Industrial Structure and Employment Growth in the 1990s in Appalachian Counties

by

Zhijun (Jeanne) Tan

Bachelor of Science,
Sun Yat-Sen University, 2002

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
September 2005

Signature of Author

Department of Urban Studies and Planning
August 10, 2005

Certified by

Karen R. Polenske
Professor of Regional Political Economy and Planning
Thesis Supervisor

Accepted by

Langley Keyes
Ford Professor of City and Regional Planning
Chairman, MCP Committee, Department of Urban Studies and Planning

Industrial Structure and Employment Growth in the 1990s in Appalachian Counties

by

Zhijun (Jeanne) Tan

Submitted to the Department of Urban Studies and Planning on August 10, 2005 in Partial Fulfillment of the Requirements for the Degree of Master in City Planning

ABSTRACT

Employment growth in the 1990s and its relationship with the initial industrial structure in 1990 are examined in the case of Appalachian counties, after controlling for labor-market conditions and other factors, such as labor mobility, natural amenities, and market size.

Spatial exploratory data analysis of the competitive employment growth (CEG) in Appalachian region shows that strong spillover effect of CEG exists among 410 counties. Counties with higher employment growth rates are concentrated in the north side of Atlanta Metropolitan area around Interstate highway I-75. Counties with lower growth rates concentrated in Central Appalachia, along the convergent border of three states, Kentucky, Virginia, and West Virginia. Another low growth rate concentration is in Northeast Pennsylvania.

The existence of spatial autocorrelation affects my empirical model's explanatory power, the significance levels, and the values of coefficients of independent variables. There is no specific theoretical base for the county-interacting mechanism of this empirical model, whereas, the magnitude of each independent variable's impact on employment growth depends on the spatial-weight matrixes. To find a better match, I compare the parameters of the model by using two different weight matrixes, i.e., weights based on physical neighbor interaction and weights based on commuting ties. Based on the result of statistical comparison, the commuting tie is more likely the way by which counties interact with each other than physical proximity in the Appalachian Region.

My empirical model is not able to explain completely the employment growth for all the counties in Appalachian Region, even after being adjusted for the spatial spillover effects, but it does provide some insight about what factors might matter for many places for their competitive employment growth from 1990 to 2000. Also, by analyzing the residuals of this model, analysts will be able to find some good candidates for case studies to understand what other determinants of economic growth might be.

Thesis Supervisor: Karen R. Polenske
Title: Professor of Regional Political Economy and Planning
Thesis Reader: Joseph Ferreira, Jr.
Title: Professor of Urban Planning and Operations Research

Acknowledgement

In this finial work for my MCP (Master in City Planning) program in the Department of Urban Studies and Planning (DUSP), I would like to express my gratitude to all those who offered me great help and advices during my two-year study at MIT, especially, to my thesis supervisor Professor Karen R. Polenske, and my thesis reader Professor Joseph Ferreira, Jr., who gave me the possibility to complete this thesis.

I am deeply indebted to my supervisor, Professor Karen R. Polenske, not only for her great help, stimulating suggestions and encouragement in the time of research for and writing of this thesis, but also for her constructive advices on my academic, professional, and personal development during my study at MIT. From her, I learned the importance of patience, hardworking, persistence, and scrupulosity in the economic research and I believe it will be benefiting me through my whole professional and personal life.

I would like express my deeply thank to my thesis reader, Professor Joseph Ferreira, Jr. He kept an eye on the progress of my work and always provided his advises/insights in the methodology and the interpretation of the result in this empirical study and keep me stay on the right track. Without his input, I might have been stuck in the middle of this research.

I am grateful for Professor Lun Anselin in the University of Illinois, Urbana-Champaign, for the software "GeoDa" developed by his laboratory and his great help in interpreting the spatial regression result in GeoDa.

I also want to thank the Department of Urban Studies and Planning for giving me the financial support and the DUSP's academic advisor Sandra Wellford and her assistant Alice Twohig for their help these two years.

I feel a deep sense of gratitude for my father and mother who formed part of my vision and taught me the good things that really matter in life. I am grateful for them, my brother, and my grandparents, for their support and selfless love. I also want to thank all of my friends, especially (but not limited to) Xin Li, Xiaoyu Shi, Duncan Kincaid, Zhan Guo, Jinhua Zhao, Hongliang Zhang, Ming Guo, Ayman Ismail, for their great help on my thesis and their friendships. The happy memory of being with them will provide a persistent inspiration for the journey in my life.

Finally, I want to give my special thank to my beloved boyfriend, Jianrong (Steve), for his love and patience.

Table of Contents

Acronyms

ARC: Appalachian Regional Commission

BLS: Bureau of Labor Statistics

CBP: County Business Pattern

CEG: Competitive Employment Growth

Coef.: Coefficient

CZs: Commuting Zones

BEA: Bureau of Economic Analysis

Eq.: Equation

LISA: Local indicators of spatial association

MAR: Marshall-Arrow-Romer (externalities)

MSA: Metropolitan Statistical Areas

OLS: Ordinary Least Squares

OMB: Office of Management and Budget

P.: Probability

R& D: Research and Development

REIS: Regional Economic Information System

SIC: Standard Industrial Classifications code

t: T statistic

ERS/USDA: Economic Research Service/United State Department of Agriculture

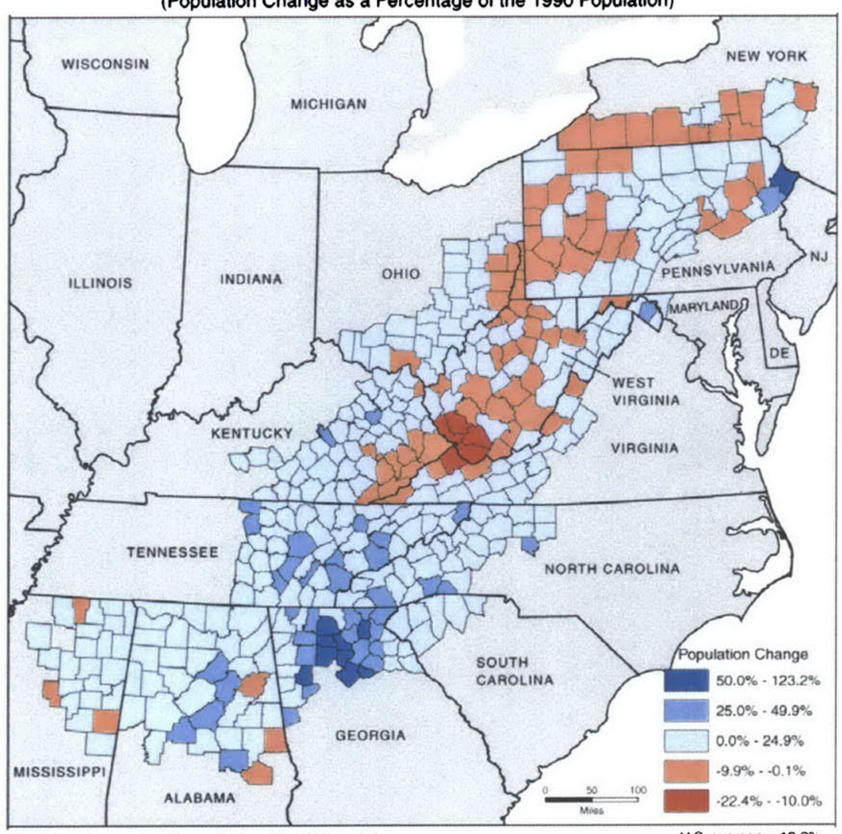
Chapter 1

Introduction and Literature Review

1.1 Introduction

Appalachia, having 22.9 million people in 2000—8 percent of the total U.S. population of 281.4 million—is known for its economic hardships and geographic isolation. Since a landmark report by the President's Appalachian Regional Commission (ARC) in 1964, its variety of economic development experiences is also studied by many analysts. In Pollard's (2003) study, he found that the Appalachian population grew 9 percent during the 1990s, which was slower than the rest of the nation (13 percent).

Figure 1: Population Change in Appalachia, 1990-2000
(Population Change as a Percentage of the 1990 Population)



Population change is the difference between the 1990 population and the 2000 population expressed as a percentage of the 1990 population.

U.S. average = 13.2%
Appalachian average = 9.1%

Source: ARC website

Moreover, these developments had some wide-ranging variations among Appalachia's subregions and economic development categories. As Pollard pointed out, "many areas continue to be almost as economically depressed as they were during the Johnson administration. Many other areas, however, have an economic status that approaches—and in some cases surpasses—that of the United States as a whole."

We might wonder why different regions have different economic growth rates. What are the factors contributing to the patterns or rates of regional economic growth? Questions like these, related to the source and mechanism of economic growth, are fundamental issues in regional economics. They have intrigued economists for centuries, ever since the publication of Adam Smith's "Inquiry into the Nature and Causes of the Wealth of Nations" (Adam Smith, 1776).

As one of the most common measurements of regional economic growth, I use employment growth as the measure in this study. For a geographic area, employment growth is directly related to business activities in the jurisdiction of that region. What makes its employment grow? Obviously, the establishment of new plants/firms, spin-offs of old firms, or the expansion of existing ones can result in positive growth in the regional employment. Conversely, the closure of old firms or reduction of employees in existing factories will cause a negative employment "growth." Therefore, to study the mechanism of employment growth, a logical approach is to study what forces attract firms to a region, what forces make them leave. What virtues of the region make existing firms prosper, and what vices make firms lose profits?

As discussed by Kusmin (1994), numerous economic-growth analysts have identified many forces or characteristics of a region associated with employment growth. They cover almost every aspect of the economy. Among those factors associated with employment growth, the industrial mix of a region is one of the most important aspects in its industrial structure, as suggested by Garcia-Mila and McGuire (1993). Also, the effects of economic diversity and industrial specialization on employment growth are implicitly suggested by those analysts who study the impact of urbanization and localization economies on the process of local economic development. In their study of the U.S. cities between 1956 and 1987, Glaeser et al. (1992) suggested that sectoral diversity matters for local growth. Yet, in their study of localization economies, Henderson et al. (1995) show evidence of a positive effect of specialization on urban growth between 1970 and 1987.

Kusmin (1994) did a comprehensive review of 35 empirical studies of factors influencing business location and/or regional economic growth. His review shows how labor-market conditions, which characterize the local economic structure, contribute to the explanation of development. Additional variables, such as mobility/stability, natural amenities, and metropolitan status of a county, also have significant impacts on economic growth.

Taking all the factors considered above, I propose a relatively universal specification/model of employment growth to conduct empirical tests.

Employment Growth = f(industrial mix, economic diversity, manufacturing specialization, labor market characteristics, natural amenity, market size).

In order to account for the "true" endogenous growth for each county, I use the competitive component of employment growth (net of the effect of national growth and sectoral growth from 1990 to 2000) from shift-share analysis as the dependent variable.[1] For sake of simplification, I use the term "competitive employment growth," CEG, in this study to refer to this endogenous growth. In Chapter 2, I provide a more precise description of these variables.

Some of these factors had been studied by other researchers, but none of those analysts includes all the factors in their growth model. Before the advent of spatial econometrics and the related computer software, analysts could not do a comprehensive examination of spatial effects in growth models due to the huge volume of observations. For example, ERS/USDA (Economic Research Service/United State Department of Agriculture) has conducted a statistical study of rural economic growth that took into account most of the factors suggested in Kusmin's literature review of 35 empirical studies, but they failed to consider the possibility of the existence of spatial autocorrelation, which might distort coefficients and significance levels of all the exploratory variables in the model. Garcia-Mila and McGuire (1993) discussed the impact of the historical economic structure on the employment growth rate, but they did not incorporate the spatial interaction of different regions' economic structures, either.

In recent years, the field of spatial econometrics has rapidly developed, providing more powerful tools and test methods for researchers to reveal the nature of economic growth and the relationship among regions. For example, Moran's I coefficient enables analysts to test for certain types of spatial autocorrelation and econometric methods have been developed that allow particular forms of weighted regression to model various spatial spillover effects.

My study has benefited from the development of spatial econometrics, including related computer software (especially GeoDa[2]), and the rich literature on economic-growth models in

---

[1] Shift-share analysts refer to Dunn (1960), in which competitive effects are originally defined.
[2] Geoda 9i—GeoDa is the latest incarnation in a long list of spatial software tools developed by Dr. Luc Anselin's Spatial Analysis Laboratory (SAL) in the Department of Geography at the University of Illinois, Urbana-Champaign.

both theoretical and empirical studies. As an attempt to test their applicability to the employment growth of Appalachian Region counties, I do not try to devise a complete model of the macroeconomics of regional economies in my study; instead, I hope to provide some compelling empirical evidence for the link between the growth/variability of regional economies and the limited, but critical, factors, i.e., industrial structure and labor-market conditions. Most importantly, I examine the heterogeneity of and spatial interdependence in my empirical model to account for the great diversity and the degree of counties' interaction in Appalachian Region. Not surprisingly, the result shows evidence of heteroskedasticity among 410 counties— a result that often occurs when there is a large difference between the size (and population) of the individual observations (counties). I also detect spatial effects by examining the interactions between neighboring counties in the empirical model.

Specifically, the driving forces or constraints of the growth are very diverse between metropolitan and non-metropolitan counties, among the three sub-regions in Appalachia, and among different economic-level groups. Finally, I examine the existence of spatial autocorrelation and show how it affects the outcome of the models and performance of different variables by using spatial-lag econometric modeling techniques.

The structure of my study is:

In Chapter 1, I review the theoretical and empirical literature about the relationship between employment growth and county's industrial structure, i.e., their industrial mix, diversity, and specialization. I also cover the interplay between employment growth and the other factors associated with the growth, such as labor-market conditions, mobility/stability, natural amenities, and metro status of a county.

In Chapter 2, I explain the methodology and variables I used in my study.

In Chapter 3, I provide the descriptive statistics of all the variables I used for the Appalachian Region, and compare the characteristics of these variables based on three county classifications, used in most of the previous studies of this region.

In Chapter 4, I present the results of my analysis. Specifically, I test the same empirical growth model on different groups of counties and compare the regression results. Then, I examine the spatial autocorrelation of employment growth among these counties. Finally, I employ two weight matrices to see how, and to what extent, they can affect the outcome of the same empirical growth model, and can explain some of the spatial interactions.

In summary, my study provides some insight into the diversity and spatial dependence of Appalachian counties by using statistical analysis and spatial econometric techniques. But the empirical model I use in this study is not perfect; I have difficulty explaining some results by just looking at the "data" themselves. Case studies and some other qualitative methods are needed in order to disentangle the puzzle of economic/employment growth, but those are beyond the scope of this study.

1.2 Literature Review: Factors Associated with Employment Growth

I classify factors associated with employment growth into two major groups: economic factors and non-economic factors. Within the economic factors, I distinguish core factors and peripheral factors. Specifically, core factors are those directly associated with the economic structure, such as industrial mix, diversity, specialization, and the size of the economy; the peripheral factors include wages and the unemployment rate. Non-economic factors include natural amenities, demographic features, and mobility etc. I will discuss them in order in this Section.

1.2.1 Industrial Structure and Employment Growth

For this study, I include three different, but related, components as parts of industrial structure: industry mix, diversity, and specialization. Most analysts explain their influences on employment growth through two perspectives: an agglomeration economies perspective and an economic-stability/business-cycle perspective.

A. Agglomeration-economies perspective

Initial industrial mix matters in the employment growth of a region because a specific industry could influence the growth of the economy by taking advantage of agglomeration economies. Agglomeration economies are general cost savings or productivity increases resulting from a geographic concentration of firms. If agglomeration economies characterize a specific industry rather than all industries, then a region with a high share of employment in an industry exhibiting agglomeration economies will experience a higher growth rate relative to regions with high concentrations of industries that do not exhibit agglomeration economies.[3] If an industry is concentrated in a region having agglomeration diseconomies, the region will be relatively better-off than other regions with a lower concentration of that industry and vice versa.

Sectoral specialization and diversity affect the total employment growth through different types of agglomeration effects: "localization economies" and "urbanization economies." Localization economies, also referred to as Marshall-Arrow-Romer (MAR) externalities (Marshall, 1890; Romer 1986), imply that a firm benefits from clustering with other firms in the same sector, whereas urbanization economies, also referred to as Jacobs externalities, imply that a firm

[3] This industry-specific form of agglomeration economies is sometimes referred to as localization economies. See the discussion of agglomeration economies by Heilbrun (1987, pp.15-18)

benefits from the clustering of many firms in the same place, regardless of type of firm because of the inter-sectoral positive effects of agglomeration (Jacobs, 1969).

Localization and urbanization economies for a particular industry affect regional economic growth in different way. If there are localization economies, regions involved in that industry are likely to specialize in just that one business activity, or a closely connected set of activities. Specialization, however, allows for full exploitation of scale economies (Henderson et al. 1995) and therefore helps that industry grow. If an industry is subject to urbanization economies, a diverse industrial environment might help it and all firms in the region thrive (Glaeser et al., 1992).

Regional analysts, such as Quigley (1998), believe that knowledge/information spillovers are one aspect of agglomeration economies that can be important in this context. The industrial mix matters here because a bigger share of a specific industry affects the economic growth through knowledge spillovers. If a specific industry devotes substantial investment in the types of research and development (R&D) that have positive spillover effects on the productivity of other industries, a region with a high share of the R&D industry may have a higher overall level of productivity, and therefore a higher growth rate than other regions. The spillover effects of R&D can be negative if a region has a large share of an industry that devotes very little investment in R&D or invests in R&D that is not transferable to other industries. In that case the relative lack of R&D spillovers will make the region grow slower than average.

B. Economic-stability and business-cycle perspective

Firms may be attracted to a region by its great economic stability and variability, because it can reduce the risk of doing business. Therefore, economic stability and diversification is associated with faster growth. The industrial mix, specialization, and diversity may affect the local economic stability through interrelatedness of industries. In this way, the industrial structure indirectly affects the growth rate through affecting the local economic stability.

From the business-cycle perspective, there are two ways to explain the relationship between inter-sectoral connectedness and a region's stability and variability. First, on the one hand, if a region has a large share of an industry that is highly interrelated with other industries through supply or demand of inputs, and the industry is highly variable, its variability could possibly be transmitted to related industries, making the cycle more intense. On the other hand, if the industry happens to be relatively stable, that stability is likely to be transmitted to industries that either provide or demand inputs from the stable industry, thus resulting in an economy that is less unstable.

Second, the intensity of the cycle in a region's economy may also be related to the breadth of the markets of the component industries of the region's economy. On the one hand, an industry that primarily produces goods and services to sell in the local market will not be able to look for alternative buyers outside the region when the local economy goes through a recession. On the other hand, if the goods and services of a majority of the region's industries are sold in the national market, these industries can sell their goods on alternative markets during a local recession, effectively diversifying the risks of local shocks.

1.2.2 Other Factors Associated with Employment Growth

Through the literature review, I identify at least three other important factors associated with employment growth: labor-market conditions, natural amenities, and size of the economy.

A. Labor-market conditions and labor mobility

The impact of labor-market conditions on the total employment growth closely relates to the business-location decisions, which drive economic growth. If a desire to hold down production costs drives business-location decisions, then, higher wages will tend to result in a relative decrease in business activity.

For the same reason, just like the low-wage factor, a high unemployment rate implies that the cost of recruiting new labor is easier and cheaper than in those regions with low unemployment rates, but it also implies that the existing cluster/industries in that area do not match the labor force quality/specialty in the region. Thus, the firms looking for a locality effect (knowledge spillover/supply) or looking for a niche/urbanization effect might benefit from the high unemployment rate.

At the same time, analysts need to be cautious that even if a search for the lowest production costs does drive changes in business activities, they may fail to find that the expected negative relationship between wages and business activity change if they fail to adjust for labor quality. When there are no effective controls for labor-quality differences, differences in the wage measure may primarily reflect differences in labor quality rather than in the cost of labor of a given quality. One way to measure labor quality is to include education-level data. I use the percentage of residents with a high school diploma in a county to indicate the education-level of its labor force.

Economic growth is usually higher in a place with a well-functioning labor market,[4] which can be indicated by high labor mobility. When technological breakthroughs or other forces create new opportunities for local employment growth, or indeed cause specific job losses, high labor mobility ensures that labor is reallocated to where it can be employed most productively. For this study, I use the percentage of residents from the same county five years ago as a proxy for initial condition of labor mobility. The higher the value, the less the county's labor force is mobile.

B. Natural amenities

Amenity is considered an important factor of economic growth because the rise of an elite "technocratic" work force creates an emphasis on privatized consumption and quality-of-life (Gorz 1985, Gouldner 1979). This factor has an impact directly on residential preferences, thus indirectly on business growth because of the ". . . non-wage incentive for privileged workers who choose to forego the possible benefits of traditional locations for the life-style rewards of amenity-rich environments. " (Kasarda and Irwin,1991, p.740). Because of the critical position of the elite "technocratic" workers in the production process, they are able to influence firms to locate in areas meeting their consumption and life-style preferences. Places offering greater environmental amenities, such as mild winter climates and ample outdoor recreational opportunities, gain more in competing for new businesses than those offering relatively few of these amenities or those with disamenities, such as high pollution. McGranahan's study (1999) found that, the average 1970-96 population change in non-metropolitan counties was 1 percent among counties low on the natural-amenities index and 120 percent among counties high on the natural-amenities index,[5] which is developed by ERS/USDA. In my study, I use the same natural amenities scale that McGranahan used to account for the amenities factor in the county.

