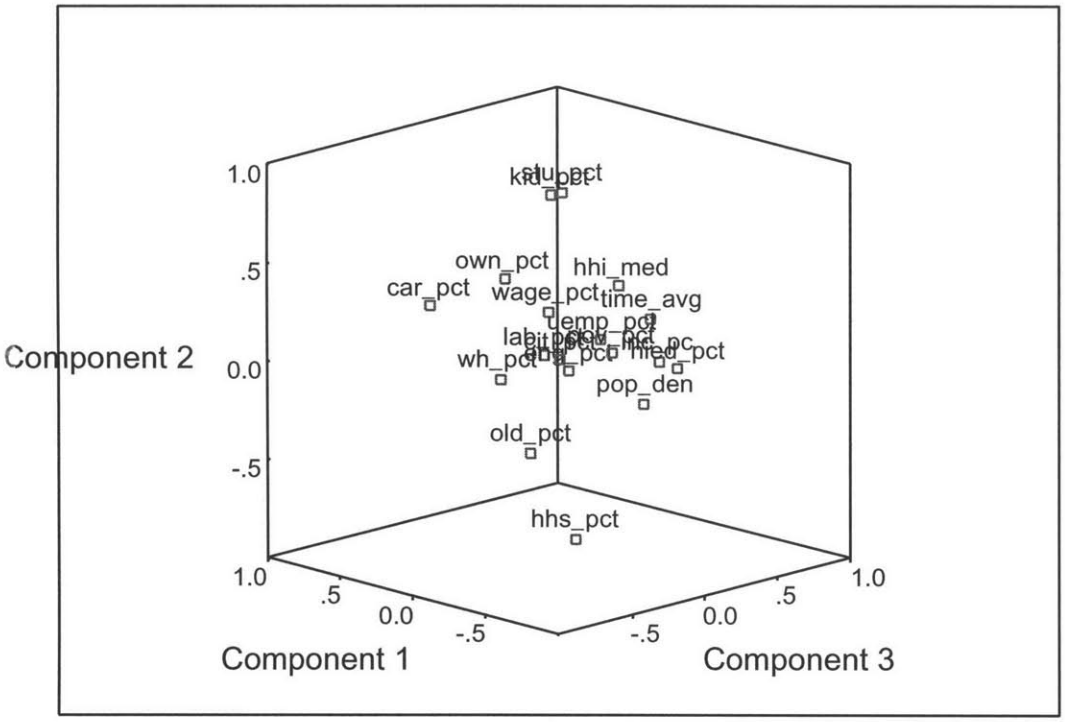
a. Rotation converged in 8 iterations.

Component Transformation Matrix

Component	1	2	3	4	5
1	.726	-.263	.555	.279	.138
2	.119	.869	.016	.479	-.003
3	.021	.199	.022	-.361	.911
4	-.676	-.102	.631	.333	.154
5	.040	.354	.542	-.672	-.358

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Component Plot in Rotated Space



Cluster

Case Processing Summary[a,b]

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
6222	100.0	0	.0	6222	100.0

a. Correlation between Vectors of Values used

b. Average Linkage (Between Groups)

Frequencies

Statistics

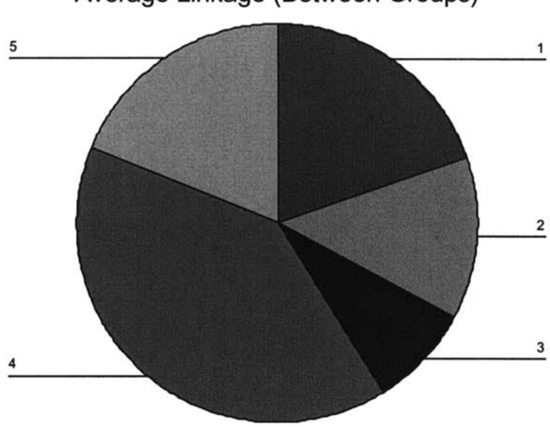
Average Linkage (Between Groups)

N	Valid	6222
	Missing	0

Average Linkage (Between Groups)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1226	19.7	19.7	19.7
	2	810	13.0	13.0	32.7
	3	530	8.5	8.5	41.2
	4	2479	39.8	39.8	81.1
	5	1177	18.9	18.9	100.0
	Total	6222	100.0	100.0	

Average Linkage (Between Groups)



Type A

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	1226	-3.95917	1.46963	-5.5E-02	.7518217
REGR factor score 2 for analysis 1	1226	-1.96958	4.37420	.4391363	.8253933
REGR factor score 3 for analysis 1	1226	-1.62255	3.97953	-2.7E-02	.6400483
REGR factor score 4 for analysis 1	1226	-2.58550	2.84962	.1200633	.7212294
REGR factor score 5 for analysis 1	1226	-4.63651	.81020	-1.38563	1.1293631
Valid N (listwise)	1226				

Type B

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	810	-3.41195	1.06636	9.56E-02	.5831813
REGR factor score 2 for analysis 1	810	-3.36093	3.36707	-5.7E-02	1.0388082
REGR factor score 3 for analysis 1	810	-.54496	7.76026	1.7391680	1.3038230
REGR factor score 4 for analysis 1	810	-4.05503	1.92422	-.3197088	.7645532
REGR factor score 5 for analysis 1	810	-3.00565	1.92113	.1287498	.4942217
Valid N (listwise)	810				

Type C

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	530	-3.73974	.64733	-.8005700	.9591462
REGR factor score 2 for analysis 1	530	-3.97427	.63449	-1.35638	.9420235
REGR factor score 3 for analysis 1	530	-1.36519	3.01809	.3992916	.7546656
REGR factor score 4 for analysis 1	530	-.38217	3.53079	1.3131145	.6523073
REGR factor score 5 for analysis 1	530	-2.44180	1.60906	8.70E-02	.6488248
Valid N (listwise)	530				

Type D

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	2479	-1.57188	2.37776	.7316201	.3508417
REGR factor score 2 for analysis 1	2479	-4.84021	2.98234	-.1983755	.7480756
REGR factor score 3 for analysis 1	2479	-2.26917	1.65396	-.3906577	.4636720
REGR factor score 4 for analysis 1	2479	-7.22555	3.04682	-7.0E-02	1.0085121
REGR factor score 5 for analysis 1	2479	-1.98102	1.51219	.2043230	.5166674
Valid N (listwise)	2479				