---

[4] A speech by William Poole, President, Federal Reserve Bank of St. Louis in Southern Illinois University at Edwardsville on April 10, 2003. http://stlouisfed.org/news/speeches/2003/4_10_03.html

[5] The ERS/USDA constructed the index by combining six measures of climate, typography, and water area that reflect environmental qualities most people prefer. These measures are warm winter, winter sun, temperate summer, low summer humidity, topographic variation, and water area. http://www.ers.usda.gov/data/NaturalAmenities/

C. Size of the economy

The size of the economy matters when rural as well as urban areas are considered, because it influences the intensity of agglomeration forces. On one hand, the level and quality of information exchanges in spillovers are sufficiently important only when the number of firms, thus the potential complementarity, is relatively high; on the other hand, the size of local markets greatly affects the firms' location choices if transportation costs are non-zero. Moreover, this variable is relevant for some non-specialized markets, such as land: high density implies high land rent, which constitutes a dispersion force. (Ciccone and Hall, 1996)

Because of the lack of good measurement of the size of the economy and the complexity of this aspect, I use the county designation of metropolitan and non-metropolitan status of 1993 as a proxy to indicate the initial market size.

1.2.3 Spatial-Interaction Effects on the Determinants

In the description of the definition of spatial econometrics, Paelinck and Klassen (1979) stressed the importance of identifying spatial interdependence, the asymmetry of spatial relations, and the relevance of factors located in "other spaces" in empirical models, especially those based on cross-sectional data. In my study, I use cross-sectional data (410 counties) to analyze the regional employment growth and its relationship with industrial structure after controlling for labor-market conditions and natural amenities. They all have potential spillover effects, so that I cannot preclude the existence of spatial autocorrelation. In other words, the employment growth in one county might not be independent from that of a neighboring county. For example, the industrial structures in neighboring counties may affect the county in question by substantial industrial interactions through seller-buyer/input-output relationships or agglomeration economies. If a sector in a county has a rapid growth, the increased demand for its input industries through backward linkages make the surrounding counties with backward linkage industries grow relatively faster. Factor-mobilization theorists (Aldashev and Möller, 2003) also suggest that labor-market conditions in neighboring areas have the same ripple effects because of the interacting equilibrium of employment and wage. For example, if the wage level of surrounding counties is low, a firm might quickly lay-off its employees and might easily relocate to one of those counties to reduce the labor cost. In addition, a county with high natural amenities may have a positive effect in attracting a firm to locate in the county or nearby counties.

I take spatial-interaction effects into account by using spatial econometric approaches that explicitly model and control for spatial autocorrelation or interdependence to avoid inefficient or inconsistent parameter estimates or specification errors. I provide a detailed discussion of the spatial model in Chapter 2, Section 2.2.3.

In summary, the literature I reviewed suggests that initial industrial structure, i.e., the composition, diversity, and specialization, might have significant effects on the employment growth. Other factors, such as the size of the economy and labor-market conditions might be associated with the growth pattern too. I consider spatial effects in the growth model to account for the interaction among regions and areas.

Chapter 2

Variable Measurement and Methodology

I use regression technique to test my main hypothesis, which is that the competitive employment growth (CEG) is highly correlated with industrial mix, industry diversity, and specialization, and the labor-market conditions, but the influences of these factors on employment growth might be distorted if analysts ignore the heterogeneity and effects of spatial autocorrelation in the regression model. I express the hypothesis in Eq. (1):

Competitive Employment Growth = f(industrial mix, economic diversity, service diversity, manufacturing specialization, labor-market characteristics, natural amenity, market size) (1)

In order to offset the effect of distortion caused by heterogeneity, I repeat the regression for each group in three classifications used by most Appalachian researchers, i.e., (1) metropolitan and non-metropolitan counties, (2) three sub-regions, and (3) distressed counties, transitional counties, competitive and attainment counties.

In order to show the effect of distortion caused by spatial autocorrelation, I examine the total competitive component of county employment growth from 1990 to 2000 in three regression models: (1) ordinary least squares (OLS) regression, (2) spatial regression based on a so-called "Queen" [6] contiguity-weight matrix, i.e., where any county that touches another at any point is considered to be contiguous, and (3) a spatial regression based on a commuting-zone-contiguity weight matrix.

2.1 Data sets and Variable Measurement

Here, I explain the definitions and measurements of these factors, including detailed definitions and summary statistics of all variables in Appendix B.

---

[6] "Queen" Contiguity includes all common points–boundaries and vertices in the definition, while "Rook" Contiguity only considers counties that share common boundaries that are lines and not just a single point (Anselin 2003b); both concepts are taken from chess.

2.1.1 Dependent Variable

Instead of using gross employment growth, I construct the dependent variable—county "competitive" employment growth from 1990 to 2000—by using shift-share analysis. Typically, shift-share studies isolate three components (national, industrial-mix, and competitive) of a region's growth by algebraic calculation, with the competitive component describing the portion of growth due to some undetermined regional effect other than national growth or industrial mix. In this way, I distinguish the difference in the growth behaviors of county economies from those that arise simply because one county has a greater share of its employment or output in a relatively fast-growth industry than does another county. The competitive component of shift-share analysis of total employment growth from 1990 to 2000 is the dependent variable CEGi (where i denotes the county) in this study. I use the employment data for 1990 and 2000 in Regional Economic Information System (REIS) 1969-2003, which are published by the Bureau of Economic Analysis.

2.1.2 Independent Variables

I have thirteen independent variables categorized into the following five groups of factors: (1) five of them are categorized as industrial-mix factors; (2) two of them represent the diversity; (3) one represents specialization; (4) four indicate the labor market condition, and (5) natural amenities and market size.

1. Industrial mix

The data set for the five industrial-mix factors contains income levels for 410 counties and the United States by eight private sectors in 1990. The data are from the Regional Economic Information System (REIS[7]) 1969-2003 provided by Bureau of Economic Analysis (BEA). This data set contains nine private sectors as well as one public sector. The nine private sectors are: Agriculture, Forestry, and Fisheries (SIC code 0); Mineral Industries (SIC code 11-14) and Construction Industries (SIC code 15-17); Manufacturing (SIC code 2 and 3); Transportation, Communications, and Utilities(SIC code 4); Wholesale Trade (SIC code 50 and 51); Retail Trade (SIC code 52-59); Finance, Insurance, and Real Estate (SIC code 6); Service Industries (SIC code 7-8). Using these income data, I describe the industrial composition of each economy by calculating the income shares for each industry.

---

[7] http://www.bea.doc.gov/bea/regional/docs/cd.asp

In order to avoid extreme linear correlation, I did not include all shares of the sector in the model. Because I did not have any a priori information on which sectors are most relevant to employment growth, I regressed Eq. (1) using OLS method with all sectors shares and selected the particular combination that provided the highest regression fit. After doing that, I incorporated the most relevant industries into the model, which are mining (**MIN90**), manufacturing (**MANFC90**), transportation (**TRNSP90**), wholesale trade (**WHTRD90**), service (**SERV90**) and public/governmental sectors (**GOV90**).

2. Diversity index—Gini coefficient

I measure two kinds of industrial diversity for each county: diversity within all private sectors and diversity within the services sector. I think there are two important aspects of "diversity". Initially high diversity within all private sectors provides opportunities for industries to gain urbanization economies. Initially low diversity within services sector particularly indicates "room" for entrepreneurship activities because entrepreneurs can find the insufficient supply more easily in an area with low diversified service sector to meet the unlimited diverse need from consumers.

I named them diversity indicator I (**BEAGINI90**) and diversity indicator II (**SGINI90**), respectively, in my study. The level of diversity is indicated by the value of the Gini coefficient (Paci, 1999). Analysts use the Gini coefficient to measure the degree of concentration (inequality) of a variable in a distribution of its elements. It compares the Lorenz curve of a ranked empirical distribution with the line of perfect equality. This line assumes that each element has the same contribution to the total summation of the values of a variable. The Gini coefficient represents the area of concentration between the Lorenz curve and the line of perfect equality as it expresses a proportion of the area enclosed by the triangle defined by the line of perfect equality and the line of perfect inequality. The Gini coefficient ranges between 0, where there is no concentration (perfect equality), and 1 where there is total concentration (perfect inequality). The closer the coefficient is to 1, the more unequal the distribution.

The Gini coefficient, among n entities (industry sectors) is given by G

$$G = \frac{n+1-2V}{n} \quad (0 \le G \le 1)$$

where $V = \sum_{i=1}^{n} Vi$ and $Vi = \sum_{j=1}^{i} Pj = \sum_{j=1}^{i} \frac{Xj}{\sum_{l=1}^{n} Xl} \begin{cases} i = 1,2,...n \\ V_0 = 0 \end{cases}$

I use 1990 sectoral employment data from the Regional Economic Information System (REIS[8]) 1969-2003 to construct the **BEAGINI90** dataset. I use the same nine industries for the private sector that I used in the industrial-composition analysis with the following standard industrial classification (SIC) codes: Agriculture, Forestry, and Fisheries (SIC code 0); Mineral Industries (SIC codes 11-14) and Construction Industries (SIC codes 15-17); Manufacturing (SIC codes 2 and 3); Transportation, Communications, and Utilities (SIC code 4); Wholesale Trade (SIC codes 50 and 51); Retail Trade (SIC codes 52-59); Finance, Insurance, and Real Estate (SIC code 6); Service Industries (SIC codes 7-8).

For the initial diversity of the service sector in 1990, I use the data set in County Business Patterns[9] (1990) provided by the Census Bureau, which contains the disaggregate employment data within the service sector, to calculate the Gini coefficient for the service sector for each county. I provide the summary statistics of Diversity Indicator I and Diversity Indicator II in Appendix B.

I use employment instead of income data for the calculation of **SGINI90**. Income data are less reliable in measuring the scale of each industrial sector, because they mix the information of the scale of industrial sector (employment) as well as wage level (Income = Employment * Wage). For example, the scale of the service sector might be underrepresented if I measure it by the income-share measure because the wage level for a service-sector worker is lower than the wages in other sectors. Thus, the employment share of a sector is a more precise way to indicate the scale of this industry.

3. Specialization index—Herfindahl measure

I construct the specialization index of the manufacturing sector (**SIM_M90**) by using the Herfindahl Index (H):

$$H = \sum_{i=1}^{n} P_i^2 \quad \left(\frac{1}{n} \leq H \leq 1\right) \text{ Where } \sum_{i=1}^{n} P_i^2 = \sum_{i=1}^{n} \frac{x_i^2}{X^2} \text{ and } X = \sum_{i=1}^{n} x_i$$

where n represents the total number of individuals (firms) in the groups (counties) we are considering, and n_i represents the number of individuals in group i, and the sum is taken over the total number of groups; therefore, pi represents the relative proportion of individuals in group

[8] http://www.bea.doc.gov/bea/regional/docs/cd.asp
[9] http://www.census.gov/epcd/cbp/view/cbpview.html

i. The Index ranges from 1/n to 1—the closer to 1 the more concentrated the group values, in other words, the more specialized the manufacturing sector.

4. Labor-market conditions and labor mobility

I use the natural logarithm of average wage for employees of the county and unemployment rate for residents of the county in 1990 as indicators for the labor-market condition. For the wage-per-job data, I use the Regional Economic Information System (REIS[10]) 1969-2003 data, while I use 1990 county unemployment-rate[11] data from the Bureau of Labor Statistics (BLS).

Labor mobility is measured by the percentage of people living in the same county since 1985, which is provided in the 1990 decennial Census[12].

5. Natural amenities and market size

I use natural amenities scale data[13] developed by Economic Research Service/U.S. Department of Agriculture (ERS/USDA) in 1993 as an indicator of the natural amenities level in the county (**SCALE**). I can measure the market size (**METRO93**) roughly with the use of a dummy variable to indicate whether this county is a metropolitan county designated by OMB (Office of Management and Budget) in 1993. I indicate metropolitan counties as 1 and non-metropolitan counties as 0.

2.2 Empirical Model and Analytical Techniques

The techniques I used include shift-share analysis, ordinary least square (OLS), and spatial-lag regression.

2.2.1 Shift-Share Analysis

I use shift-share analysis to separate statistically in any time interval the component of a region's growth that reflects the activity-mix of the region from those components that reflect overall national growth rates and changes in the region's competitive position. Other things

[10] http://www.bea.doc.gov/bea/regional/docs/cd.asp
[11] http://www.bls.gov/lau/home.htm
[12] 1990 Summary Tape File 3 (STF 3) - Sample data, Detailed Tables
[13] The natural-amenity scale is a combining index for warm winter, winter sun, temperate summer, low summer humidity, topographic variation, and water area. http://www.ers.usda.gov/data/NaturalAmenities/

being equal, a region will grow faster if it specializes in "growth industries," just as it will tend to have a low wage level if it specializes in low-wage activities or a high skill level if it specializes in high-skill activities. But shift-share analysis does not tell us why regions grow or improve. It says nothing about how a region's ability to hold its share of existing activities or to attract new ones is affected by the region's economic structure.

2.2.2 Ordinary Least Squares (OLS) Regression

The specification I use for the analysis of the determinants and the spatial autocorrelation in the competitive employment growth is based on the factors I discussed in Chapter 1.

The competitive employment growth for each county (**CEG**) from 1990 to 2000 is explained by the initial conditions of the county: industrial mix (**MIN90**, **MANFC90**, **TRNSP90**, **WHTRD90**, **SERV90**, **GOV90**), diversity of the whole economy (**BEAGINI90**), diversity of the service sector (**SGINI90**), specialization of manufacturing industry (**SIM_M90**), labor-market condition (**LNWAGE90**, **UNEMPL90**, **PCTHSGRAD90**), labor mobility (**PSAMECNT90**), natural-amenities level (**SCALE**), and market size (**METRO93**).

Thus, I have included the most important determinants for explaining the employment growth that are identified in most of the regional literature. I fit the following model, Equation (2), for explaining CEG for the whole Appalachian region, metropolitan counties and non-metropolitan counties, the three sub-regions, and the distressed, transitional, and better performing counties:

$$
\begin{aligned}
CEG_i = \alpha_0 &+ \beta_0\, MIN90_i + \beta_1\, MANFC90_i + \beta_2\, TRNSP90_i + \beta_3 WHTRD90_i + \beta_4 SERV90_i + \beta_5 GOV90i \\
&+ \beta_6\, BEAGINI90_i + \beta_7\, SGINI90_i + \beta_8\, SIM\_M90_i \\
&+ \beta_9\, LNWAGE90_i + \beta_{10}\, UNEMPL90_i + \beta_{11}\, PCTHSGRAD90 + \beta_{12}\, PSAMECNT90_I \\
&+ \beta_{13}\, SCALE_i + + \beta_{14}\, METRO93_i + \varepsilon_i \qquad (2)
\end{aligned}
$$

2.2.3 Spatial-Weights Matrix

The main difference between the traditional OLS regression model and spatial model is that, the spatial model integrates a spatial-weights matrix in the equation. The motivation for incorporating such a matrix is that the rate of economic development of regions/areas, named "economic agents" by econometricians, are thought to be influenced by economic activity in proximate agents. The function of the spatial-weights matrix in the model is to identify which agents interact with one another. —that is, which agents are "neighbors" according to some connectivity metric. (Anselin, 2002)

Generally speaking, in order to construct the spatial-weights matrix, I consider that regions and areas interact through two kinds of connection:

A. Physical connection—where economic agents are interacting through the being neighbors, or through existing transportation network, such as highways, rails, ports and airport.

B. Socioeconomic connection—commuting tie within labor market areas; trade/supply chain; innovation/learning region. I can define the spatial-weights matrices by using some social or economic indices, such as commuter flows or trade flows. (Bao, 2004)

Spatial econometrics provides very little formal guidance in the choice of the "correct" spatial weights in any given application. As Anselin indicates, when an analyst focuses on a model for substantive spatial dependence, the spatial weights should match the spatial-interaction patterns suggested by the theoretical framework (for example, a spatial-reaction function implying a specific range of interaction). In practice, analysts can eliminate bad choices by using model-validation techniques, such as a comparison of goodness-of-fit, or cross-validation. "Fortunately, empirical investigations can increasingly exploit both time and space dimensions (spatial panel data analysis), which opens up a number of opportunities to relax the structure of the weights matrix and employ non-parametric or semi-parametric methods to estimate a generic covariance structure, avoiding some of the strong priors required in the cross-sectional setting". (Anselin 2002, p. 259)

In this study, in order to present the difference between physical connection and socioeconomic connectedness, I use both physical (geographically neighboring) and socioeconomic connection (connections through commuting activity) to construct two different spatial-weights matrices in the spatial-regression model: (1) the Queen contiguity-weights matrix (Queen) and (2) commuting-zones weights matrix (CZs).

Actually, I could use GeoDa to create two contiguity-based spatial weights: Rook Contiguity and Queen Contiguity. They are similar in indicating the "physical" neighbors of counties. The only difference between them is that Rook Contiguity defines neighbors by including counties that share a common boundary while Queen Contiguity adds counties that touch only at a single point. (Anselin 2003b). In my study, I use the Queen-Contiguity definition to create the spatial weights for the "physical" connection of counties, because many analysts prefer the Queen-Contiguity to account for a more extensive definition of "neighbors" (Anselin 2002). All weights of 410 counties form the "Queen" spatial weights matrix in this study although in the case of the Appalachian Region, the Rook-Contiguity method will not affect much the scope of neighboring because few neighboring counties touched only at isolated points. There are virtually no differences between the two spatial weights for Appalachia.

For the socioeconomic connection of counties, I use the 1990 commuting-zone delineation of counties developed by ERS/USDA to define connectedness of labor. In other words, I consider all counties designated as being in the same commuting zone as "neighbors" in the "CZs" spatial-weights matrix. The motivation of delineating commuting zones is to define local labor market areas based on commuting ties among counties, which are economically meaningful (Killian, and Tolbert, 1993).

According to the documentation provided by the ERS/USDA[14], the delineation of commuting zones begins with counties as basic building blocks of commuting zones, as do those who construct MSAs (metropolitan statistical areas), BEA county groups, BTAs (Basic Trading Areas), and aggregated counties solely on the basis of commuting ties. The delineation does not consider absolute population size until later in the process. Also, the delineation is not confined to state boundaries. In other word, counties are grouped without regard for state lines, which makes the delineation more robust in reflecting the socioeconomic connections among counties.

In my study, because I am just studying those 410 counties within the Appalachian Region, for those counties along the border, I cannot catch the spatial effect from the physical counties outside the region. For the socioeconomic connection aspect, the same issue exists for those counties in a commuting zone that crosses the boundary of Appalachia because the information of counties in the same commuting zone but located outside the region is missing. Thus, my results may be biased to some extent.

2.2.4 Spatial Autocorrelation

Spatial autocorrelation means the value at any one point in space is dependent on values at the surrounding points. That is, the arrangement of values is not just random spatially. Positive

---

[14] http://www.ers.usda.gov/briefing/rurality/LMACZ/LMACZ1990.pdf

spatial correlation means that similar values tend to be near each other. Negative spatial correlation means that different values tend to be near each other. The existence of spatial autocorrelation is caused by spatial interaction/spillover effects.

In spatial econometrics, the spatial autocorrelation can be measure by Moran's I coefficient[15]. It is constructed by a variable and its spatially lagged transformation, after standardizing the variable such that the mean is zero and variance is one. The expression of the spatial lag for a standardized variable as z_i is

$$[w_z]_i = \sum_j w_{ij} z_j$$

where w_{ij} are elements of a row-standardized spatial weights matrix. For all z_i with a row-standardized spatial weights matrix, Moran's I coefficient of spatial autocorrelation is expressed by the following formula

$$I = \frac{\sum_i \sum_j z_i w_{ij} z_j}{\sum_i z_i^2}$$

or the slope of the regression line of the spatially lagged variable $[w_z]_i$ on the original variable, z_i (Anselin 1996).

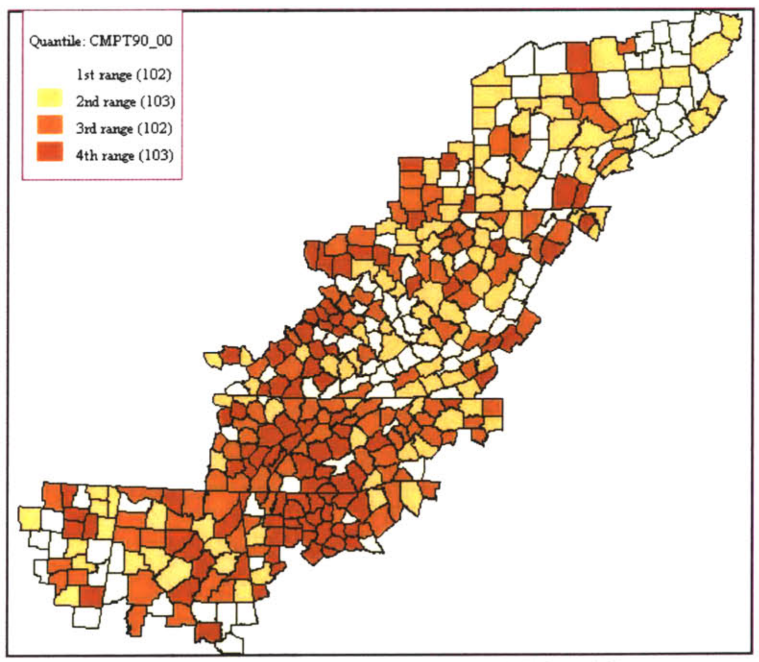
A positive Moran's I coefficient means that the observed values of locations within a certain distance or defined as "neighbors" in the spatial-weights matrix w_{ij} tend to be similar; it is negative when they tend to be dissimilar. In other words, the sign decides the pattern of similarity across space. The absolute value of Moran's I coefficient indicates the strength of spatial autocorrelation. Generally speaking, the higher the absolute value is, the stronger the spatial autocorrelation. If it is approximately zero, the observed values are arranged randomly and independently over space (Goodchild, 1986).

[15] There are numbers of ways to measure it. The most common ways are Moran's I and Geary's C Statistics. I choose Moran's I because the software GeoDa can compute it directly.

A. Visualizing and testing spatial autocorrelation

With Geoda[16], an analyst can conduct tests and visualization of both global (test for clustering) and local (test for clusters) using Moran's I statistic. The global Moran's I statistic accounts for spatial autocorrelation effects of all observations in the study area, while the local Moran statistic represents the spatial effect confined to those neighboring observations. Analysts visualize the global test by means of a Moran scatterplot (Anselin 1996), in which the slope of the regression line corresponds to Moran's I. Significance is based on a permutation test, which is a special case of randomization tests, i.e. tests that use randomly generated numbers for statistical inference. An analyst visualizes the local Moran statistic (Anselin 1995) in the form of significance and cluster maps. Here, I use the CEG variable as an example to demonstrate these methods.

Figure 2.1: Quartile Map for Competitive Employment Growth in Appalachian Region, 1990 to 2000



Source: Regional Economic Information System (REIS), 1990

Figure 2.1 shows the spatial distribution of the competitive employment growth rate in Appalachia. The darker color represents the higher growth rates. As shown in the map, most of the faster-growing counties are concentrated in Southern Appalachia.

Table 2.1 shows basic statistics for both the unadjusted employment growth rates (EG) and the competitive employment growth rates (CEG). The average employment growth rate (EG) from 1990 to 2000 in

[16] Geoda 9i—GeoDa is the latest incarnation in a long list of spatial software tools developed by Dr. Luc Anselin's Spatial Analysis Laboratory (SAL) in the Department of Geography at the University of Illinois, Urbana-Champaign.

Appalachian counties is 21%, and the standard deviation is 23%. The mean of CEG is -15.9%, and the standard deviation is 25%. Both EG and CEG is highly varied among the 410 counties. For CEG, it is ranging from a minimum -58% (far lower than the national average: 0) to the maximum 168%.

Table 2.1: Descriptive statistics of total employment growth rate and its competitive component from 1990 to 2000

Variable	# of Counties	Mean	Std. Dev.	Min	Max
Total Employment Growth Rate	410	21%	0.23	-27%	166%
Competitive Employment Growth Rate	410	-16%	0.25	-58%	168%

Source: Regional Economic Information System (REIS) 1969-2003, 1990 and 2000. Percentage calculated by the author. http://www.bea.doc.gov/bea/regional/docs/cd.asp

a. Global test: Moran scatterplot

In Geoda, the visualization of spatial autocorrelation is through the Moran scatterplot and cluster map (Anselin 2003a), which I use later in Chapter 4, Section 4.2. Basically, the main function of the Moran scatterplot is to classify the spatial autocorrelation into two categories, referred to as spatial clusters and spatial outliers. Each of them has two sub-categories. The cluster map, which I give later, shows the geographical distribution of these four kinds of spatial autocorrelation.

Figure 2.2: Example of Moran Scatterplot

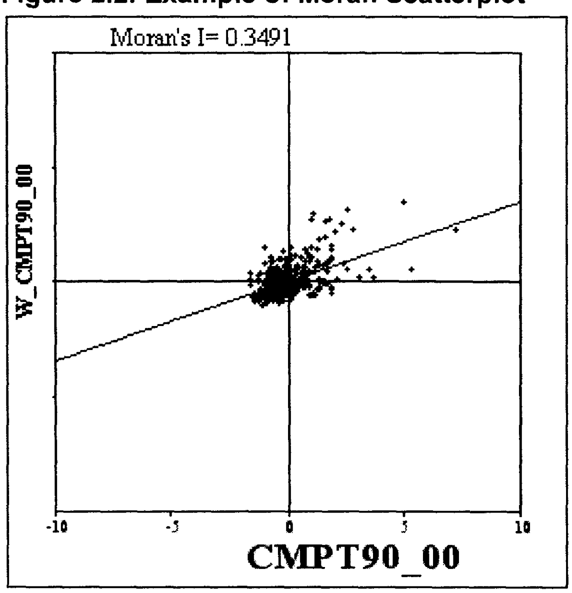


Figure 2.2 is an example of a scatterplot of the dependent variable of 410 counties—CEG (competitive employment growth from 1990 to 2000). The original variable (**CMPT90_00**) is plotted on the X-axis, and the weighted average of its neighbors (**W_CMPT90_00**) is plotted on the Y-axis. As explained in more detail in Anselin (1996), each quadrant of the Moran scatterplot corresponds to a different type of spatial correlation. The lower-left and upper-right quadrants indicate positive spatial autocorrelation, respectively of low values surrounded by neighboring low values, or high values surrounded by neighboring high values. Consequently, counties plotted within these quadrants are referred to as clusters of counties

that have low/low or high/high growth rates. In contrast, the upper-left and lower-right quadrants suggest negative spatial autocorrelation, respectively of low values surrounded by neighboring high values, or high values surrounded by neighboring low values. These are therefore referred to as counties that are spatial outliers.

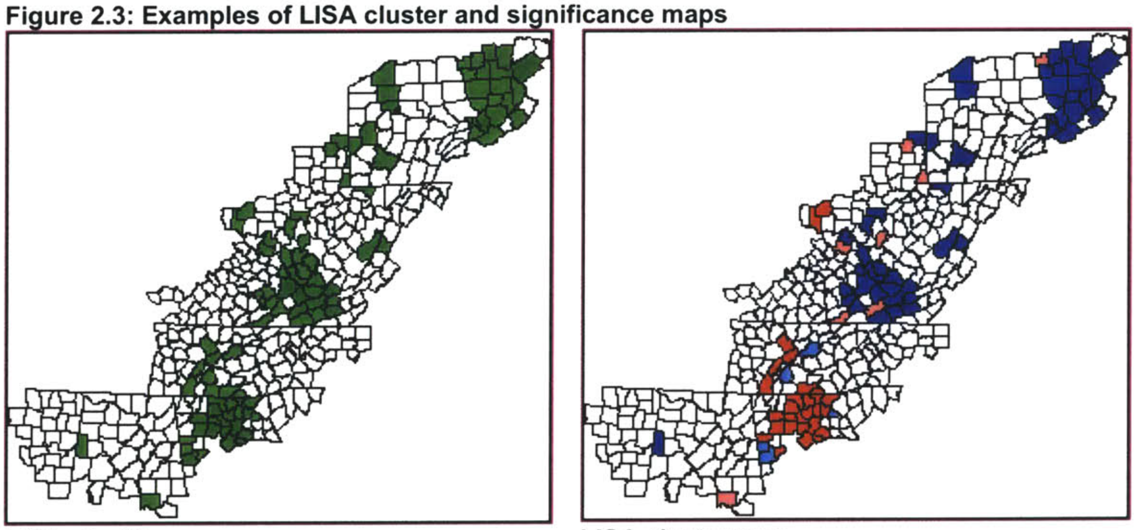
The scatterplot also provides a visual indication of the sign and strength of spatial autocorrelation in the form of the slope of the regression line, which is shown on top of the graph. Thus, the scatterplot allows for an informal investigation of the leverage (influence) of specific observations (locations) on the autocorrelation measure.

It is important to note that the scatterplot provides the classification, but does not indicate "significance". The latter is obtained by applying a Local Moran (LISA) test, as shown in Anselin (1995).

b. Local test: significance and cluster maps

The local tests of spatial autocorrelation are visualized by maps depicting the locations with significant Local Moran statistics (LISA[17] significance maps) and classifying those locations by type of association (LISA cluster maps).

Figure 2.3: Examples of LISA cluster and significance maps



LISA significance map **LISA cluster map**

[17] LISA: Local indicators of spatial association

The significance map on the left-hand panel of Figure 2.3 indicates counties with a local Moran statistic significantly different from zero. Significance is indicated by darker shades of green, with the darkest corresponding to $p = 0.05$ (based on 9999 random permutations of the observed CEG values across the 410 counties). A tighter significance criterion will eliminate some (but not many) locations from the map.

In the matching cluster map in the right-hand panel of Figure 2.3, the dark red and dark blue locations are indications of spatial clusters (respectively, high surrounded by high, and low surrounded by low). In contrast, the light red and light blue are indications of spatial outliers (respectively, high surrounded by low, and low surrounded by high).

Further discussion of the spatial exploratory analysis for CEG is provided in Chapter 4, Section 4.2.1.

2.2.5 Spatial-Regression Models

There are two ways in which I incorporate spatial interaction in the model specifications: a spatial-lag model and spatial-error model (Anselin 1988). Briefly speaking, the difference between them is that, for the spatial-lag model, an analyst incorporates a spatially lagged dependent variable (WY) on the right-hand side of the regression model, while in the spatial-error model, an analyst either models the spatial-error autocorrelation directly, following the general principles of geostatistics, or by utilizing a spatial-autoregressive process for the error term (for a recent review of these models, see Anselin and Bera, 1998; Anselin, 2001b)

To be more precise, the **spatial-lag model** is a spatial-reaction function that expresses how the magnitude of the dependent variable for one county depends on the magnitudes of the dependent variable for neighboring counties (Brueckner, 2002). This provides the theoretical basis for a so-called spatial-lag model, or, mixed regressive spatial autoregressive model (Anselin, 1988):

$$Y=\rho WY+\beta X+\varepsilon,$$

Where, Y is an $n\times 1$ vector of observations on the dependent variable, W is an $n \times n$ spatial-weights matrix that formalizes the network of directly interacting counties, ρ is the spatial autoregressive parameter, X is the usual $n \times k$ matrix of observations on the exogenous variables, with an associated $k \times 1$ regression coefficient vector β, and ε is a vector or random error terms.

The **spatial-error model** is a second form of spatial dependence, whereby the only interaction among counties would be through correlated error terms. Hence,

$Y = \beta X + \varepsilon$ with $\varepsilon = \lambda W \varepsilon + \upsilon$

where λ is a spatial-autoregressive coefficient and υ is a standard spherical error term, and the actual error is a linear combination of u and a weighted sum of errors in neighborhing counties. Ignoring spatial dependence in the error term does not lead to biased least-squares estimates, but the estimate of their variance will be biased, yielding a misleading inference (for further discussion, see, among others, Anselin, 1988a).

Thus far, I have explained the measurement of all variables in my empirical model and the details of the major method that I will use in the following analysis. In Chapter 3, I provide the descriptive statistics for the variables in the model. Then, in Chapter 4, I analyze the result of the regressions and compare the coefficients and significant level of variables among groups of counties within each classification. Also, I examine and compare the Ordinary Least Squares (OLS) model and spatial-lag model in Chapter 4.

Chapter 3

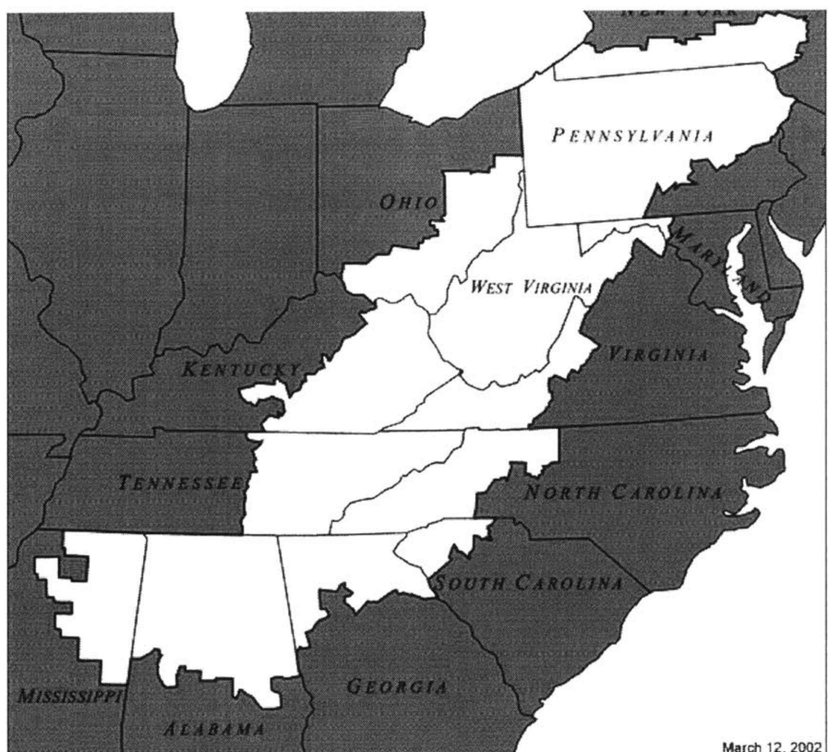
Descriptive Statistics of Variables

Before interpreting the regressions I did for the industrial structure and other determinants on competitive employment growth for Appalachian counties, I provide descriptive statistics for the dependent variable (competitive employment growth) and its determinants. First, I provide information about the geographical boundary of the Appalachian Region and its basic social characteristics. Then, in order to account for the heterogeneity of the Appalachian Region, I analyze the statistic for the dependent variable (competitive employment growth) and its determinants based on the following three classifications used by most analysts of the Appalachian Region: (1) metropolitan and non-metropolitan counties, (2) three sub-regions: Northern, Central, and Southern, and (3) distressed counties, transitional counties, competitive and attainment counties.

3.1 Geographical Definition of the Appalachian Region

Appalachia refers to the area of and around the Appalachian Mountains. The Appalachian mountain chain is the major mountain system of eastern North America, covering more than 1, 500 miles of territory from the Canadian province of Quebec to northern Alabama. The definition of Appalachia used in this study comes from the Appalachian Regional Commission (ARC)[18], created by the Appalachian Regional

Figure 3.1 Appalachian Region Boundaries



Source: Appalachian Regional Commission

18 The ARC is a partnership between the federal government and the governments of 13 states, whose purpose is to improve conditions in the Appalachian region.

Development Act of 1965 (and most recently reauthorized in 2002). The region covers 410 counties[19] in the following 13 states (Figure 3.1): all of West Virginia, as well as southern New York; most of Pennsylvania; southeastern Ohio; the western portions of Maryland, Virginia, and the Carolinas; the eastern portions of Tennessee and Kentucky; the northern portions of Georgia and Alabama; and northeastern Mississippi.

1. <u>Employment growth and its competitive component</u>

The total U.S. private employment in 1990 was 115 million, and it increased by 22.4% to 140.7 million in 2000 (Table 3.1). For the Appalachian Region, the private employment[20] growth rate from 1990-2000 is 19.5%, which is 3 percentage points lower than the U.S. average.

Table 3.1: Total employment and growth for the U.S. and Appalachian Region, 1990 and 2000

	United States(in millions)			**Appalachian Region(in millions)**		
	1990	2000	Growth	1990	2000	Growth
Total Private Employment	115.00	140.70	22.4%	8.53	10.19	19.5%

Source: Regional Economic Information System (REIS), 1990 and 2000, percent calculated by the author

As previously shown in Table 2.1, the mean of the competitive employment growth (CEG) rate from 1990 to 2000 in the Appalachian Region is -15.9%. Thus, the employment growth rates in Appalachian counties are 15.9% lower than the U.S. average. The CEG rate also has a high variation among the 410 counties, from the minimum -0.58 to the maximum 1.68.

The frequency distribution of the CEG rate[21] (Figure 3.2) shows that the majority of Appalachian counties (over 320 counties)

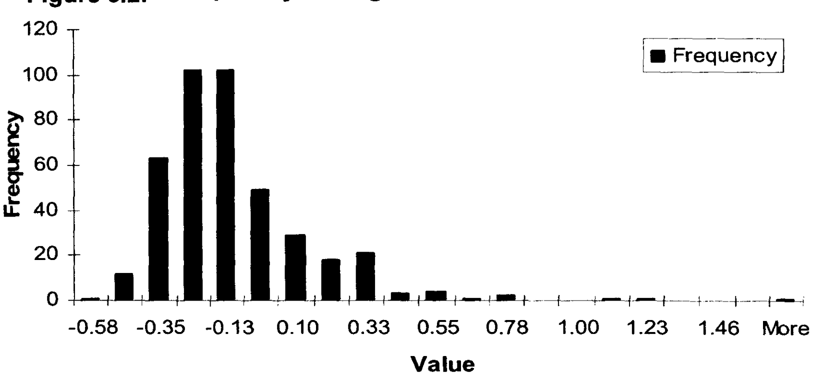


Figure 3.2: Frequency Histogram of CEG for 410 Counties

Source: Regional Economic Information System (REIS), 1990 and 2000

---

[19] In addition to the 410 counties, the Appalachian region contains eight Virginia cities that are independent of any county authority (that is, they function like counties). For analytical purposes, the Appalachian Regional Commission—following the practice of the U.S. Bureau of Economic Analysis (BEA)—incorporates each independent city within an adjacent county, and this report follows that practice.
[20] Exclude governmental sector.
[21] Value shown in the Figure 3.2 is ratio of year 2000 to year 1990.

have negative competitive employment growth (relative to the US average for the decade).

I showed the spatial distribution of the CEG rate in Appalachia in Figure 2.1. From that map, we notice that most of the faster-growing counties are concentrated in Southern Appalachia, spatially near Atlanta. For details of the top/bottom 20 counties, refer to Appendix A.

2. Industrial mix

Employment grew differently from 1990 to 2000 in Appalachia when compared with the United States for each sector. Table 3.2 lists the number of employees in 1990 and 2000 and the growth rate in private employment by sectors for the United States and the Appalachian Region.

Notably, employment in the agriculture sector increases by 46% from 1990 to 2000 in the United States, while in Appalachia, the employment growth of this sector is negative. The mining sector also dramatically shrank in Appalachia during this period.

Although total employment grew slower in the Appalachia Region (AR) than in the United States (US), some sectors greatly outperformed others in employment growth. For example, the wholesale trade, retail trade, and finance, insurance, and real estate (FIRE) sectors grew faster than the U.S average level. (Wholesale trade: U.S. 12.9% vs. AR 17.4%; Retail trade: U.S. 18.9% vs. AR 20.9%; FIRE: U.S. 23.1% vs. 30.5%).

Table 3.2: Industrial mix and growth rate of the U.S. and the Appalachia Region, 1990-2000

		US (in millions)			AR (in millions)		
Sectors	**SIC code**	**1990**	**2000**	**Growth**	**1990**	**2000**	**Growth**
Total private employment		115.00	140.70	22.4%	8.53	10.19	19.5%
Agricultural services, forestry,fishing&others	**0**	1.45	2.12	45.9%	0.07	0.07	-5.2%
Mining	**11-14**	1.04	0.78	-24.9%	0.15	0.08	-44.4%
Construction	**15-17**	7.26	9.45	30.1%	0.58	0.74	27.5%
Manufacturing	**2-3**	19.69	19.11	-2.9%	2.04	1.94	-4.5%
Transportation and public utilities	**4**	6.55	8.24	25.9%	0.46	0.56	21.9%
Wholesale trade	**50-51**	6.72	7.58	12.9%	0.42	0.49	17.4%
Retail trade	**52-59**	22.89	27.22	18.9%	1.76	2.12	20.9%
Finance, insurance, and real estate	**6**	10.71	13.19	23.1%	0.53	0.69	30.5%
Services	**7-8**	38.67	52.99	37.0%	2.46	3.35	36.0%

Source: Regional Economic Information System (REIS) 1990 and 2000, percent calculated by the author

3. <u>Diversity indices and specialization index</u>

Table 3.3 shows the descriptive statistics for diversity and specialization indices. The diversity of the whole economy (**BEAGINI90**) is less varied (StdDev/Mean = 0.074/0.50) than the diversity of the service sector (**SGINI90**) among counties (StdDev/Mean = 0.163/0.51). The variance of manufacturing specialization (**SIM_M90**) is also very large (Std.Dev. = 0.32 with a mean of 0.60).

Table 3.3: Industrial diversity and specialization of Appalachian counties, 1990

Variable	#of Observations*	Mean	Std. Dev.	Min	Max
BEAGINI90	399	0.499	0.074	0.000	0.680
SGINI90	399	0.506	0.163	0.000	0.781
SIM_M90	312	0.597	0.323	0.100	1.000

Source: County Business Pattern (CBP) 1990, numbers are calculated by the author.
Note: BEAGINI90: Diversity index of the whole economy; SGINI90: Diversity index of service sector; SIM_M90: Manufacturing specialization indicator.
*Counties are excluded from this summary statistics analysis, because CBP did not provide the data in the service/manufacturing sector for these counties to avoid disclosure of confidential information.

4. Labor-market conditions, labor mobility, amenities, and market size

From the statistical summary in Table 3.4, we confirm that Appalachian counties in 1990 are diverse in many aspects. For example, the unemployment rate (**UNEMPL90**) varied from 2% to 22%; the percentage of residents with high school education (**PHSGRAD90**) ranged from 35% to 87%. The "scale" measure for natural amenities is a standardized score using a scale that has a mean of 0 and standard deviation of 1.0 across all US Counties. I set the metro93 dummy variable to 1.0 for the 109 counties (26.6%) that were classified as metropolitan counties in 1993.

Table 3.4: Descriptive statistics of other variables for Appalachian counties, 1990

Variable	# of Obs	Mean	Std. Dev.	Min	Max
UNEMPL90	410	8.1%	3.4%	2.7%	21.9%
LNWAGE90	410	9.752	0.186	9.310	10.346
PHSGRAD90	410	61.2%	10.2%	35.5%	87.2%
PSAMECNT90	410	83.8%	6.6%	53.9%	94.7%
SCALE	410	0.132	1.163	-3.720	3.550

Source: Regional Economic Information System (REIS), 1990; Census 1990; natural amenities scale from ERS/USDA, numbers are calculated by the author

3.2 Classification 1—Metropolitan and Non-metropolitan Counties

According to the amount of population and its distribution, Appalachia is not very rural. More than three-fifths of the Appalachian population lives in metropolitan areas; about one-fourth resides in large metropolitan (metro) areas such as Atlanta, Pittsburgh, and Birmingham. However, many Appalachian areas remain sparsely populated. For example, nearly half of Appalachia's 410 counties had fewer than 30,000 people in 2000; 33 of those counties had fewer than 10,000 residents. (Pollard 2005)

The distinguishing characteristics and different development pattern of metro and non-metro counties in the United State remain or maybe are even more significant in the Appalachia Region. Here, I just stress those features relevant to this study.