Type E

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	1177	-4.82193	1.58859	-1.18914	.9757784
REGR factor score 2 for analysis 1	1177	-2.88899	3.48838	.6103573	.8782626
REGR factor score 3 for analysis 1	1177	-1.95375	1.31278	-.5259156	.4796976
REGR factor score 4 for analysis 1	1177	-6.31010	2.08788	-.3495920	1.0081423
REGR factor score 5 for analysis 1	1177	-1.62786	2.28556	.8851822	.3909326
Valid N (listwise)	1177				

Appendix E. Statistical output for San Francisco

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std.	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
POP_DEN	4676	.02037	5677.358	364.0362	402.8071	3.773	.036
HH_DEN	4676	.00692	2850.685	145.2065	186.5122	4.662	.036
WH_PCT	4676	.00000	100.00000	70.99109	25.08235	-1.081	.036
KID_PCT	4676	.00000	63.63636	21.85520	8.7293722	-.143	.036
OLD_PCT	4676	.00000	100.00000	12.53599	9.8054592	2.922	.036
CIT_PCT	4676	.00000	100.00000	78.31525	16.14834	-1.341	.036
HHS_PCT	4676	.00000	100.00000	56.75344	17.86749	-.041	.036
ENG_PCT	4676	.00000	100.00000	69.97415	17.65834	-1.100	.036
STU_PCT	4676	.00000	63.63636	16.25425	7.0456285	.061	.036
CAR_PCT	4676	.00000	100.00000	83.09352	17.18913	-1.838	.036
TIME_AVG	4676	1.00000	64.14706	25.28134	5.2767400	.463	.036
HIED_PCT	4676	.00000	100.00000	30.52564	18.90769	.630	.036
LAB_PCT	4676	.13404	100.00000	68.57957	11.66726	-1.338	.036
UEMP_PCT	4676	.00000	59.28144	5.4979584	5.3752391	2.574	.036
HHI_MED	4676	4999.000	150001.0	44776.92	20417.29	1.437	.036
WAGE_PCT	4676	6.10583	100.00000	79.59059	12.35386	-1.252	.036
INC_PC	4676	2609.000	138397.0	20439.89	11022.80	2.528	.036
POV_PCT	4676	.00000	100.00000	8.6929346	9.2732848	2.268	.036
OWN_PCT	4676	.00000	100.00000	60.78269	26.96330	-.548	.036
Valid N (listwise)	4676						

Cc ɜlations

Correlations

		POP_DEN	HH_DEN	WH_PCT	KID_PCT	OLD_PCT	CIT_PCT	HHS_PCT	ENG_PCT	STU_PCT	CAR_PCT	TIME_AVG	HIED_PCT	LAB_PCT	UEMP_PCT	HHI_MED	WAGE_PCT	INC_PC	POV_PCT	OWN_PCT
POI	Pearson Correlation	1.000	.915	-.407	-.145	-.011	-.423	.130	-.408	-.122	-.630	.008	-.069	-.053	.171	-.311	-.078	-.197	.292	-.445
	Sig. (2-tailed)	.	.000	.000	.000	.444	.000	.000	.000	.000	.000	.582	.000	.000	.000	.000	.000	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
HH_	Pearson Correlation	.915	1.000	-.274	-.284	.073	-.285	.307	-.255	-.256	-.637	-.010	.039	-.016	.101	-.295	-.134	-.082	.232	-.460
	Sig. (2-tailed)	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.490	.008	.271	.000	.000	.000	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
WH	Pearson Correlation	-.407	-.274	1.000	-.288	.142	.545	.259	.589	-.270	.384	-.107	.399	.204	-.493	.398	.099	.476	-.532	.336
	Sig. (2-tailed)	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
KID	Pearson Correlation	-.145	-.284	-.288	1.000	-.484	-.145	-.754	-.265	.897	.322	.172	-.379	.107	.243	.002	.230	-.350	.164	.133
	Sig. (2-tailed)	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.897	.000	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
OLI	Pearson Correlation	-.011	.073	.142	-.484	1.000	.110	.478	.173	-.424	-.097	-.096	.032	-.654	-.077	-.105	-.698	.143	-.087	.092
	Sig. (2-tailed)	.444	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.029	.000	.000	.000	.000	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
CIT	Pearson Correlation	-.423	-.285	.545	-.145	.110	1.000	.275	.918	-.154	.318	-.020	.216	.000	-.174	.177	-.072	.289	-.257	.274
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.171	.000	.973	.000	.000	.000	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
HH	Pearson Correlation	.130	.307	.259	-.754	.478	.275	1.000	.348	-.723	-.303	-.129	.267	-.160	-.110	-.244	-.421	.239	.015	-.306
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.313	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
EN	Pearson Correlation	-.408	-.255	.589	-.265	.173	.918	.348	1.000	-.218	.281	-.019	.344	.015	-.270	.254	-.085	.393	-.316	.307
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.203	.000	.302	.000	.000	.000	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
STI	Pearson Correlation	-.122	-.256	-.270	.897	-.424	-.154	-.723	-.218	1.000	.263	.159	-.296	.072	.198	.079	.208	-.269	.137	.193
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
CA	Pearson Correlation	-.630	-.637	.384	.322	-.097	.318	-.303	.281	.263	1.000	-.046	-.046	.162	-.228	.362	.229	.148	-.415	.516
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.002	.002	.000	.000	.000	.000	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
TIN	Pearson Correlation	.008	-.010	-.107	.172	-.096	-.020	-.129	-.019	.159	-.046	1.000	-.035	.021	.050	.035	.028	-.035	.012	.174
	Sig. (2-tailed)	.582	.490	.000	.000	.000	.171	.000	.203	.000	.002	.	.016	.155	.001	.016	.058	.015	.417	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
HIE	Pearson Correlation	-.069	.039	.399	-.379	.032	.216	.267	.344	-.296	-.046	-.035	1.000	.226	-.414	.607	.136	.718	-.362	.259
	Sig. (2-tailed)	.000	.008	.000	.000	.029	.000	.000	.000	.000	.002	.016	.	.000	.000	.000	.000	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
LA	Pearson Correlation	-.053	-.016	.204	.107	-.654	.000	-.160	.015	.072	.162	.021	.226	1.000	-.227	.228	.720	.139	-.283	.024
	Sig. (2-tailed)	.000	.271	.000	.000	.000	.973	.000	.302	.000	.000	.155	.000	.	.000	.000	.000	.000	.000	.106
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
UE	Pearson Correlation	.171	.101	-.493	.243	-.077	-.174	-.110	-.270	.198	-.228	.050	-.414	-.227	1.000	-.429	-.204	-.402	.533	-.318
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001	.000	.000	.	.000	.000	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
HH	Pearson Correlation	-.311	-.295	.398	.002	-.105	.177	-.244	.254	.079	.362	.035	.607	.228	-.429	1.000	.349	.757	-.556	.634
	Sig. (2-tailed)	.000	.000	.000	.897	.000	.000	.000	.000	.000	.000	.016	.000	.000	.000	.	.000	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
WA	Pearson Correlation	-.078	-.134	.099	.230	-.698	-.072	-.421	-.085	.208	.229	.028	.136	.720	-.204	.349	1.000	.055	-.310	.110
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.058	.000	.000	.000	.000	.	.000	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
IN	Pearson Correlation	-.197	-.082	.476	-.350	.143	.289	.239	.393	-.269	.148	-.035	.718	.139	-.402	.757	.055	1.000	-.450	.390
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.015	.000	.000	.000	.000	.000	.	.000	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
PC	Pearson Correlation	.292	.232	-.532	.164	-.087	-.257	.015	-.316	.137	-.415	.012	-.362	-.283	.533	-.556	-.310	-.450	1.000	-.542
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.313	.000	.000	.000	.417	.000	.000	.000	.000	.000	.000	.	.000
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676
OV	Pearson Correlation	-.445	-.460	.336	.133	.092	.274	-.306	.307	.193	.516	.174	.259	.024	-.318	.634	.110	.390	-.542	1.000
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.106	.000	.000	.000	.000	.000	.
	N	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676	4676

Factor Analysis

Communalities

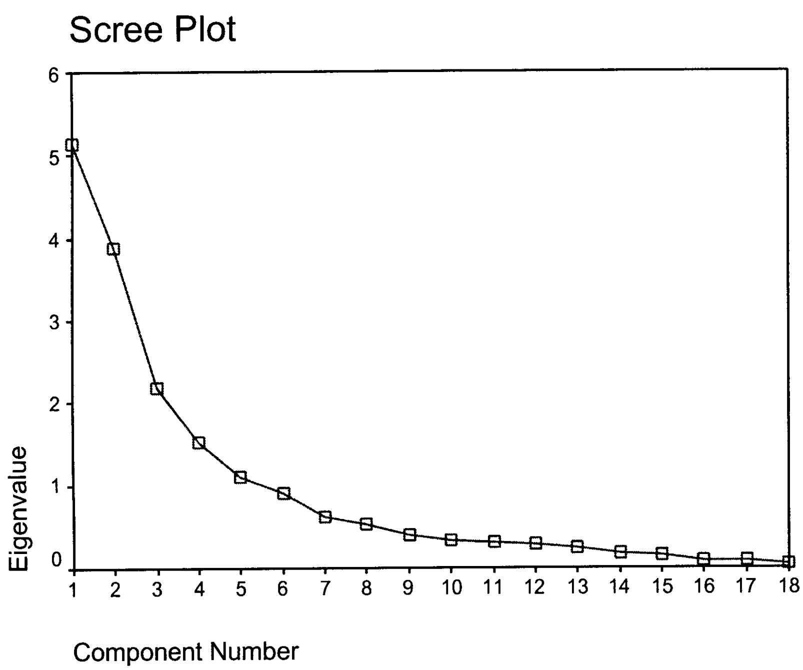
	Initial	Extraction
POP_DEN	1.000	.633
WH_PCT	1.000	.702
KID_PCT	1.000	.863
OLD_PCT	1.000	.871
CIT_PCT	1.000	.897
HHS_PCT	1.000	.836
ENG_PCT	1.000	.900
STU_PCT	1.000	.794
CAR_PCT	1.000	.770
TIME_AVG	1.000	.562
HIED_PCT	1.000	.746
LAB_PCT	1.000	.830
UEMP_PCT	1.000	.531
HHI_MED	1.000	.860
WAGE_PCT	1.000	.830
INC_PC	1.000	.771
POV_PCT	1.000	.659
OWN_PCT	1.000	.767

Extraction Method: Principal Component Analysis.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.142	28.566	28.566	5.142	28.566	28.566	3.848	21.380	21.380
2	3.890	21.609	50.175	3.890	21.609	50.175	3.409	18.937	40.317
3	2.175	12.082	62.257	2.175	12.082	62.257	2.913	16.182	56.499
4	1.514	8.411	70.669	1.514	8.411	70.669	2.475	13.747	70.247
5	1.103	6.127	76.796	1.103	6.127	76.796	1.179	6.549	76.796
6	.908	5.042	81.838						
7	.619	3.440	85.279						
8	.520	2.890	88.169						
9	.396	2.202	90.371						
10	.320	1.780	92.151						
11	.314	1.744	93.895						
12	.284	1.577	95.472						
13	.238	1.323	96.796						
14	.173	.963	97.758						
15	.165	.916	98.675						
16	9.819E-02	.545	99.220						
17	8.574E-02	.476	99.697						
18	5.459E-02	.303	100.000						

Extraction Method: Principal Component Analysis.



Component Matrix[a]

	Component				
	1	2	3	4	5
POP_DEN	-.503	-.291	-.500	-.166	.130
WH_PCT	.789	-7.74E-02	6.519E-02	.233	-.120
KID_PCT	-.343	.801	.310	-1.59E-02	8.885E-02
OLD_PCT	.143	-.701	.414	-.388	-.193
CIT_PCT	.614	-.133	.422	.486	.296
HHS_PCT	.216	-.851	-.107	.224	5.920E-02
ENG_PCT	.700	-.200	.342	.404	.300
STU_PCT	-.291	.771	.294	-.110	.129
CAR_PCT	.465	.505	.434	.112	-.314
TIME_AVG	-5.62E-02	.186	7.237E-02	-.202	.691
HIED_PCT	.658	-.141	-.422	-.215	.263
LAB_PCT	.242	.482	-.601	.422	1.345E-02
UEMP_PCT	-.631	-2.04E-02	.231	.106	.259
HHI_MED	.720	.360	-.168	-.418	9.436E-02
WAGE_PCT	.185	.659	-.536	.268	-6.20E-02
INC_PC	.759	-.107	-.225	-.315	.183
POV_PCT	-.727	-.181	9.810E-02	.139	.261
OWN_PCT	.602	.377	.312	-.406	3.040E-02

Extraction Method: Principal Component Analysis.

a. 5 components extracted.