1. Employment growth and its competitive component

It is not surprising to see that the average employment growth of metro counties is higher than non-metropolitan counties, both in terms of gross growth or its competitive component. The variance of metro counties is bigger than for their non-metro counterparts.

Table 3.5: Employment growth rate and its competitive component for Metropolitan counties and non-metropolitan counties, 1990-2000

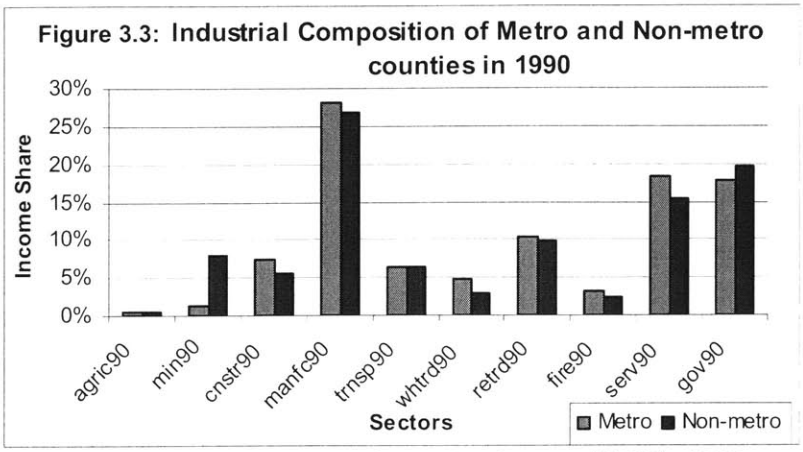
	Variable	# of Obs	Mean	Std. Dev.	Min	Max
Metropolitan	**GR90_00**	109	27.4%	28.3%	-0.103	1.669
	CMPT90_00	109	-13.4%	28.4%	-0.459	1.197
Non-metropolitan	**GR90_00**	301	20.1%	24.1%	-0.277	2.163
	CMPT90_00	301	-16.8%	24.1%	-0.578	1.683

Source: Regional Economic Information System (REIS), 1990 and 2000, numbers are calculated by the author

Note: GR90_00: employment growth rate, 1990 to 2000; CMPT90_00: competitive employment growth rate, 1990 to 2000.

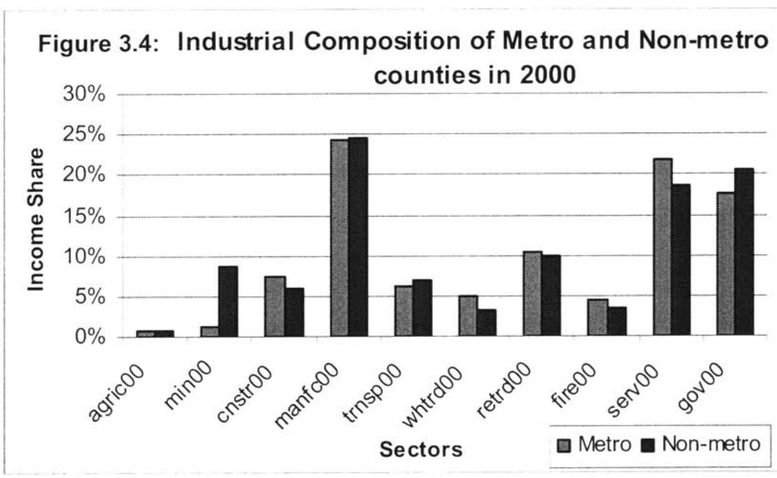
2. Industrial mix

Figure 3.3 shows the industrial composition of metro and non-metro counties in Appalachia. The mining sector (**MIN90**) is overrepresented in non-metro counties, while other sectors have similar shares. The manufacturing sector dominated in the economy, followed by government sector (**GOV90**) and services sector (**SERV90**).



Source: Regional Economic Information System (REIS), 1990

Comparing the industrial composition between 1990 and 2000 (Figure 3.4), I notice that the share of manufacturing sector (**MANF90** and **MANF00**) declines slightly (from around 27% to below 25%) while the share of the services sector (**SERV90** and **SERV00**) increases.



Source: Regional Economic Information System (REIS), 2000

3. <u>Diversity indices and the specialization index</u>

In Table 3.6, I show that the average level of manufacturing specialization (**SIM_M90**) of non-metro counties is higher than the one of metro counties. Interestingly, the table shows that the mean of the diversity index for the whole economy (**BEAGINI90**) of non-metro counties is higher than that of metro counties, The difference of 0.014 is statistically significant (α= 0.1). This finding means that non-metro economies are more diverse than metro economies in Appalachia.

Table 3.6: Industrial diversity and specialization of metropolitan counties and non-metropolitan counties, 1990

Variable	# of Observations		Mean		Std. Dev.	
	Metro	Non-metro	Metro	Non-metro	Metro	Non-metro
BEAGINI90	97	237	0.489	0.503	0.073	0.073
SGINI90*	97	237	0.486	0.513	0.209	0.141
SIM_M90*	97	237	0.403	0.677	0.292	0.301

Source: County Business Pattern (CBP) 1990, numbers are calculated by the author.
Note: BEAGINI90: Diversity index of the whole economy; SGINI90: Diversity index of service sector; SIM_M90: Manufacturing specialization indicator.
*76 Counties are excluded from this summary statistics analysis because CBP did not provide the data in the service/manufacturing sector for these counties to avoid disclosure of confidential information.

4. Labor-market conditions, labor mobility, and amenities

After conducting statistical tests, I find that all the means of variables in Table 3.7 are significantly different ($\alpha = 0.05$) between metro and non-metro counties. On average, metro counties have higher wage (LNWAGE90), lower unemployment rate (UNEMPL90), and more educated labor force (PHSGRAD90) in 1990 but their variances are less than non-metro counties. It is interesting that the natural-amenities indicator (SCALE) is better in metro counties than their non-metro counterparts.

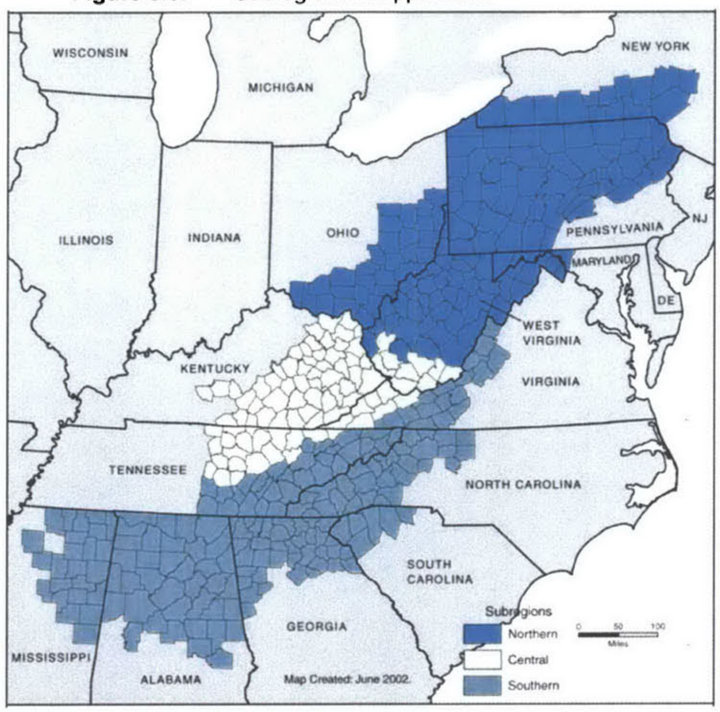
Table 3.7: Descriptive statistics of other variables for Metropolitan and Non-metropolitan Counties, 1990

Variable	# of Observations		Mean		Std. Dev.	
	Metro	Non-metro	Metro	Non-metro	Metro	Non-metro
lnwage90	109	301	9.851	9.716	15.8%	18.3%
unempl90	109	301	6.4%	8.6%	1.789	3.622
phsgrad90	109	301	68.3%	58.6%	7.7%	9.7%
psamecnt90	109	301	81.8%	84.6%	7.5%	6.1%
scale	109	301	0.332	0.060	0.930	1.230

Source: Regional Economic Information System (REIS), 1990; Census 1990; natural amenities scale from ERS/USDA, numbers are calculated by the author

3.3 Classification 2—Three Subregions

Figure 3.5: Subregions in Appalachia



Source: ARC website

The Appalachian Regional Commission (ARC) divides the region into three subregions based on the location of the county: Northern, Central, and Southern. (Figure 3.5) Northern Appalachia contains every Appalachian county in New York, Pennsylvania, Maryland, and Ohio, as well as 46 of West Virginia's 55 counties. Central Appalachia includes West Virginia's nine southernmost counties, all of Appalachian Kentucky, the southwestern tip of Virginia, and the northwest part of Tennessee's Appalachian area. Finally, southern Appalachia includes most of the Appalachian portion of Virginia and Tennessee, as well as the entire Appalachian sections of the Carolinas, Georgia, Alabama, and Mississippi (Appalachian Regional Commission, 1974).

1. Employment growth and its competitive component

Table 3.8: Employment growth and its competitive component of three sub-regions, 1990-2000

Variable	#of Observations			Mean			Std. Dev.		
	N	C	S	N	C	S	N	C	S
GR90_00	144	87	179	15.2%	19.2%	28.9%	15.5%	23.9%	30.5%
CMPT90_00	144	87	179	-23.2%	-18.3%	-9%	18.2%	23.4%	29.2%

Source: Regional Economic Information System (REIS), 1990 and 2000, numbers are calculated by the author. N: northern region; C: central region; S: southern region

After conducting the statistics test, I find that the mean of the northern, central and southern regions are statistically significantly different from each other (α = 0.00).

These summary statistics Table 3.8 show that

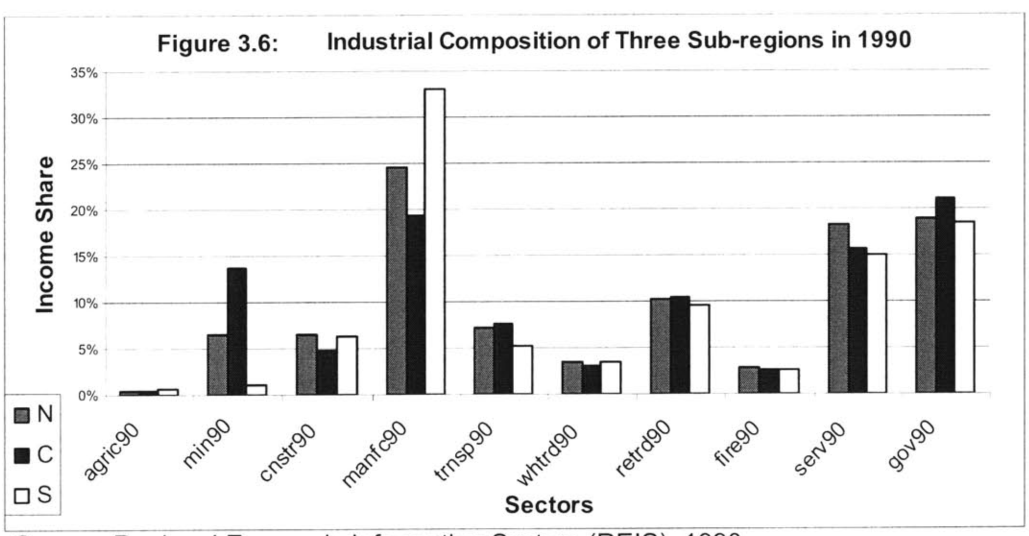
First, on average, gross employment growth in the Southern subregion is the best among the three regions. I notice that employment in Central Appalachia grew faster than that in the Northern subregion, although the Central counties were thought to be most economically distressed. (Isserman, 1996) This result might occur because a region with a weak initial economic base has more room to grow.

Second, although all three subregions have a negative employment growth from 1990 to 2000 (in terms of the competitive component), the southern region has the lowest absolute value, -9%. It is consistent with the fact found by the Brandow Company's (2001) research on number of establishments in Appalachia. The Southern subregion has the highest birth rate of establishments, followed by the Central and Northern subregions. They indicate that the Southern subregion's establishment birth and death rates and job-creation and destruction rates are more closely aligned with those of the United States as a whole than with either of the other two subregions.

2. Industrial mix

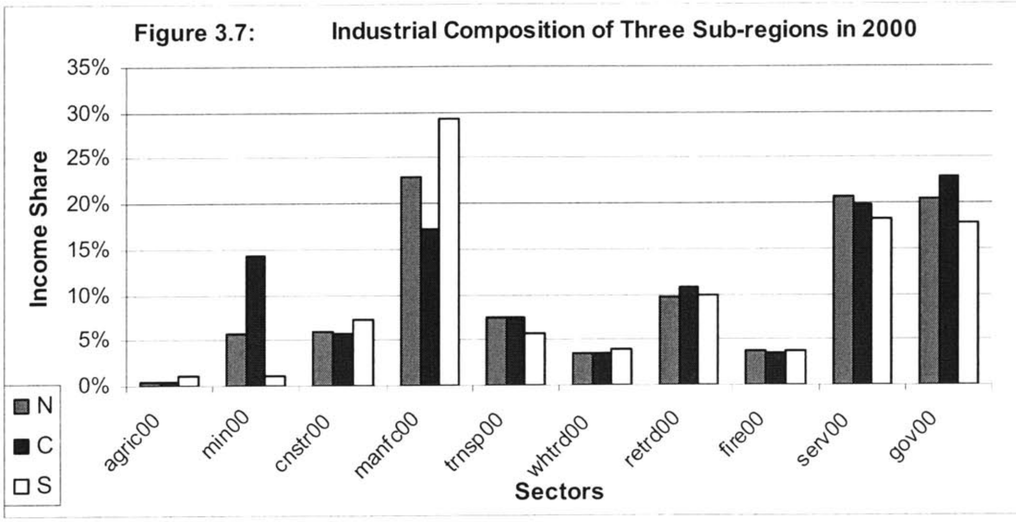
Figure 3.6 shows that the Central subregion most closely approximates the conventional view of Appalachia, where most of its activity is in non-metro areas and it relies on mining and manufacturing.

Although, on average, the shares of manufacturing sector are highest among all sectors for three subregions, it is more dominant in Southern counties.



Source: Regional Economic Information System (REIS), 1990

From Figure 3.7, I find that the pattern did not change from 1990 to 2000.



Source: Regional Economic Information System (REIS), 2000

3. Diversity indices and specialization index

Table 3.9: Industrial diversity and specialization of three subregions, 1990

Variable	#of Observation			Mean			Std. Dev.		
	N	C	S	N	C	S	N	C	S
BEAGINI90	117	59	158	0.487	0.495	0.509	0.078	0.064	0.074
SGINI90*	117	59	158	0.509	0.500	0.504	0.164	0.145	0.168
SIM_M90*	117	59	158	0.514	0.752	0.601	0.328	0.286	0.312

Source: County Business Pattern (CBP) 1990, numbers are calculated by the author.
Note: BEAGINI90: Diversity index of the whole economy; SGINI90: Diversity index of service sector; SIM_M90: Manufacturing specialization indicator.
*76 Counties are excluded from these summary statistics analysis because CBP did not provide the data in the service/manufacturing sector for these counties to avoid disclosure of confidential information.

Using the statistics test, I find that the means of the diversity level (**BEAGINI90** and **SGINI90**) are actually not different among the northern, central, and southern subregions, but the mean of the manufacturing-specialization level index (**SIM_M90**) for Central Appalachian is significantly higher ($\alpha = 0.01$) than ones for the Northern and Southern subregions. This finding means that the diversification of the economy and the service sector is similar among the three subregions, while Central Appalachia counties have the highest level of manufacturing specialization, followed by Southern and Northern Appalachia.

4. Labor-market conditions, labor mobility, and amenities

Table 3.10: Labor-market conditions, mobility, and amenities of three subregions, 1990

Variable	#of Observation			Mean			Std. Dev.		
	N	C	S	N	C	S	N	C	S
lnwage90	144	87	179	9.834	9.680	9.722	0.154	0.226	0.165
unempl90	144	87	179	8.57%	10.85%	6.28%	3.069	3.760	2.169
phsgrad90	144	87	179	69.5%	50.0%	60.0%	0.069	0.073	0.074
psamecnt90	144	87	179	84.5%	86.8%	81.8%	0.064	0.050	0.067
scale	144	87	179	-0.361	-0.220	0.700	0.877	1.074	1.163
metro93	144	87	179	0.326	0.080	0.307	0.471	0.274	0.463

Source: Regional Economic Information System (REIS), 1990; Census 1990; natural amenities scale from ERS/USDA; numbers are calculated by the author.

Using the statistics test for the three subregions' means of the wage level in 1990, I find that the mean of the northern, central, and southern subregions are different from each other. I show that, on average, the Central subregion has the lowest wage level, while the Northern Appalachia has the highest wage level in 1990.

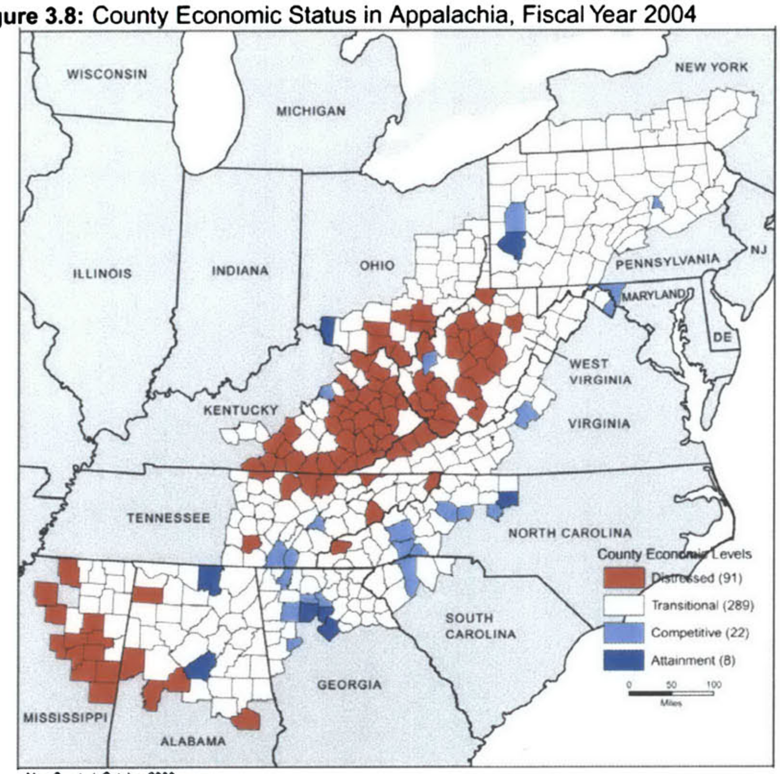
The result is not surprising when considering Jensen's (1998) findings. He found that relative to the United States, Northern Appalachia has lower entry rates and increasingly lower wages and productivity, the Central subregion has higher entry rates and lower (but relatively unchanged between 1963 and 1992) wages and productivity, and the Southern counties have higher entry rates and lower (but less so) wages and productivity.

Another interesting finding is, the natural amenities (**SCALE**) of southern counties are much better than those of the central and northern Appalachia. (0.7 vs. -0.22 and -0.36)

3.4 Classification 3—Overrepresented Distressed Counties in Appalachia

The economic status of Appalachia varies widely—some areas are as economically advanced as the United States as a whole, while many areas remain economically depressed. The Appalachian Regional Commission has developed a system that uses three indicators of economic viability—per capita income, poverty, and unemployment—to classify counties into the following four categories of economic development: distressed, transitional, competitive, and attainment counties. (Detail definition in Appendix C) Considering the limited observations and the similarity of the latter two categories, I combine the competitive and attainment counties and call them "better-performing" counties. (Figure 3.8)

Figure 3.8: County Economic Status in Appalachia, Fiscal Year 2004



Map Created: October 2003
Data Sources: U.S. Bureau of Labor Statistics, LAUS, 1999-2001;
U.S. Bureau of Economic Analysis, REIS, 2000;
U.S. Census Bureau, 2000 Census, SF3

Source: ARC website

1. Employment growth and its competitive component

Like the differences among three sub-regions, the differences among these three economic-development groups are also statistically significant. Of course, the competitive and attainment counties, on average, have the highest growth rate, and the distressed counties grew much slower than their regional counterparts (Table 3.11).