Rotated Component Matrix[a]

	Component				
	1	2	3	4	5
POP_DEN	-.229	-.378	-.620	6.136E-02	.221
WH_PCT	.470	-.197	.606	8.612E-02	-.260
KID_PCT	-.224	.866	-4.80E-02	.194	.150
OLD_PCT	.128	-.314	5.991E-02	-.852	-.162
CIT_PCT	8.573E-02	-.133	.928	-2.19E-02	.103
HHS_PCT	-5.11E-02	-.848	.235	-.240	-3.96E-02
ENG_PCT	.208	-.224	.890	-4.12E-02	.113
STU_PCT	-.133	.843	-7.15E-02	.145	.201
CAR_PCT	.281	.557	.488	4.865E-02	-.375
TIME_AVG	5.547E-02	.186	2.468E-02	-3.73E-03	.724
HIED_PCT	.711	-.397	9.016E-02	.156	.227
LAB_PCT	.176	-1.10E-02	5.315E-02	.890	-6.76E-02
UEMP_PCT	-.622	.148	-.129	-.130	.298
HHI_MED	.886	.189	9.545E-02	.137	.108
WAGE_PCT	.227	.215	-5.43E-02	.849	-9.90E-02
INC_PC	.804	-.263	.179	-6.83E-03	.154
POV_PCT	-.706	-4.18E-02	-.232	-.117	.302
OWN_PCT	.673	.453	.277	-.176	4.539E-02

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

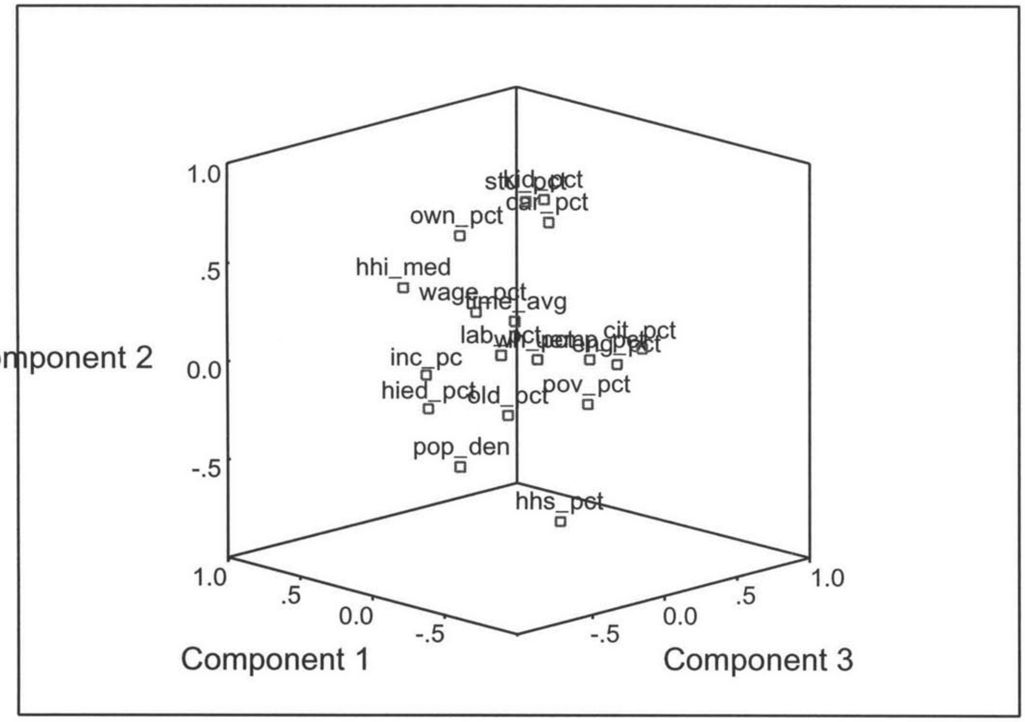
a. Rotation converged in 8 iterations.

Component Transformation Matrix

Component	1	2	3	4	5
1	.787	-.156	.581	.071	-.117
2	.145	.836	-.028	.528	.035
3	-.245	.475	.533	-.655	-.025
4	-.547	-.212	.579	.530	-.200
5	-.029	-.080	.204	.082	.972

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Component Plot in Rotated Space



Cluster

Case Processing Summary[a,b]

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
4676	100.0	0	.0	4676	100.0

a. Correlation between Vectors of Values used

b. Average Linkage (Between Groups)

Frequencies

Statistics

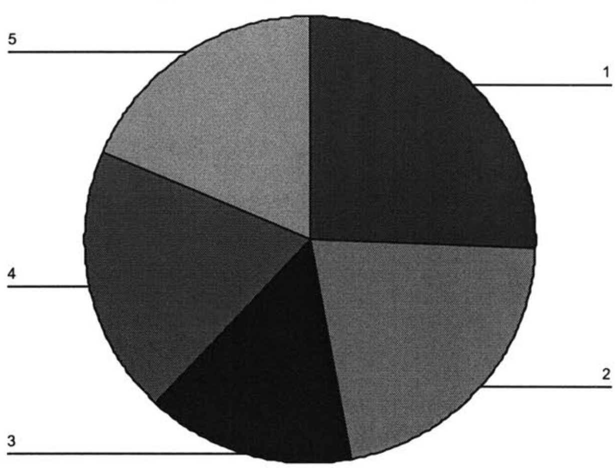
Average Linkage (Between Groups)

N	Valid	4676
	Missing	0

Average Linkage (Between Groups)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1199	25.6	25.6	25.6
	2	1000	21.4	21.4	47.0
	3	695	14.9	14.9	61.9
	4	916	19.6	19.6	81.5
	5	866	18.5	18.5	100.0
	Total	4676	100.0	100.0	

Average Linkage (Between Groups)



Type A

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	1199	-1.16570	4.52542	.9289882	.8531361
REGR factor score 2 for analysis 1	1199	-3.08633	1.61207	-.1488193	.6326573
REGR factor score 3 for analysis 1	1199	-3.83995	1.66917	.3155026	.5919530
REGR factor score 4 for analysis 1	1199	-6.95655	1.73074	-.6775470	1.1250063
REGR factor score 5 for analysis 1	1199	-4.01675	3.34837	-.1259848	.9454148
Valid N (listwise)	1199				