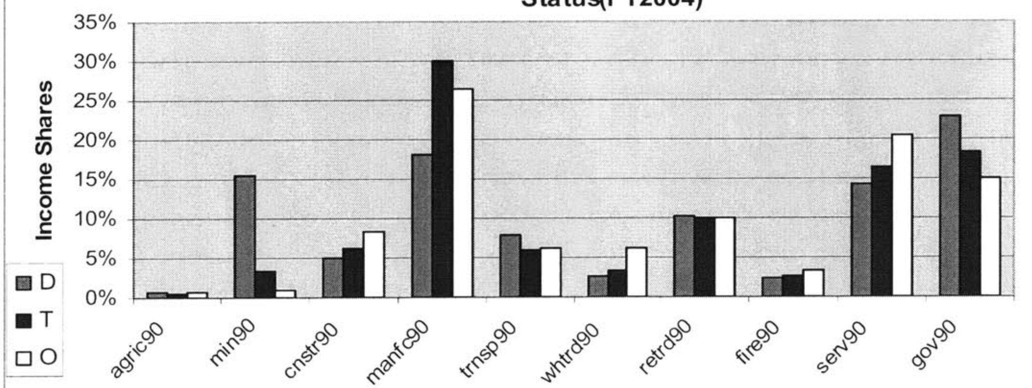
Table 3.11: Employment growth rate and its competitive component for three economic levels, 1990-2000

Variable	# of Observations			Mean			Std. Dev.		
	D	T	O	D	T	O	D	T	O
GR90_00	91	289	30	13.9%	21.8%	48.9%	0.214	0.217	0.454
CMPT90_00	91	289	30	-23.2%	-16.0%	6.4%	0.206	0.226	0.446

Source: Regional Economic Information System (REIS), 1990 and 2000
Note: D: distressed counties; T: transitional counties; O: competitive and attainment counties. Based on ARC designation of fiscal year 2004.

2. Industrial mix

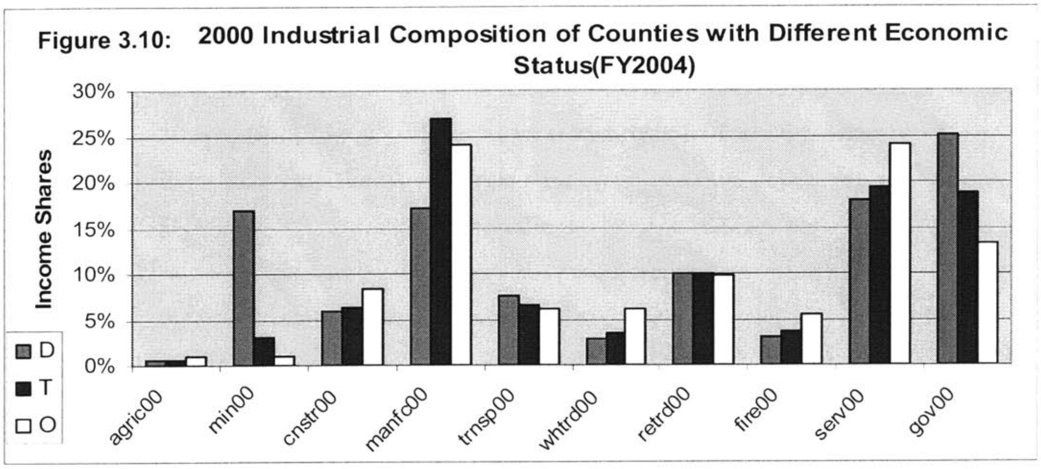
Figure 3.9: 1990 Industrial Composition of Counties with Different Economic Status(FY2004)



Source: Regional Economic Information System (REIS), 1990

In Figure 3.9, comparing the 1990 industrial composition of counties grouped to three kinds of economic status, I show that the mining sector is significantly overrepresented in distressed counties, followed by the governmental sector. The transitional counties have the highest average comparative income share (30%) in the manufacturing sector.

The same pattern still exists in the 2000 industrial composition (Figure 3.10).



Source: Regional Economic Information System (REIS), 2000

3. Diversity indices and specialization index

As observed in Table 3.12, different economic types of counties are similar both in terms of the whole economy and the services sector, while the distressed counties have the highest specialization level within the manufacturing sector.

Table 3.12: Industrial diversity and specialization of three economic levels, 1990

Variable	# of Observations			Mean			Std. Dev.		
	D	T	O	D	T	O	D	T	O
BEAGINI90	58	249	27	0.493	0.502	0.490	0.074	0.063	0.112
SGINI90*	58	249	27	0.512	0.513	0.422	0.148	0.153	0.244
SIM_M90*	58	249	27	0.873	0.559	0.352	0.203	0.312	0.269

Source: County Business Pattern (CBP) 1990, calculated by author.
*Note: Counties are excluded from this summary statistics analysis because CBP did not provide the data in the service/manufacturing sector for these counties to avoid disclosure of confidential information.

4. Labor-market conditions, labor mobility, and amenities

The distressed counties have the lowest wage level and the highest unemployment rate compared to their regional counterparts. Labor mobility and the natural-amenity scale are also highest in better-performing counties. (Table 3.13)

Table 3.13: Labor-market conditions, mobility, and amenities of three economic levels, 1990

Variable	# of Observations			Mean			Std. Dev.		
	D	T	O	D	T	O	D	T	O
lnwage90	91	289	30	9.679	9.763	9.868	0.226	0.167	0.144
unempl90	91	289	30	12.2%	7.1%	4.6%	3.537	2.172	1.065
phsgrad90	91	289	30	51.6%	63.3%	70.2%	0.080	0.089	0.077
psamecnt90	91	289	30	87.6%	83.4%	76.7%	0.048	0.062	0.076
scale	91	289	30	-0.302	0.219	0.606	1.066	1.165	1.084
metro93	91	289	30	0.022	0.287	0.800	0.147	0.453	0.407

Source: Regional Economic Information System (REIS), 1990; Census 1990; natural amenities scale from ERS/USDA

So far, I have examined the competitive employment growth, industrial structure, labor-market conditions, and natural amenities in the Appalachian Region as a whole and their characteristics between metropolitan counties and non-metropolitan counties, among Northern, Central, and Southern Appalachia, and among different economic-development groups.

In general, both of the actual employment growth (**ER90_00**) and the competitive employment growth (**CMPT90_00**) are significantly different (1) between metropolitan and non-metropolitan counties, (2) among distressed counties, transitional counties and better forming counties, and (3) among three subregions.

For most of the characteristics I discussed above, i.e., Industrial diversity and specialization, labor-market conditions, and amenities, there are strong differences (1) between metropolitan and non-metropolitan counties and (2) among different economic-development groups, but they are not significant among three subregions.

In Chapter 4, I use regression and spatial analyses for the empirical model and finally summarize my findings in the conclusion part.

Chapter 4

Regression and Spatial Analysis

In the previous chapter, I have carefully examined the descriptive statistics of the dependent variable and thirteen independent variables, and their differences among various groups of counties in Appalachia, i.e., metropolitan (metro) and non-metropolitan (non-metro) counties, three sub-regions, and different economic-development groups. In this chapter, I focus on the following simple linear regression model for the competitive growth rate, which is one possible specification of the general relationship specified by Eq. (2) in Section 2.2.2.

$$CEG_i=\alpha_0+ \beta_0\ MIN90_i +\beta_1\ MANFC90_i+ \beta_2\ TRNSP90_i +\beta_3 WHTRD90_i+ \beta_4 SERV90_i+ \beta_5 GOV90i +\beta_6\ BEAGINI90_i+ \beta_7\ SGINI90_i+ \beta_8\ SIM\_M90_i + \beta_9\ LNWAGE90_i +\beta_{10}\ UNEMPL90_i+ \beta_{11}\ PCTHSGRAD90+ \beta_{12}\ PSAMECNT90_I + \beta_{13}\ SCALE_i + + \beta_{14}\ METRO93_i +\varepsilon_i$$

Using the optimal least squares (OLS) method of fit, I fit the same model specification for each group of counties within the three types of county classification (metro, economic development and geographic region). After analyzing and comparing the results, I examine possible spatial autocorrelation effects and compare the OLS and spatial regression models.

4.1 OLS Regression Result Comparison among Different Groups

There are three types of county classifications used by most studies of the Appalachian Region: (1) metro and non-metro segregation, (2) three sub-regions (northern, central and southern Appalachian), and (3) county economic status classification (distressed, transitional and better-performing counties). In this section, I regress the specification on groups within each of the three classifications and compare the regression results.

From Table 4.1 (on page 54), I note that, the value of R^2 and adjusted R^2 are significantly different between metro and non-metro models: 0.73 and 0.69 for metro model; 0.24 and 0.19 for non-metro model. This means that the regression model has more explanatory power for the metro counties than for the non-metro counties. The model seems to explain the competitive employment competitive growth rate (1990-2000) better by the determinants for southern than for northern Appalachia[22]. When I classify Appalachian counties by their economic status in

[22] Industry mix and diversity do not appear to matter for central Appalachia and F test for central Appalachia is too large

2000, I find that the values of R^2 and adjusted R^2 for the better-performing county model are much higher than the distressed- and transitional-county models (0.82 and 0.61 vs. 0.41 and 0.18 vs. 0.38 and 0.34). In summary, it seems that this empirical model better explains the competitive employment growth from 1990 to 2000 by using the determinants in metro counties, or in southern Appalachia counties, or or in better-performing counties.

Besides the discrimination of explanatory power of the specification among county groups, both the significance levels (P>t.) and the value of the coefficients (Coef.) for those statistically significant variables also varied greatly among groups within each classification. I conduct the following comparison analysis for industrial mix, industrial diversity and specialization, labor-market conditions, and others, respectively.

4.1.1 Industrial Mix

When I examine the industrial-mix factors for metro and non-metro counties, the share of manufacturing, transportation, service, and governmental sectors in 1990 have significantly negative correlations with the growth rates from 1990 to 2000 for metro counties. I also notice that the mining sector matters only for non-metro counties. The reason might be that (1) the mining sector is small outside the mountainous central Appalachian Region; (2) the rural economy is too small and varied for other sectors to have a visible and consistent effect on employment growth across all rural Appalachian counties; (3) technically thinking, the significances of these variables may be diluted by the noise of other variables.

In the case of the three sub-regions, almost every sector's share I included in the model is significantly correlated with the growth rates for counties in southern Appalachia, while for counties in the central region, none of these is relevant. It might be because the economy of most counties in Central Appalachia is dominated by mining sector and their whole economy is very small. Therefore the impact of any single sector does not showed in the growth model.

For counties in northern Appalachia, the share of the wholesale trade sector in 1990 has a significantly positive effect (coef. =2.05) while that of the service sector has a significantly negative one (coef. = -0.65) on the county competitive employment growth rates. In comparing the impacts of wholesale trade sector between Northern and Southern Appalachia counties, I find a much larger coefficient of wholesale trade sectors for Northern Appalachia over Southern Appalachia, (2.05 vs. 1.39). The pattern reverses for services sector; the impacts of services sector are almost three times higher in counties of the southern region than those in the Northern Appalachia (-1.66 vs. -0.65).

I obtain a similar pattern when I compare counties at different economic levels. Almost all the shares of sectors I used in the model have significantly strong impacts on the growth rates for transitional counties, but the wholesale trade sector has no statistically significant effects. However, none of the shares are statistically significant for those distressed counties. For those better-performing counties, only the share of the services sector in 1990 matters[23] to their competitive employment growth rates (CEG) from 1990 to 2000.

4.1.2 Diversity Indices and Specialization Index

The industrial diversity indicators and specialization indicator also perform differently, but they are less diverse than industrial-mix factors among groups within each classification. Generally speaking, Diversity Indicator I (**BEAGINI90**), which measures the diversity of the whole economy in 1990, has a statistically significant negative impact (P<0.05) on the growth rates within non-metro counties. It is less significant for metro counties (P < 0.2) and for transitional counties (P < 0.10). This diversity indicator is not a significant factor in explaining the difference of growth rates among counties either within each sub-region or within distressed counties and better performing counties.

Diversity Indicator II (**SGINI90**), which measures the diversity of the service sector in 1990, only contributes to the model for the different growth rates among better-performing counties. It has a positive impact on the CEG rate (Coef. =0.50). This means that the more diverse services a better-performing county provided in 1990, the more was its competitive employment growth between 1990 and 2000. It makes sense because service diversity is one of the important indicators of economic health, as illustrated by previous studies (e.g., Quigley 1998).

As the specialization indicator, I used the Herfindahl measure for the manufacturing sector in each county, indicated by **SIM_M90** in the model. I found a statistically significant (P < 0.10) positive relationship between the level of manufacturing sector specialization and competitive employment growth only in Southern Appalachia counties and in the transitional counties.

[23] Only 27 counties are in this group and they include most of the half dozen outliers with especially high growth rates. So, instead of examining further at this point, I will discuss it until accounting for spatial autocorrelation.

This empirical result is consistent with the prediction of some growth theorists and findings from many prior empirical studies, which usually confirm that the diversity of an economy will help economic growth (Glaeser et al., 1992), but it might disappoint those theorists who believe that specialization of a particular manufacturing industry allows for full exploitation of scale economies and therefore helps "growth" (Henderson et al. 1995).

4.1.3 Labor-Market Conditions and Labor Mobility

The impact of labor-market conditions on employment growth is a little surprising because the initial wage level at 1990 is not relevant, meanwhile the education level of labor force has a statistically significant negative impact on the competitive employment growth within non-metro counties, Northern Appalachian counties, and transitional counties (coef. =-0.42, -0.81 and -0.66 respectively). This finding means that a county with a higher education level of labor force in 1990 has a *lower* growth rate from 1990 to 2000 compared to other counties in the same group. It might be that low-wage Appalachia counties were more attractive to firms seeking low skilled labor in the 1990s.

I use the percentage of people who were living in the same county five years ago as a proxy for labor mobility. The coefficients of this variable (**PSAMECNT90**) are negative for all county groups; meaning that more mobile counties (with *fewer* of the same people living in the same county between 1985 and 1990), experienced a higher competitive employment growth in the next decade (after controlling for the other factors in the model). Thus, mobility has a statistically significant positive impact on CEG for the overall model and all the sub-market models in Table 4.1. PSAMECNT90 is the only factor with consistent and highly statistically significant effects across all the models runs. It is not surprising because economic growth is usually higher in a place with a well-functioning labor market[24], which can be indicated by high labor mobility. Higher labor mobility ensures that labor is more quickly reallocated to where it can be employed most productively.

---

[24] A speech by William Poole, President, Federal Reserve Bank of St. Louis in Southern Illinois University at Edwardsville on April 10, 2003. http://stlouisfed.org/news/speeches/2003/4_10_03.html

4.1.4 Others: Amenities and Market Characteristics

Because I included natural amenities, people mobility, and metro status in 1990 as control variables in this analysis, I discuss them briefly.

The natural-amenity scale[25] has little impact on employment growth according to the regression results. Metro status in 1990 also only has a modest impact on the employment growth from 1990 to 2000, but it is confined to the southern Appalachian counties only.

[25] Refer to the added discussion in Chapter 2, Section 2.1.2.

Table 4.1: Regression Result Summary of Different Groups

		All		Metro and Non-metro				Three Sub-regions						Economic Status in 2000					
				Metro		Non-metro		Northern		Central		Southern		Distressed		Transitional		Others	
Sum-mary	Observations	312		94		218		109		55		148		55		230		27	
	F test(Prob.>F)	0.00		0.00		0.00		0.00		0.25		0.00		0.07		0.00		0.01	
	R^2	0.43		0.73		0.24		0.37		0.33		0.57		0.41		0.38		0.82	
	Adj R^2	0.40		0.69		0.19		0.27		0.08		0.52		0.18		0.34		0.61	
	Variables	**Coef.**	**P>t**	**Coef.**	**P>t**	**Coef.**	**P>t**	**Coef.**	**P>t**	**Coef.**	**P>t**	**Coef.**	**P>t**	**Coef.**	**P>t**	**Coef.**	**P>t**	**Coef.**	**P>t**
Industry Mix	MIN90	-0.17	***	D		-0.12	***	-0.05		0.02		-0.23		-0.01		-0.16	***	D	
	MANFC90	-0.40	***	-1.25	***	-0.05		0.15		0.23		-1.19	***	0.07		-0.72	***	-0.72	
	TRNSP90	-0.49	*	-1.53	***	-0.17		0.56	*	0.95		-1.98	***	-0.89		-0.90	***	-3.37	
	WHTRD90	1.85	***	0.50		1.13	*	2.05	***	1.06		1.39	**	-0.77		0.58		0.77	
	SERV90	-1.10	***	-2.08	***	-0.52		-0.65	***	-0.17		-1.66	***	0.23		-1.32	***	-2.22	**
	GOV90	-0.83	***	-1.87	***	-0.28		-0.39	*	-0.75		-1.39	***	-0.52		-0.94	***	-2.40	
Diversity &Specialization	BEAGINI90	-0.31	***	-0.35	*	-0.51	***	-0.14		-0.06		-0.07		-0.53		-0.30	**	-0.15	
	SGINI90	0.06		0.00		0.04		-0.02		0.05		0.11	*	0.09		-0.01		0.50	**
	SIM_M90	0.07	**	0.06		0.03		0.06		-0.07		0.11	**	-0.04		0.08	**	0.45	*
Labor Market Condi-tion	LNWAGE90	-0.09		0.11		-0.08		-0.08		-0.57	*	0.21	*	-0.30	*	0.05		0.82	
	UNEMPL90	0.00		0.01		0.00		-0.01	*	0.03	**	0.00		0.02	**	0.01		0.10	
	PHSGRAD90	-0.45	***	-0.59	*	-0.42	***	-0.81	***	-0.27		-0.70	**	-0.60	*	-0.66	***	-0.16	
	PSAMECNT9	-1.79	***	-2.58	***	-0.97	***	-0.89	***	-2.17	***	-1.83	***	-1.90	***	-1.32	***	-2.71	***
Others	SCALE	0.01	*	0.01		0.01		-0.02	*	0.00		0.01		0.02		0.01		0.05	
	METRO93	0.05	**	D		D		-0.03		0.06		0.07	**	0.08		0.02		0.23	*
Cons-tants	_CONS	2.89	***	2.53	*	1.94	***	2.02	**	7.16	***	0.60		4.80	***	1.52	*	-5.74	

*** <.05
** <.10
* <.20
Note: D-dropped

4.2 Accounting for Spatial Autocorrelation in the Model

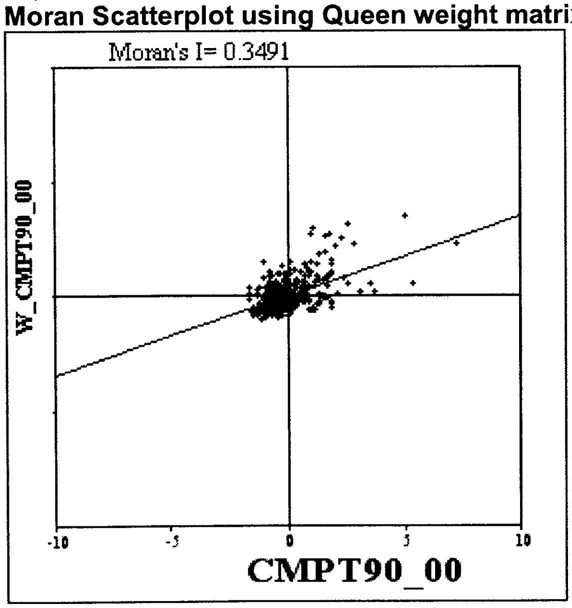
I have performed comprehensive OLS regressions to examine the relationship between competitive employment growth and its determinants in the previous section. But none of these analyses take into account the spatial factor. Spatial factors deal with the interactions among different counties and have been identified as being important in the economic growth literature (Nijkamp, 1998). In this section, I will focus on the spatial aspect of the models presented in Table 4.1 above. As discussed in the methodology chapter, I will firstly carry out the spatial exploratory data analysis (EDA) and visualize the spatial autocorrelation of competitive employment growth. I will then conduct the spatial regressions on CEG based on two different spatial weight matrixes and, finally, compare the results.

4.2.1 Spatial Exploratory Data Analysis for Spatial Autocorrelation

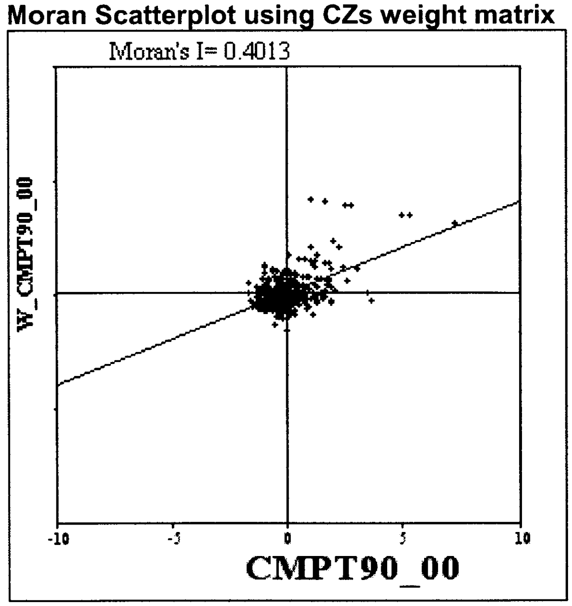
As discussed in Chapter 2, I test for the existence of spatial autocorrelation using Moran's I value, which can be visualized by Moran's I scatterplot, using the GeoDa software. Figure 4.1 provides two Moran's I scatterplots of competitive employment growth from 1990 to 2000. The left diagram used Queen-contiguity weight matrix and the right diagram used Commuting-Zones contiguity weight matrix.

Figure 4.1

Moran Scatterplot using Queen weight matrix



Moran Scatterplot using CZs weight matrix



In the two scatterplots, the signs of the slope of the regression lines are positive, indicating that the observed value of locations defined as "neighbors" tend to be similar in both cases. In other words, counties tend to have higher competitive employment growth rates where the growth rates are higher in the neighboring counties.

Both of Moran's I coefficients are far from zero, which indicates the existence of spatial autocorrelation, and, furthermore, it is strong. The Moran's I value of **CMPT90_00** using the Queen-contiguity weight is 0.3491, which is lower than the one (0.4013) using Commuting Zones contiguity weights. These values suggest that the spatial autocorrelation is stronger among counties and its socioeconomic "neighbors" within the same commuting zones, than the one among counties and its physical "neighbors" confined to having a common border or point. I return to this question later after reporting model results.