Type B

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	1000	-2.33795	2.66733	-9.4E-02	.6882499
REGR factor score 2 for analysis 1	1000	-3.96247	1.72025	-1.05335	.9976501
REGR factor score 3 for analysis 1	1000	-3.77106	1.95506	-.1565612	.9102089
REGR factor score 4 for analysis 1	1000	-2.62231	2.81793	.7919178	.6087151
REGR factor score 5 for analysis 1	1000	-2.80466	3.49901	7.45E-02	.9543178
Valid N (listwise)	1000				

Type C

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	695	-4.17045	1.07134	-1.08674	.9300875
REGR factor score 2 for analysis 1	695	-3.60562	2.99693	.5598772	.8416974
REGR factor score 3 for analysis 1	695	-3.13676	2.43653	2.56E-02	.9713146
REGR factor score 4 for analysis 1	695	-6.50683	2.87295	-.3474912	.9041011
REGR factor score 5 for analysis 1	695	-.93387	5.26320	1.0629225	.9149524
Valid N (listwise)	695				

Type D

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	916	-2.53737	.88827	-.4618086	.5257395
REGR factor score 2 for analysis 1	916	-2.08025	3.32332	.2316039	.5849095
REGR factor score 3 for analysis 1	916	-1.04338	1.90520	.7516149	.4030587
REGR factor score 4 for analysis 1	916	-2.85535	2.53790	.1563124	.6924672
REGR factor score 5 for analysis 1	916	-4.29223	1.21117	-.5054478	.6243335
Valid N (listwise)	916				

Type E

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	866	-2.05195	3.62227	.1828378	.6696544
REGR factor score 2 for analysis 1	866	-2.28454	3.42867	.7280793	.7859515
REGR factor score 3 for analysis 1	866	-5.65404	.82346	-1.07163	1.0577913
REGR factor score 4 for analysis 1	866	-6.31375	2.00678	.1371655	.7467999
REGR factor score 5 for analysis 1	866	-4.32778	4.16607	-.2299761	.8782280
Valid N (listwise)	866				

Appendix F. Statistical output for Dallas

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std.	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
POP_DEN	3510	.21340	1590.323	168.4519	152.9531	2.392	.041
HH_DEN	3510	.06628	917.74194	66.88581	75.01253	3.833	.041
WH_PCT	3510	.00000	100.00000	74.50468	28.19257	-1.397	.041
KID_PCT	3510	.00000	62.89105	26.34913	9.2585280	-.422	.041
OLD_PCT	3510	.00000	92.30769	9.3919848	8.5127116	1.936	.041
CIT_PCT	3510	12.04128	100.00000	90.20157	12.03110	-2.398	.041
HHS_PCT	3510	.00000	100.00000	53.18149	17.68373	.386	.041
ENG_PCT	3510	4.36123	100.00000	77.91797	16.14433	-1.886	.041
STU_PCT	3510	.00000	57.50000	18.91544	7.6530692	-.215	.041
CAR_PCT	3510	.00000	100.00000	91.25161	9.3939260	-3.371	.041
TIME_AVG	3510	7.75000	53.19643	23.70467	4.7493268	.618	.041
HIED_PCT	3510	.00000	88.98810	24.28996	19.39297	.711	.041
LAB_PCT	3510	2.28690	100.00000	71.47353	12.51054	-.759	.041
UEMP_PCT	3510	.00000	59.25926	6.5165835	6.1106421	2.458	.041
HHI_MED	3510	4999.000	150001.0	35664.93	19299.97	1.751	.041
WAGE_PCT	3510	8.12183	100.00000	83.67042	11.91183	-1.203	.041
INC_PC	3510	944.00000	158592.0	16125.99	10867.33	3.445	.041
POV_PCT	3510	.00000	100.00000	12.59284	14.34606	1.921	.041
OWN_PCT	3510	.00000	100.00000	61.12050	29.65965	-.732	.041
Valid N (listwise)	3510						

Correlations

Correlations

		POP_DEN	HH_DEN	WH_PCT	KID_PCT	OLD_PCT	CIT_PCT	HHS_PCT	ENG_PCT	STU_PCT	CAR_PCT	TIME_AVG	HIED_PCT	LAB_PCT	UEMP_PCT	HHI_MED	WAGE_PCT	INC_PC	POV_PCT	OWN_PCT
POP_DEN	Pearson Correlation	1.000	.914	-.247	-.054	-.151	-.424	.138	-.354	-.120	-.203	-.165	-.009	.164	.093	-.218	.098	-.121	.189	-.466
	Sig. (2-tailed)	.	.000	.000	.001	.000	.000	.000	.000	.000	.000	.000	.586	.000	.000	.000	.000	.000	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
HH_DEN	Pearson Correlation	.914	1.000	-.124	-.273	-.092	-.262	.369	-.181	-.318	-.142	-.199	.115	.226	.008	-.191	.087	-.005	.098	-.499
	Sig. (2-tailed)	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.623	.000	.000	.769	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
WH_PCT	Pearson Correlation	-.247	-.124	1.000	-.264	.048	.340	.179	.389	-.220	.466	-.082	.470	.245	-.592	.516	.226	.464	-.674	.350
	Sig. (2-tailed)	.000	.000	.	.000	.004	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
KID_PCT	Pearson Correlation	-.054	-.273	-.264	1.000	-.372	-.182	-.797	-.329	.876	.018	.326	-.387	-.025	.206	-.070	.089	-.382	.258	.179
	Sig. (2-tailed)	.001	.000	.000	.	.000	.000	.000	.000	.000	.286	.000	.000	.134	.000	.000	.000	.000	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
OLD_PCT	Pearson Correlation	-.151	-.092	.048	-.372	1.000	.188	.372	.189	-.291	-.101	-.182	-.042	-.674	.058	-.091	-.704	.102	.025	.130
	Sig. (2-tailed)	.000	.000	.004	.000	.	.000	.000	.000	.000	.000	.000	.013	.000	.001	.000	.000	.000	.140	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
CIT_PCT	Pearson Correlation	-.424	-.262	.340	-.182	.188	1.000	.145	.893	-.082	.238	.135	.202	-.054	-.155	.259	-.066	.246	-.359	.394
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.001	.000	.000	.000	.000	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
HHS_PCT	Pearson Correlation	.138	.369	.179	-.797	.372	.145	1.000	.234	-.760	-.079	-.339	.300	.008	-.117	-.155	-.231	.247	-.061	-.401
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.655	.000	.000	.000	.000	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
ENG_PCT	Pearson Correlation	-.354	-.181	.389	-.329	.189	.893	.234	1.000	-.158	.218	.082	.344	.014	-.213	.348	-.004	.357	-.415	.353
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.418	.000	.000	.812	.000	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
STU_PCT	Pearson Correlation	-.120	-.318	-.220	.876	-.291	-.082	-.760	-.158	1.000	.024	.312	-.295	-.080	.169	.049	.057	-.269	.167	.300
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.	.147	.000	.000	.000	.000	.004	.001	.000	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
CAR_PCT	Pearson Correlation	-.203	-.142	.466	.018	-.101	.238	-.079	.218	.024	1.000	-.004	.175	.289	-.459	.320	.340	.200	-.553	.386
	Sig. (2-tailed)	.000	.000	.000	.286	.000	.000	.000	.000	.147	.	.814	.000	.000	.000	.000	.000	.000	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
TIME_AVG	Pearson Correlation	-.165	-.199	-.082	.326	-.182	.135	-.339	.082	.312	-.004	1.000	-.212	.038	.079	-.023	.078	-.195	.023	.255
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.814	.	.000	.025	.000	.174	.000	.000	.167	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
HIED_PCT	Pearson Correlation	-.009	.115	.470	-.387	-.042	.202	.300	.344	-.295	.175	-.212	1.000	.282	-.464	.696	.257	.744	-.495	.190
	Sig. (2-tailed)	.586	.000	.000	.000	.013	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
LAB_PCT	Pearson Correlation	.164	.226	.245	-.025	-.674	-.054	.008	.014	-.080	.289	.038	.282	1.000	-.349	.203	.770	.114	-.379	-.103
	Sig. (2-tailed)	.000	.000	.000	.134	.000	.001	.655	.418	.000	.000	.025	.000	.	.000	.000	.000	.000	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
UEMP_PCT	Pearson Correlation	.093	.008	-.592	.206	.058	-.155	-.117	-.213	.169	-.459	.079	-.464	-.349	1.000	-.476	-.365	-.411	.649	-.281
	Sig. (2-tailed)	.000	.623	.000	.000	.001	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
HHI_MED	Pearson Correlation	-.218	-.191	.516	-.070	-.091	.259	-.155	.348	.049	.320	-.023	.696	.203	-.476	1.000	.318	.813	-.616	.581
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.004	.000	.174	.000	.000	.000	.	.000	.000	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
WAGE_PCT	Pearson Correlation	.098	.087	.226	.089	-.704	-.066	-.231	-.004	.057	.340	.078	.257	.770	-.365	.318	1.000	.117	-.424	.039
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.812	.001	.000	.000	.000	.000	.000	.000	.	.000	.000	.021
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
INC_PC	Pearson Correlation	-.121	-.005	.464	-.382	.102	.246	.247	.357	-.269	.200	-.195	.744	.114	-.411	.813	.117	1.000	-.484	.299
	Sig. (2-tailed)	.000	.769	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
POV_PCT	Pearson Correlation	.189	.098	-.674	.258	.025	-.359	-.061	-.415	.167	-.553	.023	-.495	-.379	.649	-.616	-.424	-.484	1.000	-.498
	Sig. (2-tailed)	.000	.000	.000	.000	.140	.000	.000	.000	.000	.000	.167	.000	.000	.000	.000	.000	.000	.	.000
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510
OWN_PCT	Pearson Correlation	-.466	-.499	.350	.179	.130	.394	-.401	.353	.300	.386	.255	.190	-.103	-.281	.581	.039	.299	-.498	1.000
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.021	.000	.000	.
	N	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510	3510

Factor Analysis

Communalities

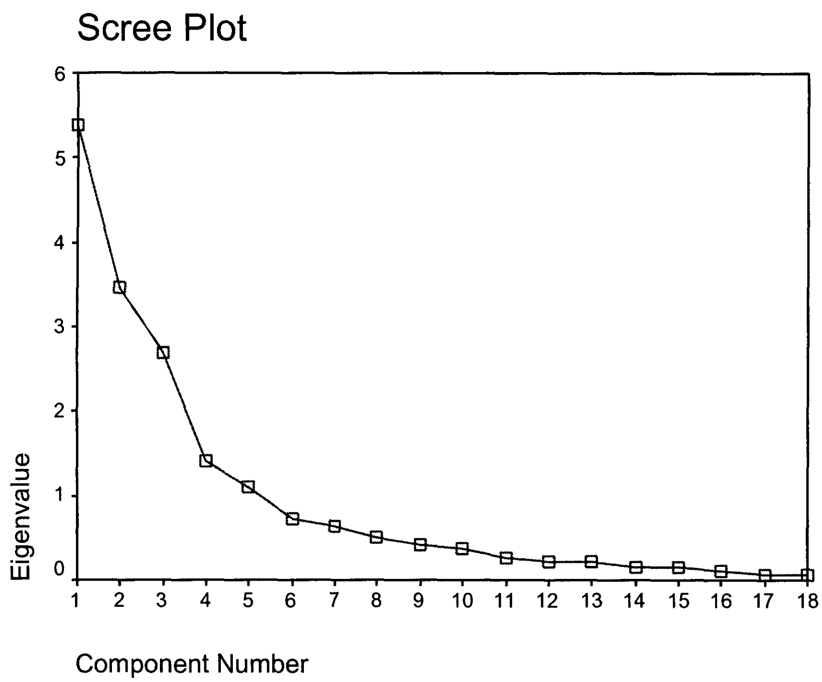
	Initial	Extraction
POP_DEN	1.000	.496
WH_PCT	1.000	.676
KID_PCT	1.000	.863
OLD_PCT	1.000	.854
CIT_PCT	1.000	.873
HHS_PCT	1.000	.862
ENG_PCT	1.000	.881
STU_PCT	1.000	.819
CAR_PCT	1.000	.702
TIME_AVG	1.000	.456
HIED_PCT	1.000	.825
LAB_PCT	1.000	.859
UEMP_PCT	1.000	.670
HHI_MED	1.000	.938
WAGE_PCT	1.000	.844
INC_PC	1.000	.870
POV_PCT	1.000	.793
OWN_PCT	1.000	.788

Extraction Method: Principal Component Analysis.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.385	29.915	29.915	5.385	29.915	29.915	3.337	18.541	18.541
2	3.465	19.247	49.162	3.465	19.247	49.162	2.902	16.122	34.663
3	2.700	14.999	64.161	2.700	14.999	64.161	2.883	16.017	50.680
4	1.412	7.843	72.004	1.412	7.843	72.004	2.576	14.313	64.993
5	1.109	6.161	78.165	1.109	6.161	78.165	2.371	13.172	78.165
6	.735	4.082	82.248						
7	.635	3.530	85.778						
8	.498	2.767	88.545						
9	.428	2.376	90.921						
10	.376	2.087	93.008						
11	.259	1.437	94.445						
12	.230	1.279	95.724						
13	.213	1.181	96.905						
14	.165	.914	97.819						
15	.148	.821	98.641						
16	.118	.655	99.295						
17	6.832E-02	.380	99.675						
18	5.852E-02	.325	100.000						

Extraction Method: Principal Component Analysis.