To visualize the local test of spatial autocorrelation, I use maps that depict the locations with significant Local Moran statistics (LISA[26] significance maps), and I classify those locations by type of association (LISA cluster maps) as follows.

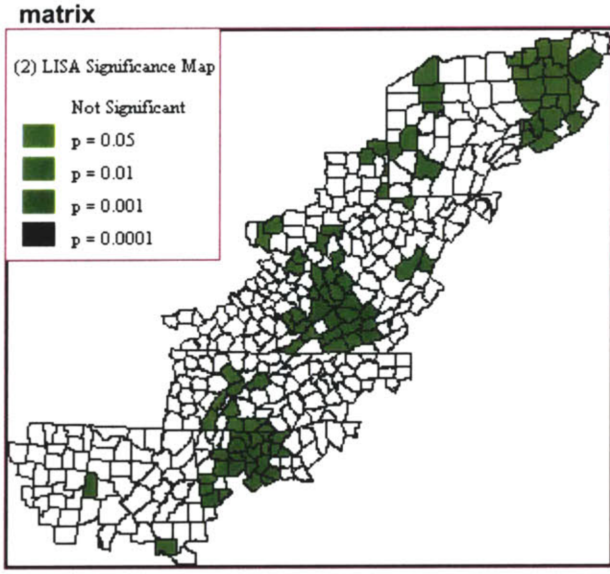
Figure 4.2 consists of both LISA significance maps and cluster maps for competitive employment growth from 1990 to 2000, in which I use a "Queen" contiguity weight matrix and a "Commuting Zones" contiguity weight matrix (based on 9999 random permutations).

---

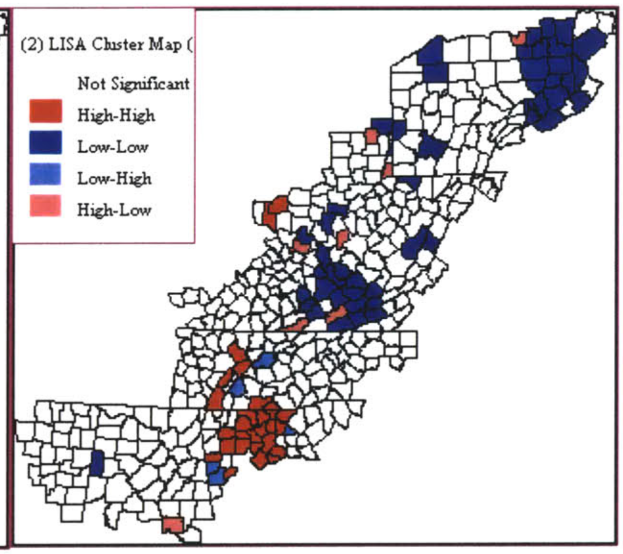
[26] LISA = local indicator of spatial autocorrelation.

Figure 4.2

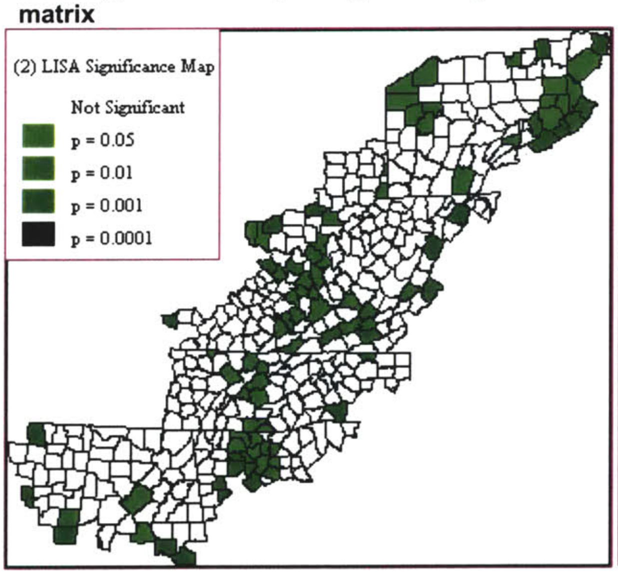
LISA significance map using Queen weight matrix



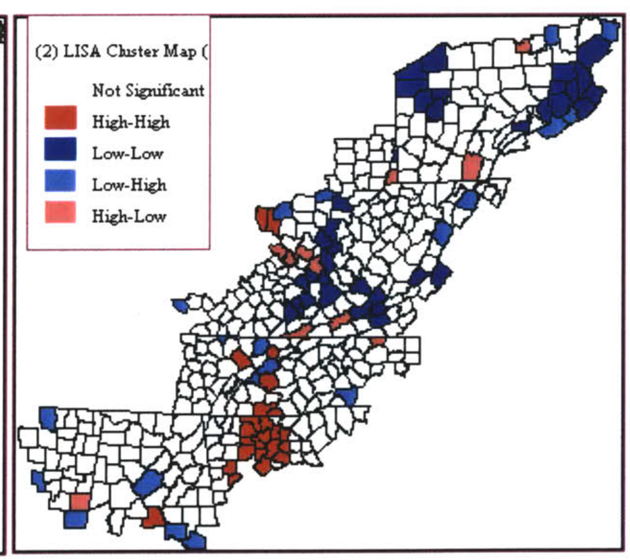
LISA cluster map using Queen weight matrix



LISA significance map using CZs weight matrix



LISA cluster map using CZs weight matrix



In the right-hand panel of Figure 4.2, I show two LISA cluster maps, depicting the locations of significant local Moran's I statistics, classified by type of spatial association (High-High, Low-

Low, Low-High, High-Low). Both of them use the default significance of P=0.05. These two right-hand panels show matching LISA significance maps for the two weights matrices.

I observe from both cluster maps that a big dark red cluster, which consists of counties with high competitive employment growth (CEG) rate from 1990 to 2000 surrounded by counties with high CEG, is in Georgia. They are counties between Atlanta and Chattanooga along the Interstate highway I-75.

I also observe two big dark blue (green) clusters that consist of counties with low CEG surrounded by low CEG in both cluster maps. They mostly consist of Central Appalachian Counties, along the convergent border of three states, Kentucky, Virginia, and West Virginia. I note another big dark blue (green) cluster in the northeast corner of Appalachian Region with most of the counties being in Northeast Pennsylvania, along the Interstate highway I-81. A smaller dark blue (green) cluster is also present in Pennsylvania State, specifically, in northwest Pennsylvania, along the interstate highway I-79.

Except for the dark blue (green) cluster (Low-Low) in Central Appalachia, all big dark clusters are constituted of counties along some interstate highway. This is consistent with many previous studies (e.g., Wilbur Smith Associates, 1998). Interstate highways do play an important role in the economic development process, by providing the vital connectivity of counties. However, an interesting finding arises when I focus on the only exception—Central Appalachia—where counties of low CEG cluster without any interstate connection. When I compare the spatial pattern of dark blue (green) polygons in Central Appalachia between two cluster maps, I find that the one constructed by commuting-zones contiguity, which indicates the spatial autocorrelation among "socioeconomic" neighbors through commuting ties, is much more scattered than the one constructed based on "physical" neighbors. I conduct additional investigations of these patterns using the following full hedonic regression analysis.

4.2.2 Spatial Regression Result Comparative Analysis

The objective here is to include spatial autocorrelation effects directly in the model of competitive employment growth and then to compare the regression result between OLS models and spatial models. The result of OLS regressions on different groups presented earlier in Section 4.1 suggests that there is some evidence of sub-market differences among the 410 counties and possible heteroskedasticity problems. The R^2 values for the model runs are much higher for the metro county submarket than for the non-metro submarket (R2=0.73 vs. 0.24).

I use several dummy variables and various interaction terms to allow for different constants and coefficients for metro and non-metro counties and for different geographic subregions. Specifically, I use the following formulation in Equation (3) to reconstruct my model:

$$y_j = (1+m)Rn \times \alpha_{Rn} + (1+m)Rs \times \alpha_{Rs} + m \times \alpha_m + m \times \sum_{i=1}^{k} \beta_i x_{ij} + n \times \alpha_n + n \times \sum_{i=1}^{k} \lambda_i x_{ij} + \varepsilon_j$$

(3)

(m -metro county dummy; n -non-metro county dummy; k is the number of variables)

Where

m =1 and n =0 for metro counties; m =0 and n =1 for non-metro counties;
Rn = 1 for northern counties, and Rs = 1 for southern counties;
y_j is CEG (competitive employment growth rate) for county j;
x_{ij} is the value of independent variable i for county j;
α_m, α_n, α_{Rn}, α_{Rs}, and are constant terms;
β_{ij} is the coefficient for metro county j's independent variable x_{ij};
λ_{ij} is the coefficient for non-metro county j's independent variable x_{ij};
ε_j is the error term for county j.

After running OLS regression in a complete form, I further tune the model according to the significance level of each variable. I show the results in Appendix F. Variables with a significance level less than 0.05 remain in the refined model. Therefore, the refined specification for the comparison between OLS regression and spatial model is expressed as Equation (4):

$$\begin{aligned} CEG_i = {} & \alpha_0 + \beta_1{*}M\_MANFC90_i + \beta_2{*}M\_TRNSP90_i + \beta_3{*}M\_SERV90_i + \beta_4{*}M\_GOV90i \\ & + \beta_5{*}PSAMECNT90_i \\ & + \beta_6{*}N\_MIN90_i + \beta_7{*}N\_MANFC90_i + \beta_8{*}N\_SERV90_i + \beta_9{*}N\_GOV90i \\ & + \beta_{10}{*}N\_DV1_i \\ & + \beta_{11}{*}N\_UNEMPL90_i + \beta_{11}{*}N\_PCTHSGRAD90 + \beta_{12}{*}N\_PSAMECNT90_i \\ & + \beta_{13}{*}N\_REGIONN + \varepsilon_i \qquad (4) \end{aligned}$$

In Table 4.2, I present the regression results for the competitive employment growth from 1990 to 2000 as the dependent variable, based on data of 410 Appalachian counties. I estimated the same model specification using OLS and Geoda's spatial lag and spatial error formulations. (Detailed results are in Appendix G). In Table 4.2, I show the results for the base model in Columns 4 and 5 and for the two spatial models in Columns 6-9. Columns 6 and 7 are the coefficients and significance levels of the spatial-lag model with Queen-contiguity weights matrix, while Columns 8 and 9 are estimations of spatial-lag model with Commuting-Zone contiguity weights matrix.

I choose the spatial-lag model, instead of the spatial-error model, because the spatial-lag model resulted in a better fit according to the value of the Lagrange Multiplier (LM) and the related robustness test (Anselin 1988b). As shown in the summary Table 4.2, the LM (lag) is higher than LM (error) for both weights matrixes, i.e., 38.28 vs. 24.09 for "Queen" contiguity weight matrix and 26.99 vs. 11.85 for "commuting zone" contiguity weights matrix. The robustness test for the lag-model is much more statistically significant than the robustness test for error model (0.00 vs. 0.84 for "Queen" and 0.00 vs. 0.15 for "Commuting Zone").

The OLS base model results are similar to those in Tables 4.1 for the separate metro and non-metro submarkets. The metro and non-metro interaction terms allow separate coefficients to be estimated for each submarket. Coefficient differences from the earlier results are due to eliminating variables that were not significant and including the regional dummy variable. The same industrial mix, diversity, and labor market condition factors are significant. The significant LM-LAG statistic in both spatial lag models (Queen and CZs) suggests that the influence of a county's economic development trend spreads well beyond the county border. The spatial-lag models fit better than the OLS model in terms of explanatory power indicating that it makes sense to attribute some of the CEG variability to spatial "spillover" effects and not just local county characteristics. In statistics, the improved fit is indicated by the increased values of R^2 for two spatial-lag models comparing with the OLS base model. The R^2 increases from 0.44 (OLS) to 0.50 (Queen-lag) and 0.49 (CZs-lag).

When we compare the additional variable (**W_CMPT90_00**) in the two spatial-lag models, although both of them have significantly positive effects on the competitive employment growth, the one weighted by commuting-zones-contiguity is 50% higher than that of the one using Queen-contiguity (0.45 vs. 0.31). This phenomenon implies that the impact of neighboring counties defined by "commuting tie" on a county's CEG is higher than that of counties confined to "physically bordering".

I also notice that almost all coefficients of variables are less in the "CZs" spatial-lag model than in the "Queen" one. It implies that the "commuting zone" method can do a little better than the "Queen" method by reducing the magnitude of the local effects and increasing the magnitude of "neighbor" effects. The characteristics of the county itself are less important when the spatial spillover effect is measured through the "commuting tie" rather than the effect from simply being a physical neighbor (Column 8 vs. Column 6). I also confirms the findings by comparing the Moran's I value between "CZs" and "Queen" method in Section 4.2.1, spatial autocorrelation effect is stronger among a county and its socioeconomic "neighbor(s)" within the same commuting zones, than the one among a county and its physical "neighbor(s)" confined to having a common border or point.

Table 4.2: Regression Result Summary of OLS model and Two Spatial-Lag models

	Model		OLS		Spatial Lag			
	Spatial Weight Matrix		N/A		Queen		CZs	
	Summaries	R^2	0.44		0.50		0.49	
			Coef.	P>t	Coef.	P>t	Coef.	P>t
	Constant terms	CONSTANT	2.54	0.00	2.22	0.00	2.22	0.00
		METRO93	0.27	0.49	0.14	0.69	-0.38	0.28
		W_CMPT90_00	N/A	N/A	0.31	0.00	0.45	0.00
Metro Counties	Industrial Mix	M_MANFC90	-1.53	0.00	-1.34	0.00	-1.00	0.00
		M_TRNSP90	-1.96	0.00	-1.71	0.00	-1.43	0.01
		M_SERV90	-2.90	0.00	-2.62	0.00	-2.21	0.00
		M_GOV90	-1.89	0.00	-1.55	0.00	-1.24	0.00
	Labor Market Condition	M_PSAMECNT	-1.87	0.00	-1.48	0.00	-1.13	0.00
Non-Metropolitan Counties	Industrial Mix	N_MIN90	-0.95	0.00	-0.82	0.00	-0.77	0.00
		N_MANFC90	-0.56	0.00	-0.52	0.00	-0.49	0.00
		N_SERV90	-1.44	0.00	-1.29	0.00	-1.26	0.00
		N_GOV90	-1.03	0.00	-0.93	0.00	-0.90	0.00
	Diversity	N_BEAGINI90	-0.33	0.06	-0.31	0.06	-0.34	0.04
	Labor Market Condition	N_UNEMPL90	0.01	0.12	0.01	0.12	0.01	0.11
		N_PHSGRAD9	-0.76	0.00	-0.66	0.00	-0.66	0.00
		N_PSAMECNT	-1.84	0.00	-1.55	0.00	-1.54	0.00
	Others	N_REGIONN	0.05	0.17	0.06	0.06	0.07	0.05
					Value	Prob.	Value	Prob.
	Test	**Lagrange Multiplier (lag)**	N/A		38.28	0.00	26.99	0.00
		Robust LM (lag)	N/A		14.21	0.00	17.20	0.00
		Lagrange Multiplier (error)	N/A		24.09	0.00	11.85	0.00
		Robust LM (error)	N/A		0.03	0.86	2.06	0.15

4.2.3 Outliers Analysis

Figure 4.1 shows that the outliers did not fit my empirical model by plotting the residuals for spatial-lag model using CZs contiguity weights matrix. Those counties with highlighted boundary line are the top/bottom seven counties with highest positive/negative residuals (Actual value – Predicted value= Residual value), which means my empirical model under/over predicts the CEG of these counties.

Specifically, there are three metro counties around Atlanta City and eleven are non-metro counties, as shown in Table 4.3 and Table 4.4. These counties are good candidates for further research (case studies) to understand what factors affecting their employment growth from 1990 to 2000 are not caught in my empirical model.

Table 4.3 Top Seven Counties with Highest Positive Residuals

Name	MSAs	Sub-region	LC_CZs	CEG	Residual
Dawson		Southern	Atlanta city, GA	1.68	1.15
Jackson		Central	Cookeville city, TN	0.78	0.91
Meigs		Southern	Cleveland city, TN	0.74	0.67
Clay		Northern	Charleston city, WV	0.32	0.56
Paulding	Atlanta MSA	Southern	Atlanta city, GA	1.20	0.52
Forsyth	Atlanta MSA	Southern	Atlanta city, GA	1.11	0.49
Potter		Northern	Olean city, NY	0.23	0.46

Note: MSAs–Metropolitan Statistical Areas; LC_CZs–largest city in its commuting zone; CEG–competitive employment growth from 1990 to 2000

Table 4.4 Bottom Seven Counties with Highest Negative Residuals

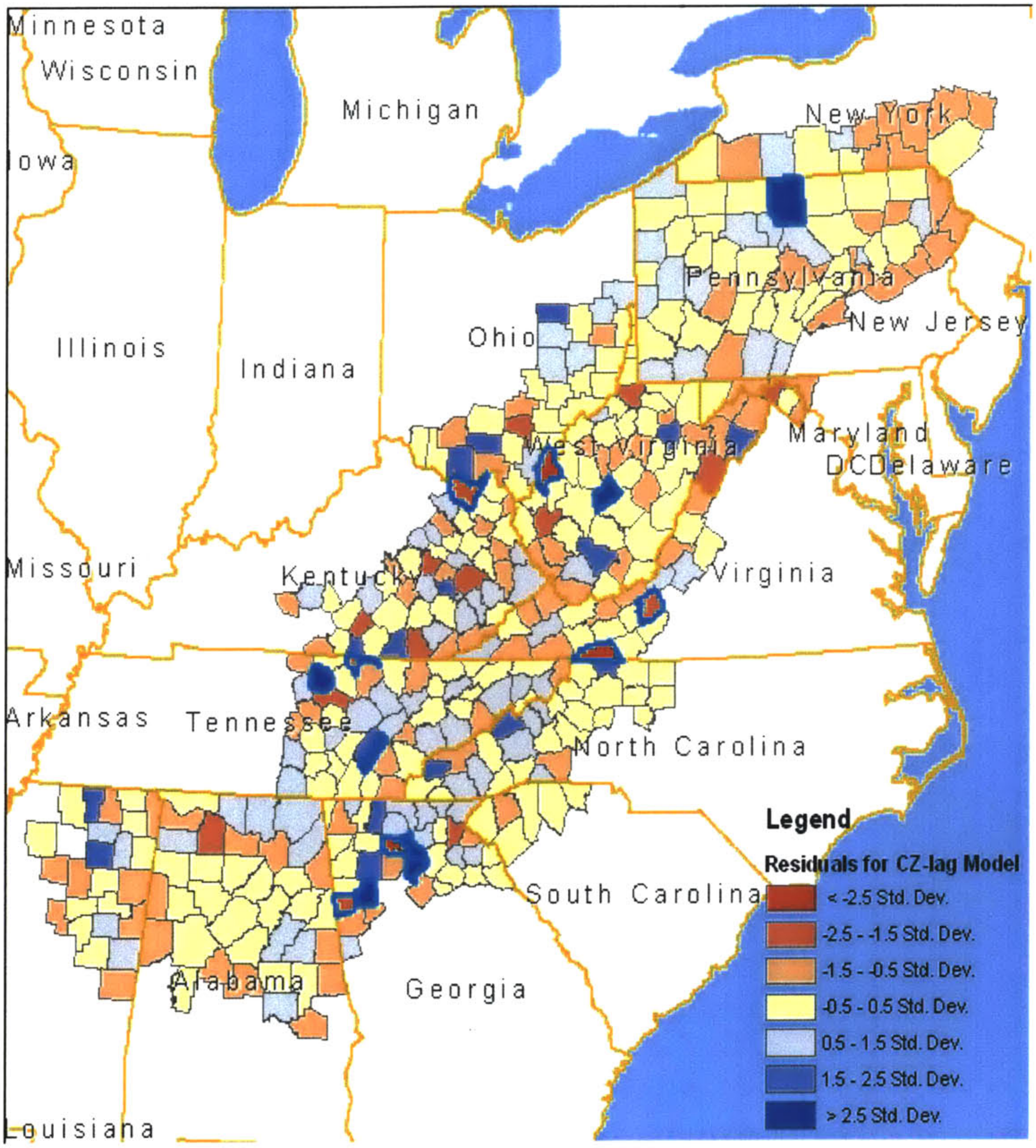
Name	MSAs	Subregion	LC_CZs	CEG	Residual
Haralson		Southern	Anniston city, AL	-0.40	-0.34
Montgomery[27]		Southern	Roanoke city, VA	-0.27	-0.35
Lewis		Central	Maysville city, KY	-0.57	-0.41
Pickett		Central	Cookeville city, TN	-0.40	-0.44
Mason		Northern	Athens city, OH	-0.56	-0.49
Grayson		Southern	Galax city, VA	-0.58	-0.49
Pickens	Atlanta MSA	Southern	Atlanta city, GA	0.11	-0.49

Note: MSAs–Metropolitan Statistical Areas; LC_CZs–largest city in its commuting zone; CEG–competitive employment growth from 1990 to 2000

In summary, the initial industrial mix (of some sectors), industrial diversity, labor market conditions and people mobility do have statistically significant effects on economic growth in my model. The spatial spillover effect of competitive employment growth is strong among the 410 counties in Appalachia. It is better caught by the commuting-tie weights lag model than the one weighted by geographical adjacency.

[27] Montgomery county + Radford city

Figure 4.1: Spatial Distribution of Residuals for CZ Spatial-Lag Model



Chapter 5

Conclusion and Further Research

According to my statistical analysis of the competitive growth rate from 1990 to 2000 and its hypothesized determinants, the empirical evidence from Appalachia suggests that the initial industrial mix (of some sectors), industrial diversity, labor-market conditions and people mobility do have statistically significant effects on economic growth in my model, after adjusting for distinct characteristics of metropolitan and non-metropolitan counties. Specifically, my results show that the share of the wholesale trade sector, diversity of the economy, and the mobility of people have positive effects, while the initial share of mining, manufacturing, services, and governmental sector have negative effects on the competitive employment growth rate.

From these results, I cannot conclude that they are the true determinants for economic growth in Appalachia region just based on statistical analysis, especially for the effect from industrial mix, because it might just mirror the technology change and industrial restructuring during the 1990s. For example, the high-tech sector is growing much faster than traditional industries during late 1990s, so that counties with more firms in this sector will result in greater growth during this period. Actually, the employment growth rate calculated from two data points (1990 and 2000) without looking at the business cycle in 1990s make it even more difficult to draw a strong conclusion based on the statistical results. I am more confident with the effects from other factors, such as diversity of economy and labor mobility, but the confirmation of the causality between these factors and the employment growth needs further research from theories or case studies.

Although with my empirical model I cannot confirm the causality between my hypothesized determinants and county competitive employment growth in Appalachia by the data themselves, there are still at least two interesting findings from the spatial analysis of this empirical model and the comparison among the traditional OLS regression model and two spatial-lag models.

First, after the spatial exploratory data analysis on the spatial distribution of the competitive employment growth (CEG) in Appalachian region, I find that a strong spillover effect of CEG exists among 410 counties. As presented in Figure 4.2, counties with high employment growth rates are concentrated in the north side of Atlanta Metropolitan area around Interstate highway

I-75. Counties with low growth rates are concentrated in Central Appalachia, along the convergent border of three states, Kentucky, Virginia, and West Virginia. Another low growth rate concentration is in Northeast Pennsylvania.

Second, I find the spillover effect of employment growth is better caught by the commuting-tie weights lag model than the one weighted by geographical adjacency. This finding implies that the impact of neighboring counties defined by "commuting ties" on a county's economic growth is higher than that of counties defined by "physically bordering". In other words, counties are more likely to interact through commuting flows than just by being next-door neighbors.

My empirical model is not able to explain completely the employment growth for all the counties in Appalachian Region, even after adjusting for the spatial-spillover effects, but it does provide some insights into what factors matter for many places for their competitive employment growth from 1990 to 2000. Further research for this study can be (1) to understand the impacts of industrial mix by looking deeply into the technology change and business cycle effects during different time periods; (2) to understand the effect of "economic base" on counties employment growth by incorporating the trade-flow data between counties and with regions outside of Appalachia; (3) to find some other non-economic factors by using some qualitative method such as case studies. The residuals analysis of this model will help us find some good candidates to understand what other determinants of economic growth would be.

Appendices

Appendix A: Top/Bottom 20 Appalachian Counties in Employment Growth from 1990 to 2000

Table A-1: Top-20 County in total employment growth, 1990-2000

STATE	R	BEA_ FIPS	County Name	1990 Private Employment	2000 Private Employment	Total Change (%)	National growth (%)	Mix (%)	Comp etitive (%)
Pennsylvania	N	42105	**Potter**	5628	9329	0.66	0.22	0.21	0.23
Kentucky	C	21053	**Clinton**	2287	3869	0.69	0.22	0.16	0.31
Tennessee	C	47087	**Jackson**	2373	5132	1.16	0.22	0.16	0.78
Alabama	S	01117	**Shelby**	32207	54503	0.69	0.22	0.19	0.27
Georgia	S	13085	**Dawson**	1854	5864	2.16	0.22	0.26	1.68
Georgia	S	13223	**Paulding**	7123	19014	1.67	0.22	0.25	1.20
Georgia	S	13117	**Forsyth**	15484	39201	1.53	0.22	0.20	1.11
Georgia	S	13135	**Gwinnett**	164341	325810	0.98	0.22	0.21	0.55
Georgia	S	13057	**Cherokee**	22212	43463	0.96	0.22	0.24	0.49
Georgia	S	13111	**Fannin**	4123	7359	0.78	0.22	0.21	0.35
Georgia	S	13281	**Towns**	2126	3777	0.78	0.22	0.23	0.32
Georgia	S	13011	**Banks**	2799	4968	0.77	0.22	0.10	0.45
Georgia	S	13291	**Union**	3493	6110	0.75	0.22	0.21	0.32
Georgia	S	13187	**Lumpkin**	3918	6765	0.73	0.22	0.20	0.30
Georgia	S	13097	**Douglas**	21510	36818	0.71	0.22	0.23	0.25
Georgia	S	13015	**Bartow**	21761	37154	0.71	0.22	0.16	0.33
Georgia	S	13195	**Madison**	3159	5274	0.67	0.22	0.18	0.26
Mississippi	S	28009	**Benton**	1078	1779	0.65	0.22	0.20	0.23
Tennessee	S	47121	**Meigs**	2118	4526	1.14	0.22	0.18	0.74
Virginia	S	51023	**Botetourt**	5824	10567	0.81	0.22	0.21	0.38

Table A-2: Top-20 county in competitive employment growth, 1990-2000

STATE	R	BEA_ FIPS	County Name	1990 Private Employment	2000 Private Employment	Total Change (%)	National growth (%)	Mix (%)	Comp etitive (%)
Ohio	N	39001	**Adams**	5717	8771	0.53	0.22	-0.30	0.62
West Virginia	N	54001	**Barbour**	3443	4068	0.18	0.22	-0.37	0.32
West Virginia	N	54015	**Clay**	1224	2003	0.64	0.22	0.09	0.32
Kentucky	C	21053	**Clinton**	2287	3869	0.69	0.22	0.16	0.31
Tennessee	C	47087	**Jackson**	2373	5132	1.16	0.22	0.16	0.78
Tennessee	C	47001	**Anderson**	35717	45009	0.26	0.22	-0.46	0.50
Georgia	S	13085	**Dawson**	1854	5864	2.16	0.22	0.26	1.68
Georgia	S	13223	**Paulding**	7123	19014	1.67	0.22	0.25	1.20
Georgia	S	13117	**Forsyth**	15484	39201	1.53	0.22	0.20	1.11
Georgia	S	13135	**Gwinnett**	164341	325810	0.98	0.22	0.21	0.55
Georgia	S	13057	**Cherokee**	22212	43463	0.96	0.22	0.24	0.49
Georgia	S	13011	**Banks**	2799	4968	0.77	0.22	0.10	0.45
Georgia	S	13123	**Gilmer**	5995	9299	0.55	0.22	-0.10	0.43
Georgia	S	13111	**Fannin**	4123	7359	0.78	0.22	0.21	0.35
Georgia	S	13015	**Bartow**	21761	37154	0.71	0.22	0.16	0.33
Georgia	S	13281	**Towns**	2126	3777	0.78	0.22	0.23	0.32
Georgia	S	13291	**Union**	3493	6110	0.75	0.22	0.21	0.32
Tennessee	S	47121	**Meigs**	2118	4526	1.14	0.22	0.18	0.74
Tennessee	S	47145	**Roane**	19873	18736	-0.06	0.22	-0.59	0.31
Virginia	S	51023	**Botetourt**	5824	10567	0.81	0.22	0.21	0.38

Table A-3: Bottom-20 county in total employment growth, 1990-2000

STATE	R	BEA_ FIPS	County Name	1990 Private Employment	2000 Private Employment	Total Change (%)	National growth (%)	Mix (%)	Competitive (%)
Ohio	N	39163	**Vinton**	2472	2240	-0.09	0.22	0.09	-0.40
West Virginia	N	54023	**Grant**	5393	4625	-0.14	0.22	-0.26	-0.11
West Virginia	N	54053	**Mason**	7479	6682	-0.11	0.22	0.23	-0.56
West Virginia	N	54029	**Hancock**	18369	16478	-0.10	0.22	0.12	-0.45
Kentucky	C	21095	**Harlan**	9962	7764	-0.22	0.22	0.08	-0.53
Kentucky	C	21135	**Lewis**	3025	2383	-0.21	0.22	0.14	-0.57
Kentucky	C	21025	**Breathitt**	3979	3386	-0.15	0.22	0.11	-0.48
Kentucky	C	21207	**Russell**	6921	5982	-0.14	0.22	0.11	-0.47
Kentucky	C	21171	**Monroe**	4074	3684	-0.10	0.22	0.14	-0.46
Kentucky	C	21159	**Martin**	3251	3010	-0.07	0.22	-0.05	-0.25
Virginia	C	51027	**Buchanan**	11869	8587	-0.28	0.22	0.04	-0.54
Virginia	C	51051	**Dickenson**	4153	3538	-0.15	0.22	0.06	-0.43
West Virginia	C	54047	**McDowell**	5763	4187	-0.27	0.22	0.02	-0.51
West Virginia	C	54059	**Mingo**	9745	8274	-0.15	0.22	0.10	-0.48
West Virginia	C	54045	**Logan**	12700	11513	-0.09	0.22	0.14	-0.46
Alabama	S	01037	**Coosa**	2310	2050	-0.11	0.22	0.01	-0.35
Alabama	S	01075	**Lamar**	6663	5989	-0.10	0.22	0.12	-0.44
Mississippi	S	28095	**Monroe**	13644	12192	-0.11	0.22	0.12	-0.45
Virginia	S	51077	**Grayson**	3726	2915	-0.22	0.22	0.14	-0.58
Virginia	S	51021	**Bland**	1831	1632	-0.11	0.22	0.14	-0.48

Table A-4: Bottom-20 county in competitive employment growth, 1990-2000

STATE	R	BEA_ FIPS	County Name	1990 Private Employment	2000 Private Employment	Total Change (%)	National growth (%)	Mix (%)	Competitive (%)
New York	N	36109	**Tompkins**	54207	55148	0.02	0.22	0.29	-0.50
New York	N	36077	**Otsego**	22825	23403	0.03	0.22	0.27	-0.46
Ohio	N	39081	**Jefferson**	27932	27239	-0.02	0.22	0.21	-0.46
Pennsylvania	N	42127	**Wayne**	15447	15697	0.02	0.22	0.24	-0.45
West Virginia	N	54053	**Mason**	7479	6682	-0.11	0.22	0.23	-0.56
West Virginia	N	54103	**Wetzel**	5090	4963	-0.02	0.22	0.23	-0.48
Kentucky	C	21135	**Lewis**	3025	2383	-0.21	0.22	0.14	-0.57
Kentucky	C	21095	**Harlan**	9962	7764	-0.22	0.22	0.08	-0.53
Kentucky	C	21025	**Breathitt**	3979	3386	-0.15	0.22	0.11	-0.48
Kentucky	C	21207	**Russell**	6921	5982	-0.14	0.22	0.11	-0.47
Kentucky	C	21171	**Monroe**	4074	3684	-0.10	0.22	0.14	-0.46
Kentucky	C	21071	**Floyd**	12387	11695	-0.06	0.22	0.17	-0.45
Virginia	C	51027	**Buchanan**	11869	8587	-0.28	0.22	0.04	-0.54
West Virginia	C	54047	**McDowell**	5763	4187	-0.27	0.22	0.02	-0.51
West Virginia	C	54059	**Mingo**	9745	8274	-0.15	0.22	0.10	-0.48
West Virginia	C	54045	**Logan**	12700	11513	-0.09	0.22	0.14	-0.46
Alabama	S	01087	**Macon**	5674	5906	0.04	0.22	0.30	-0.48
Mississippi	S	28095	**Monroe**	13644	12192	-0.11	0.22	0.12	-0.45
Virginia	S	51077	**Grayson**	3726	2915	-0.22	0.22	0.14	-0.58
Virginia	S	51021	**Bland**	1831	1632	-0.11	0.22	0.14	-0.48

Source: REIS, 1990 and 2000; percentage calculated by the author.
Note: R: Sub-Region; N: Northern Appalachia; C: Central Appalachia; S: Southern Appalachia

Appendix B: Definition of All Variables

Table B-1: Source and Definition of Variables

Variable Name	Source	Definition
Dependent Variable		
cmpt90_00	Bureau of Economic Analysis(BEA) 1990 and 2000	Total employment growth net of national growth and industrial mix effects (competitive component of shift-share analysis)
Explanatory Variables		
manfc90	BEA 1990	Percentage of income from manufacturing industry (SIC code=2&3); County as work-site
whtrd90	BEA 1990	Percentage of income from wholesale industry (SIC code=5); County as work-site
transp90	BEA 1990	Percentage of income from transportation industry (SIC code=4); County as work-site
serv90	BEA 1990	Percentage of income from services industry (SIC code=7&8); County as work-site
gov90	BEA 1990	Percentage of income from public administration sector (SIC code=9); County as work-site
beagini90	County Business Pattern 1990	Diversity Indicator I: Gini coefficient for private sectors' employment at the level of 1-digit SIC code at the county in 1990
sgini90	County Business Pattern 1990	Diversity Indicator II: Gini coefficient for service sector's employment at the level of 2-digit SIC code at the county in 1990
sim_m90	County Business Pattern 1990	Specialization Indicator: Simpson's index for manufacturing sector at the level of 2-digit SIC code at the county in 1990
unempl90	Bureau of Census 1990	Unemployment rate in 1990 (county as residence-place)
lnwage90	Bureau of Census 1990	Natural Logarithm of average wage in 1990 (county as work-site)
phsgrad90	Bureau of Census 1990	Percentage of people with high school education
psamecnt90	Bureau of Census 1990	Percentage of people living in the same county since 1985
scale	USDA/Economic Research Service	Natural amenity scale level
metro93	Office of Management and Budget (OMB) 1993	Whether or not metropolitan county in 1993 designation, 1=metro; 0=non-metro

Table B-2: Summary Statistics Table of Variables

Variable	# of Obs	Mean	Std. Dev.	Min	Max
CMPT90_00	410	-0.159	0.254	-0.578	1.683
MANFC90	410	0.271	0.149	0.006	0.708
WHTRD90	410	0.032	0.023	0.000	0.201
SERV90	410	0.158	0.071	0.000	0.615
GOV90	410	0.192	0.079	0.056	0.573
BEAGINI90	410	0.499	0.074	0.000	0.680
SGINI90*	399	0.506	0.163	0.000	0.781
SIM_M90**	334	0.597	0.323	0.100	1.000
UNEMPL90	410	0.081	0.034	0.027	0.219
LNWAGE90	410	0.098	0.002	0.093	0.103
PSAMECNT90	410	0.838	0.066	0.539	0.947
PHSGRAD90	410	0.612	0.102	0.355	0.872
SCALE	410	0.132	1.163	-3.720	3.550
METRO93	410	0.266	0.442	0.000	1.000

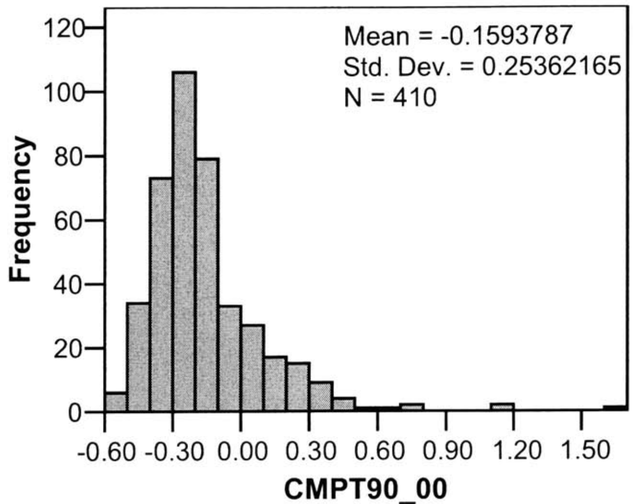
Appendix C: Distressed Designation and County Economic Status Classification System

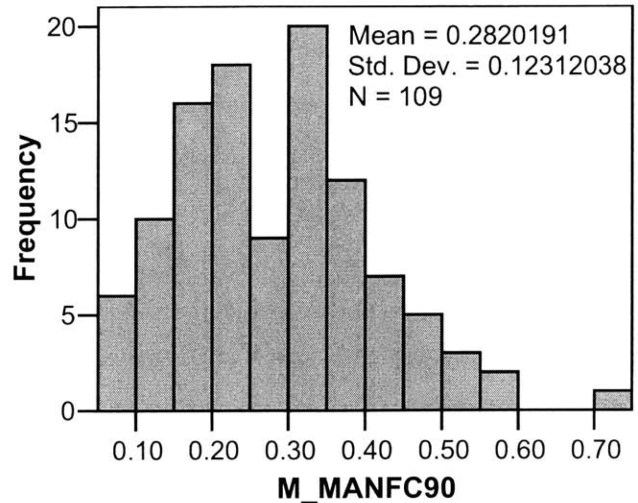
The Appalachian Regional Commission (ARC) developed a county economic classification system to target counties in need of special economic assistance. Four economic levels were created based on the comparison of three county economic indicators (three–year average unemployment, per–capita market income, and poverty) to their respective national averages. Thresholds have been established to define the four economic levels that classify counties: distressed, transitional, competitive, and attainment. Using the defined thresholds, ARC computes county economic levels each fiscal year using the most current data available. The county's economic level is then used in the distribution of funds for the fiscal year.

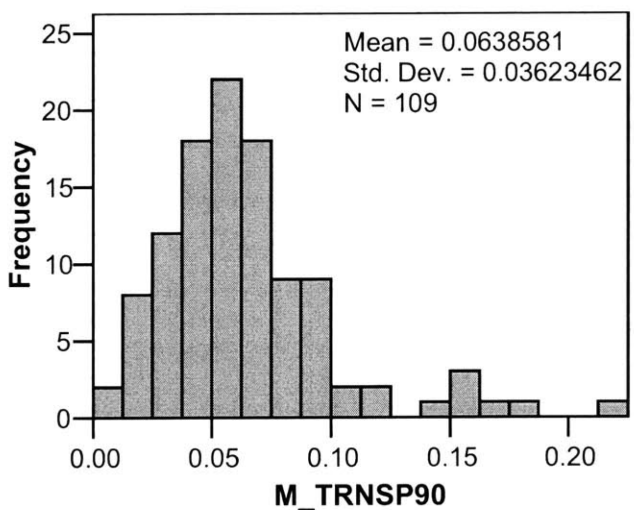
	County Economic Indicators		
County Economic Levels	**Three-Year Average Unemployment Rate**	**Per Capita Market Income**	**Poverty Rate**
Distressed	150% or more of U.S. average	67% or less of U.S. average	150% or more of U.S. average
Transitional	All counties not in other classes. Individual indicators vary.		
Competitive	100% or less of U.S. average	80% or more of U.S. average	100% or less of U.S. average
Attainment	100% or less of U.S. average	100% or more of U.S. average	100% or less of U.S. average

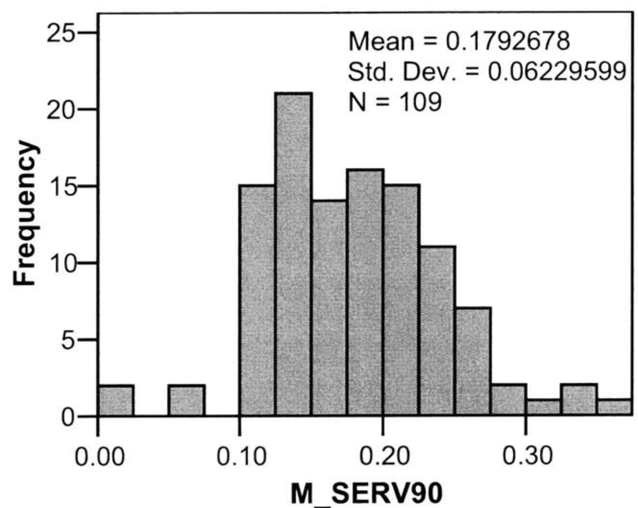
Source: Both text and table in this Appendix are from ARC website, http://www.arc.gov/search/method/cty_econ.jsp

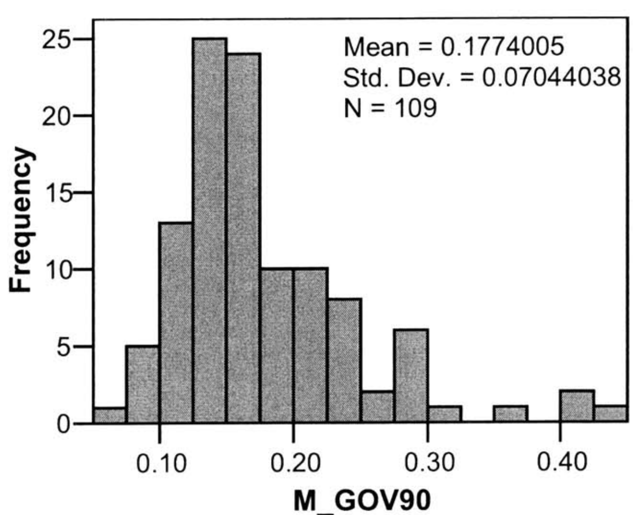
Appendix D: Distribution Histogram of Variables in Reduced Form Model