Component Matrix[a]

	Component				
	1	2	3	4	5
POP_DEN	-.288	-.148	-.601	-.156	7.362E-02
WH_PCT	.774	1.412E-02	9.805E-03	2.691E-02	-.276
KID_PCT	-.422	.792	.194	-.140	-1.06E-02
OLD_PCT	1.732E-02	-.646	.563	-.207	-.277
CIT_PCT	.521	-8.09E-02	.543	.488	.247
HHS_PCT	.234	-.846	-.225	.201	-3.42E-03
ENG_PCT	.614	-.155	.436	.434	.319
STU_PCT	-.308	.768	.312	-.181	6.448E-02
CAR_PCT	.542	.314	9.362E-04	.165	-.532
TIME_AVG	-.101	.448	.282	.346	.215
HIED_PCT	.742	-.137	-.219	-.315	.330
LAB_PCT	.353	.370	-.694	.326	9.602E-02
UEMP_PCT	-.703	-.114	.201	5.441E-02	.346
HHI_MED	.779	.263	6.899E-02	-.443	.247
WAGE_PCT	.364	.545	-.602	.215	7.529E-02
INC_PC	.749	-.158	-1.98E-02	-.440	.299
POV_PCT	-.842	-.166	4.568E-02	-4.85E-02	.225
OWN_PCT	.493	.417	.575	-.176	-9.70E-02

Extraction Method: Principal Component Analysis.

a. 5 components extracted.

Rotated Component Matrix[a]

	Component				
	1	2	3	4	5
POP_DEN	-.238	-.276	-2.58E-02	.276	-.535
WH_PCT	-.184	.696	.334	5.018E-02	.209
KID_PCT	.890	-9.47E-02	-.197	9.807E-02	-.116
OLD_PCT	-.309	6.694E-02	1.262E-04	-.863	9.900E-02
CIT_PCT	-.112	.153	.120	-7.39E-02	.904
HHS_PCT	-.919	-2.07E-02	3.671E-02	-.109	5.954E-02
ENG_PCT	-.221	.139	.250	-1.70E-02	.866
STU_PCT	.898	-8.65E-02	-6.76E-02	2.613E-02	-2.53E-03
CAR_PCT	8.474E-02	.814	-2.13E-02	.138	.116
TIME_AVG	.432	-.108	-.192	.191	.429
HIED_PCT	-.287	.172	.822	.180	6.984E-02
LAB_PCT	-.100	.248	7.795E-02	.883	-3.73E-02
UEMP_PCT	.139	-.721	-.313	-.182	1.534E-02
HHI_MED	.185	.336	.872	8.672E-02	.153
WAGE_PCT	.110	.302	.139	.848	-5.09E-02
INC_PC	-.203	.188	.884	-3.13E-02	.101
POV_PCT	9.318E-02	-.735	-.387	-.194	-.237
OWN_PCT	.494	.461	.347	-.243	.388

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

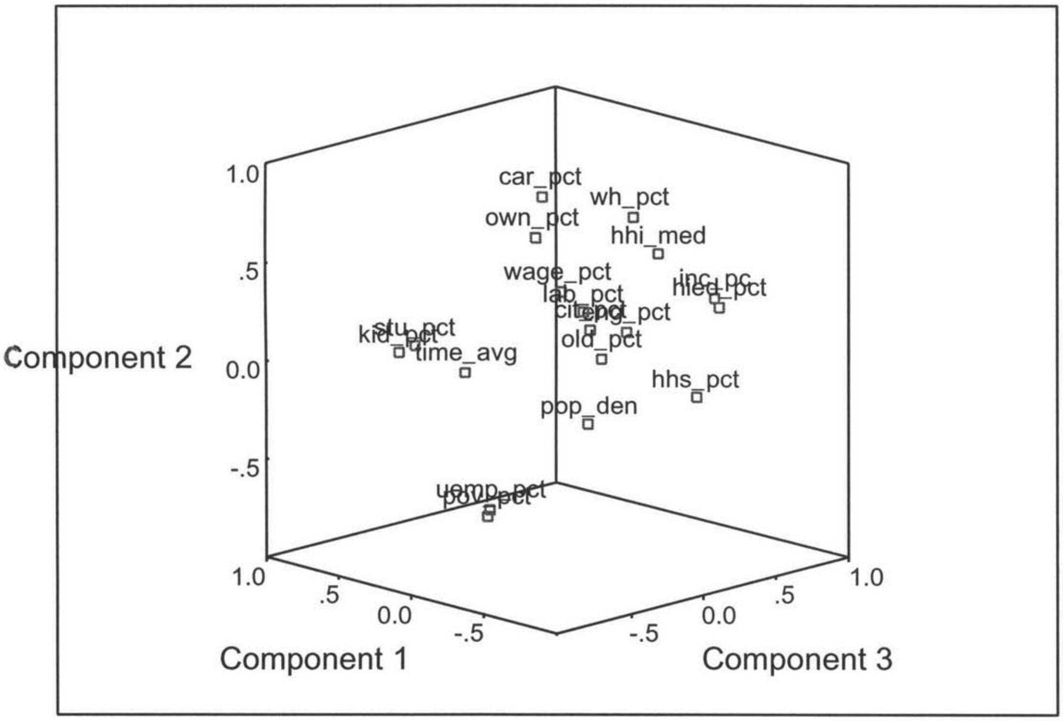
a. Rotation converged in 6 iterations.

Component Transformation Matrix

Component	1	2	3	4	5
1	-.239	.627	.622	.157	.371
2	.862	.215	-.013	.458	.020
3	.367	.000	-.031	-.717	.592
4	-.253	.057	-.624	.405	.615
5	.028	-.746	.472	.295	.364

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Component Plot in Rotated Space



Cluster

Case Processing Summary[a,b]

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
3510	100.0	0	.0	3510	100.0

a. Correlation between Vectors of Values used

b. Average Linkage (Between Groups)

Frequencies

Statistics

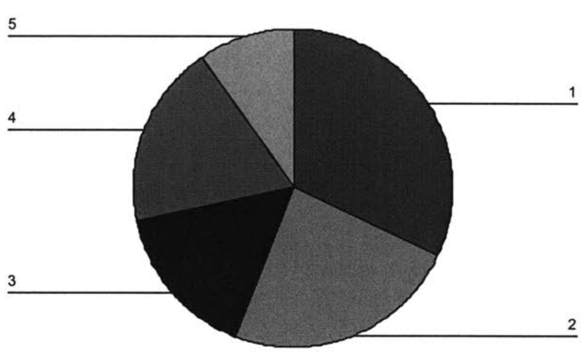
Average Linkage (Between Groups)

N	Valid	3510
	Missing	0

Average Linkage (Between Groups)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1126	32.1	32.1	32.1
	2	834	23.8	23.8	55.8
	3	554	15.8	15.8	71.6
	4	652	18.6	18.6	90.2
	5	344	9.8	9.8	100.0
	Total	3510	100.0	100.0	

Average Linkage (Between Groups)



Type A

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	1126	-2.37854	2.25021	9.30E-02	.6258153
REGR factor score 2 for analysis 1	1126	-1.29033	1.87634	.5075302	.4259647
REGR factor score 3 for analysis 1	1126	-2.26905	.83423	-.5734094	.3956712
REGR factor score 4 for analysis 1	1126	-4.33009	2.16332	-.3365303	.7914509
REGR factor score 5 for analysis 1	1126	-1.71424	2.30495	.4203638	.5583690
Valid N (listwise)	1126				

Type B

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	834	-1.59022	2.74270	.7347097	.5822489
REGR factor score 2 for analysis 1	834	-2.89482	1.27751	.1079922	.6804259
REGR factor score 3 for analysis 1	834	-1.41121	2.22043	-.1610852	.6303040
REGR factor score 4 for analysis 1	834	-2.45766	2.47741	.3012717	.6896203
REGR factor score 5 for analysis 1	834	-5.52860	1.59971	-.9374701	1.2806465
Valid N (listwise)	834				

Type C

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	554	-3.22573	.92134	-1.26434	.9694702
REGR factor score 2 for analysis 1	554	-3.23530	1.33865	-.1928597	.6778682
REGR factor score 3 for analysis 1	554	-2.05349	4.03954	4.30E-02	.6981376
REGR factor score 4 for analysis 1	554	-1.02361	2.86054	1.1332380	.5928583
REGR factor score 5 for analysis 1	554	-4.08883	1.41161	-.1425747	.7458039
Valid N (listwise)	554				

Type D

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	652	-3.17420	2.14270	-.2263603	1.0096770
REGR factor score 2 for analysis 1	652	-4.70221	1.46284	.1385772	.6107641
REGR factor score 3 for analysis 1	652	-1.08903	7.92172	1.3382666	1.2615267
REGR factor score 4 for analysis 1	652	-7.18732	1.61761	-.5715461	1.0579298
REGR factor score 5 for analysis 1	652	-2.98500	1.84889	.1170600	.4828452
Valid N (listwise)	652				

Type E

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1 for analysis 1	344	-2.53048	2.72893	.3794078	.7125027
REGR factor score 2 for analysis 1	344	-8.47196	.52401	-1.87515	1.6015598
REGR factor score 3 for analysis 1	344	-1.34349	1.24279	-.3382918	.5301673
REGR factor score 4 for analysis 1	344	-5.95330	1.70000	-.3706201	.9510669
REGR factor score 5 for analysis 1	344	-.63803	3.48027	.9046037	.5375006
Valid N (listwise)	344				

Bibliography

Alonso, William. 1964. *Location and land use; toward a general theory of land rent*. Cambridge, MA: Harvard University Press.

Ayse Can. 1998. *GIS and Spatial Analysis of Housing and Mortgage Markets*. Journal of Housing Research, 9(1).

Chiang, Alpha C. 1984. *Fundamental Methods of Mathematical Economics*, 3rd ed. New York, NY: McGraw-Hill Book Company.

DiPasquale, Denise, and William C. Wheaton. 1996. *Urban Economics and Real Estate Markets*. Englewood Cliffs, NJ: Prentice Hall.

Gujarati, Damodar N. 1978. *Basic Econometrics*. New York, NY: McGraw-Hill Book Company.

Isard, Walter, Iwan J. Azis, Matthew P. Drennan, Ronald E. Miller, Sidney Salzman, and Erik Thorbecke. 1998. *Methods of Interregional and Regional Analysis*. Aldershot, England: Ashgate.

Kennedy, Peter. 1998. *A Guide to Econometrics*, 4th ed. Cambridge, MA: MIT Press.

McMillen, Daniel P. and John F. McDonald. 1998. *Suburban Subcenters and Employment Density in Metropolitan Chicago*. Journal of Urban Economics, 43.

Monmonier, Mark. 1996. *How to Lie with Maps*, 2nd ed. Chicago, IL: The University of Chicago Press.

O'Sullivan, Arthur. 2000. *Urban Economics*, 4th ed. New York, NY: McGraw-Hill Book Company.

Pindyck, Robert S., and Daniel L. Rubinfeld. 1991. *Econometric Models and Economic Forecast*, 3rd ed. New York, NY: McGraw-Hill Book Company.

Ricardo, David. 1965 (1817; reprinted). *The Principles of Political Economy and Taxation*. London: J. M. Dent and Son.

Treyz, George I. 1993. *Regional Economic Modeling: A Systematic Approach to Economic Forecasting and Policy Analysis*. Boston, MA: Kluwer Academic Publisher.

Isard, Walter, Iwan J. Azis, Matthew P. Drennan, Ronald E. Miller, Sidney Saltzman, and Erik Thorbecke. 1998. *Methods of Interregional and Regional Analysis*. Brookfield, VT: Ashgate.

Wheaton, William C. 1977. *Residential Decentralization, Land Rents, and the Benefits of Urban Transportation Investment*. The American Economic Review, 67(2).

Wheaton, William C. 1996. *Telecommunications Technology and Real Estate: Some Perspective*. (Unpublished)