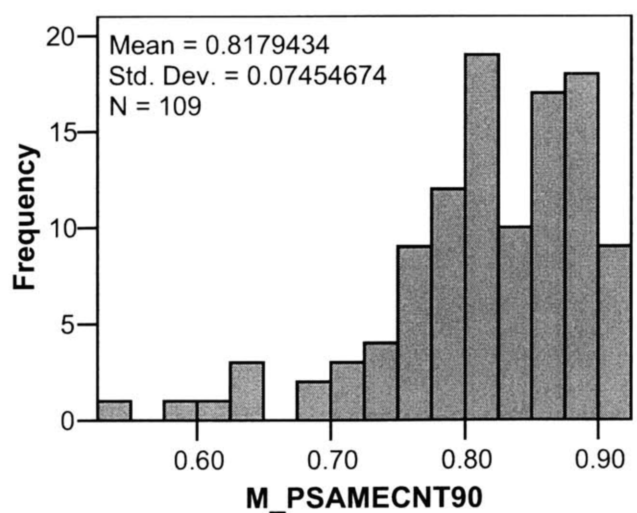


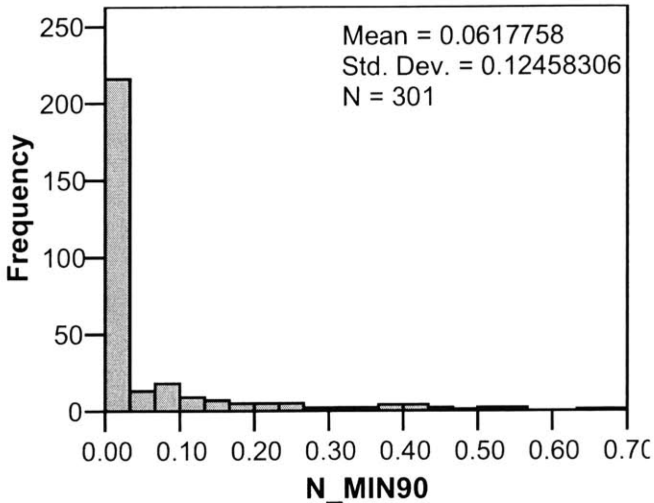


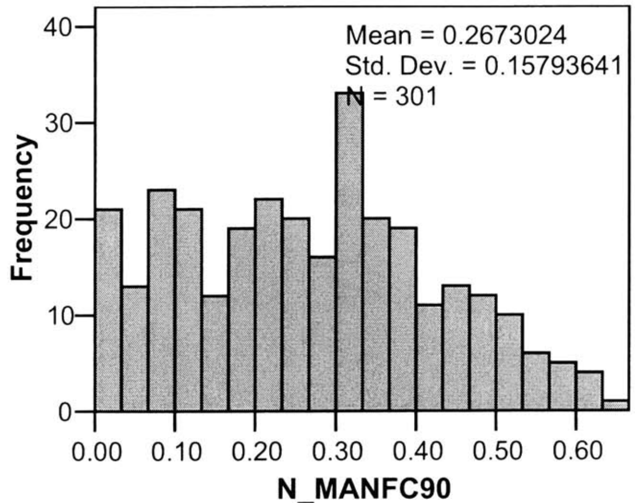


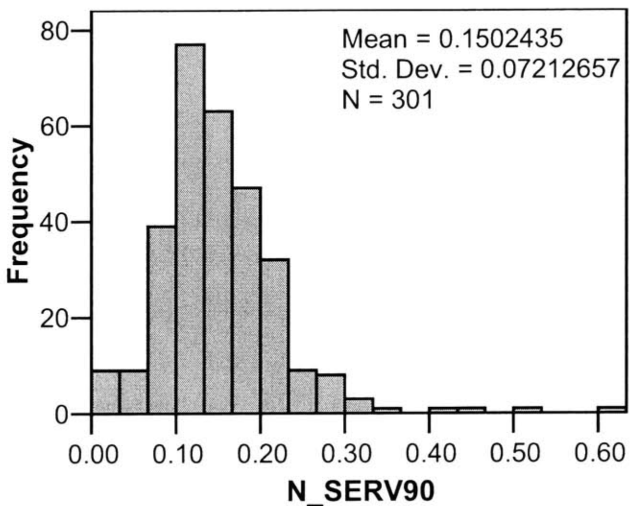


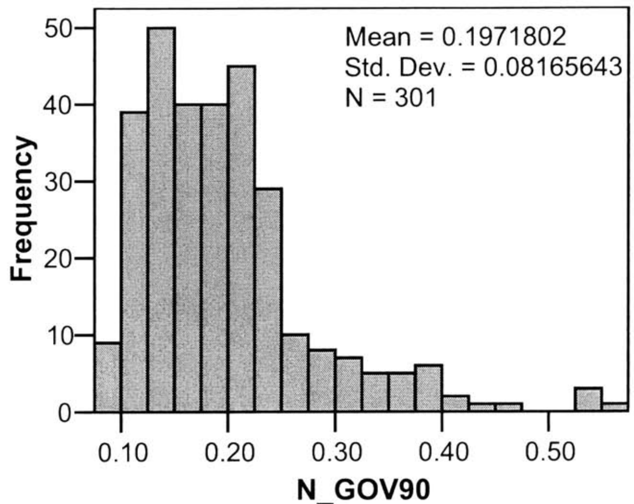


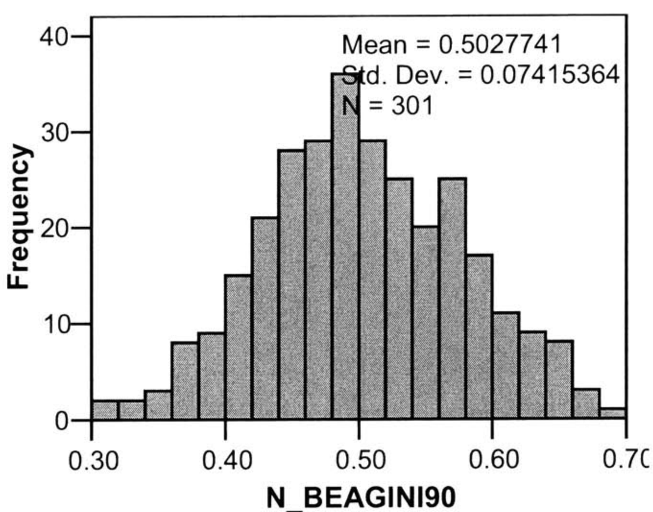


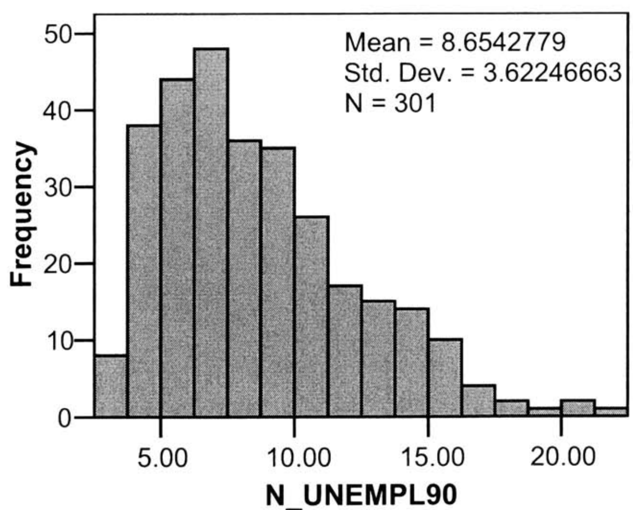

Mean = 0.0617758
Std. Dev. = 0.12458306
N = 301
Frequency
N_MIN90

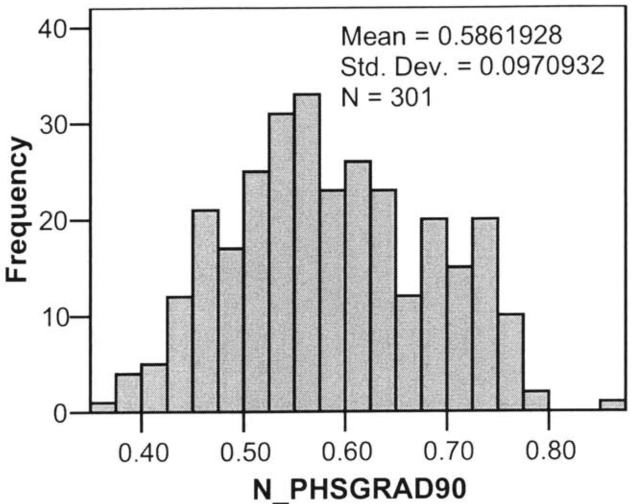

Mean = 0.2673024
Std. Dev. = 0.15793641
N = 301
Frequency
N_MANFC90

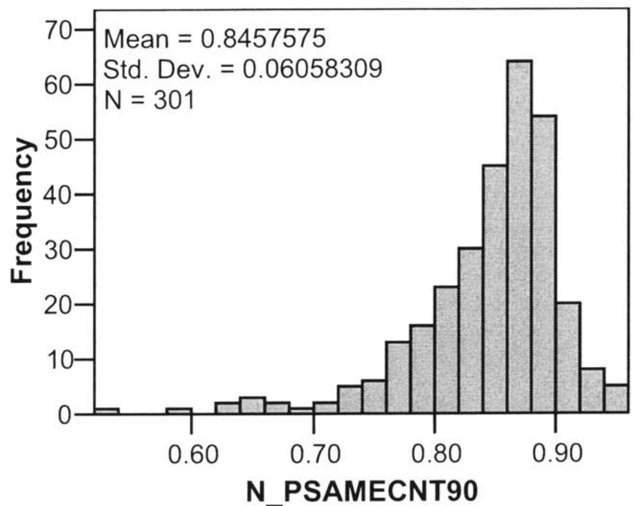

Mean = 0.1502435
Std. Dev. = 0.07212657
N = 301
Frequency
N_SERV90


Mean = 0.1971802
Std. Dev. = 0.08165643
N = 301
Frequency
N_GOV90


Mean = 0.5027741
Std. Dev. = 0.07415364
N = 301
Frequency
N_BEAGINI90


Mean = 8.6542779
Std. Dev. = 3.62246663
N = 301
Frequency
N_UNEMPL90


Mean = 0.5861928
Std. Dev. = 0.0970932
N = 301
Frequency
40
30
20
10
0
0.40
0.50
0.60
0.70
0.80
N_PHSGRAD90


Mean = 0.8457575
Std. Dev. = 0.06058309
N = 301
Frequency
70
60
50
40
30
20
10
0
0.60
0.70
0.80
0.90
N_PSAMECNT90

Appendix E: Ordinary Least Square (OLS) Regression Comparison

D-1: OLS Regressions for Metropolitan and Non-Metropolitan Counties

Table D-1-1: Metropolitan Counties

metro93 =	1			Number of observations	94
Source	SS	df	MS	F(12, 81)	17.91
Model	5.694247	12	0.4745	Prob > F	0
Residual	2.146195	81	0.0265		
Total	7.840442	93	0.0843		

R-squared	0.7263	Adj R-squared	0.6857		Root MSE	0.16278
CMPT90_00	**Coef.**	**Std.Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
MANFC90**	-1.18	0.328	-3.59	0.00	-1.829	-0.525
TRNSP90**	-1.61	0.741	-2.18	0.03	-3.084	-0.137
WHTRD90	0.41	0.929	0.44	0.66	-1.438	2.257
SERV90**	-2.34	0.560	-4.17	0.00	-3.451	-1.222
GOV90**	-1.86	0.404	-4.6	0.00	-2.662	-1.054
BEAGINI90	-0.40	0.256	-1.56	0.12	-0.908	0.110
SGINI90	-0.02	0.094	-0.24	0.81	-0.208	0.164
SIM_M90	0.10	0.083	1.2	0.23	-0.066	0.266
LNWAGE90	-0.01	0.177	-0.05	0.96	-0.361	0.343
UNEMPL90	0.01	0.013	1	0.32	-0.012	0.038
PSAMECNT90**	-2.53	0.369	-6.85	0.00	-3.265	-1.796
SCALE	0.03	0.019	1.37	0.18	-0.012	0.062
METRO93	(dropped)					
_cons	3.28	1.714	1.91	0.06	-0.135	6.687

Table D-1-2: Non-metro Counties

metro93 =	0			Number of observations	218
Source	SS	df	MS	F(12, 205)	4.54
Model	1.554445	12	0.1295	Prob > F	0
Residual	5.846221	205	0.0285		
Total	7.400667	217	0.0341		

R-squared	0.21	Adjusted R-squared	0.1638		Root MSE	0.16887
CMPT90_00	**Coef.**	**Std.Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
MANFC90	0.10	0.129	0.75	0.46	-0.158	0.350
TRNSP90	-0.18	0.451	-0.41	0.68	-1.072	0.705
WHTRD90*	1.36	0.775	1.76	0.08	-0.165	2.890
SERV90**	-0.47	0.220	-2.14	0.03	-0.905	-0.038
GOV90	-0.20	0.229	-0.89	0.38	-0.656	0.249
BEAGINI90**	-0.50	0.190	-2.65	0.01	-0.876	-0.128
SGINI90	0.03	0.100	0.32	0.75	-0.165	0.230
SIM_M90	0.04	0.045	0.92	0.36	-0.047	0.129

LNWAGE90**	-0.26	0.075	-3.51	0.00	-0.408	-0.115
UNEMPL90	0.00	0.006	0.64	0.52	-0.007	0.014
PSAMECNT90**	-0.85	0.265	-3.21	0.00	-1.371	-0.327
SCALE	0.02	0.011	1.47	0.14	-0.005	0.037
METRO93	(dropped)					
_cons	3.30	0.799	4.14	0.00	1.730	4.879

D-2: OLS Regressions for Three Sub-regions

Table D-2-1: Northern Appalachia

subregion	=	Northern		Number of observations	109
Source	SS	df	MS	F(13, 95)	3.21
Model	0.585716	13	0.0451	Prob > F	0.0005
Residual	1.333789	95	0.014		
Total	1.919506	108	0.0178		

R-squared	0.3051	Adj R-squared	0.2101		Root MSE	0.11849
CMPT90_00	Coef.	Std.Err.	t	P>t	[95% Conf.	Interval]
MANFC90	0.20	0.172	1.17	0.25	-0.141	0.542
TRNSP90	0.50	0.437	1.15	0.26	-0.367	1.370
WHTRD90**	2.33	0.896	2.6	0.01	0.547	4.105
SERV90**	-0.66	0.235	-2.82	0.01	-1.127	-0.195
GOV90	-0.34	0.296	-1.16	0.25	-0.932	0.244
BEAGINI90	-0.16	0.168	-0.95	0.35	-0.493	0.175
SGINI90	0.00	0.083	-0.05	0.96	-0.169	0.161
SIM_M90**	0.09	0.046	2.03	0.05	0.002	0.183
LNWAGE90**	-0.21	0.101	-2.09	0.04	-0.412	-0.010
UNEMPL90	-0.01	0.006	-0.86	0.39	-0.017	0.007
PSAMECNT90**	-0.66	0.329	-2	0.05	-1.314	-0.006
SCALE	-0.02	0.015	-1.43	0.16	-0.052	0.008
METRO93	-0.04	0.030	-1.19	0.24	-0.095	0.024
_cons	2.47	1.065	2.32	0.02	0.360	4.588

Table D-2-2: Central Appalachia

subregion	=	Central		Number of observations	55
Source	SS	df	MS	F(13, 41)	1.55
Model	0.957117	13	0.0736	Prob > F	0.1398
Residual	1.943231	41	0.0474		
Total	2.900348	54	0.0537		

R-squared	0.33	Adj R-squared	0.1176		Root MSE	0.21771
CMPT90_00	Coef.	Std.Err.	t	P>t	[95% Conf.	Interval]
MANFC90	0.19	0.426	0.44	0.67	-0.674	1.045
TRNSP90	0.91	1.189	0.77	0.45	-1.492	3.312
WHTRD90	0.91	2.148	0.42	0.67	-3.427	5.250
SERV90	-0.20	0.837	-0.23	0.82	-1.887	1.495
GOV90	-0.81	0.770	-1.05	0.30	-2.364	0.745
BEAGINI90	-0.05	0.617	-0.09	0.93	-1.299	1.194
SGINI90	0.06	0.250	0.26	0.80	-0.440	0.569
SIM_M90	-0.04	0.148	-0.3	0.77	-0.342	0.254
LNWAGE90**	-0.60	0.289	-2.06	0.05	-1.180	-0.012
UNEMPL90*	0.03	0.017	1.8	0.08	-0.004	0.064

PSAMECNT90**	-2.03	0.780	-2.61	0.01	-3.608	-0.458
SCALE	0.01	0.036	0.16	0.87	-0.066	0.077
METRO93	0.04	0.116	0.34	0.73	-0.194	0.273
_cons	7.13	3.158	2.26	0.03	0.754	13.508

Table D-2-3: Southern Appalachia

subregion	=	Southern		Number of observations	148
Source	SS	df	MS	F(13, 134)	12.99
Model	5.106233	13	0.3928	Prob > F	0
Residual	4.05327	134	0.0302		
Total	9.159503	147	0.0623		

R-squared	0.5575	Adj R-squared	0.5145		Root MSE	0.17392
CMPT90_00	**Coef.**	**Std.Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
MANFC90**	-1.06	0.234	-4.53	0.00	-1.526	-0.598
TRNSP90**	-1.93	0.666	-2.89	0.00	-3.245	-0.610
WHTRD90*	1.42	0.794	1.79	0.08	-0.147	2.992
SERV90**	-1.83	0.380	-4.83	0.00	-2.585	-1.083
GOV90**	-1.33	0.305	-4.36	0.00	-1.935	-0.728
BEAGINI90	-0.14	0.252	-0.57	0.57	-0.640	0.356
SGINI90	0.13	0.090	1.45	0.15	-0.047	0.308
SIM_M90**	0.13	0.060	2.08	0.04	0.006	0.245
LNWAGE90	0.05	0.132	0.35	0.73	-0.216	0.308
UNEMPL90	0.00	0.009	0.44	0.66	-0.014	0.023
PSAMECNT90**	-1.62	0.290	-5.58	0.00	-2.189	-1.043
SCALE	0.01	0.014	0.44	0.66	-0.021	0.033
METRO93*	0.07	0.037	1.93	0.06	-0.002	0.145
_cons	1.57	1.298	1.21	0.23	-0.998	4.136

D-3: OLS Regressions for Distressed, Transitional and Other counties

Table D-3-1: Distressed Counties

distr04 =	Distressed		
Source	SS	df	MS
Model	0.535324	13	0.04118
Residual	0.913349	41	0.02228
Total	1.448672	54	0.02683

Number of obs	55
F(13, 41)	1.85
Prob > F	0.0675

R-squared	0.3695	Adj R-squared	0.1696		Root MSE	0.14925
CMPT90_00	**Coef.**	**Std.Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
MANFC90*	0.02	0.266	0.06	0.96	-0.523	0.553
TRNSP90	-1.01	0.785	-1.29	0.21	-2.597	0.575
WHTRD90	-1.21	1.552	-0.78	0.44	-4.347	1.924
SERV90	0.30	0.581	0.51	0.61	-0.878	1.470
GOV90	-0.44	0.565	-0.78	0.44	-1.581	0.701
BEAGINI90	-0.59	0.433	-1.36	0.18	-1.461	0.287
SGINI90	0.11	0.172	0.61	0.54	-0.243	0.454
SIM_M90	0.03	0.123	0.24	0.81	-0.219	0.278
LNWAGE90*	-0.35	0.182	-1.91	0.06	-0.714	0.020
UNEMPL90*	0.02	0.011	1.78	0.08	-0.002	0.040
PSAMECNT90	-1.24	0.761	-1.64	0.11	-2.782	0.292
SCALE	0.02	0.024	0.95	0.35	-0.025	0.070
METRO93	0.06	0.116	0.52	0.61	-0.174	0.294
_cons	4.36	2.216	1.97	0.06	-0.119	8.832

Table D-3-2: Transitional Counties

distr04 =	Transitional		
Source	SS	df	MS
Model	3.356331	13	0.25818
Residual	6.382971	216	0.02955
Total	9.739303	229	0.04253

Number of obs	230
F(13, 216)	8.74
Prob > F	0

R-squared	0.3446	Adj R-squared	0.3052		Root MSE	0.1719
CMPT90_00	**Coef.**	**Std.Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
MANFC90**	-0.45	0.167	-2.68	0.01	-0.775	-0.118
TRNSP90*	-0.80	0.449	-1.78	0.08	-1.685	0.084
WHTRD90	1.16	0.740	1.57	0.12	-0.298	2.618
SERV90**	-1.32	0.279	-4.73	0.00	-1.872	-0.771
GOV90**	-0.76	0.252	-3.01	0.00	-1.255	-0.263
BEAGINI90	-0.26	0.196	-1.33	0.19	-0.648	0.126
SGINI90	-0.01	0.080	-0.18	0.86	-0.172	0.143
SIM_M90**	0.10	0.046	2.24	0.03	0.012	0.194
LNWAGE90**	-0.24	0.094	-2.52	0.01	-0.422	-0.052
UNEMPL90	0.01	0.007	1	0.32	-0.007	0.021

PSAMECNT90**	-1.24	0.245	-5.05	0.00	-1.720	-0.754
SCALE	0.01	0.011	1.2	0.23	-0.008	0.034
METRO93	0.03	0.028	1.05	0.29	-0.026	0.085
_cons	3.69	0.916	4.03	0.00	1.888	5.498

Table D-3-3: Better Performing Counties

distr04 =	Others				Number of obs	27
Source	SS	df	MS		F(13, 13)	4.52
Model	2.463077	13	0.18947		Prob > F	0.0053
Residual	0.545106	13	0.04193			
Total	3.008184	26	0.1157			

R-squared	0.8188	Adj R-squared	0.6376		Root MSE	0.20477
CMPT90_00	**Coef.**	**Std.Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
MANFC90	-0.74	1.117	-0.66	0.52	-3.151	1.676
TRNSP90	-3.52	3.023	-1.16	0.27	-10.050	3.011
WHTRD90	0.69	2.527	0.27	0.79	-4.768	6.149
SERV90*	-2.28	1.225	-1.86	0.09	-4.929	0.364
GOV90	-2.50	1.817	-1.38	0.19	-6.426	1.426
BEAGINI90	-0.12	0.652	-0.18	0.86	-1.528	1.289
SGINI90*	0.52	0.255	2.02	0.06	-0.035	1.066
SIM_M90*	0.47	0.236	2.01	0.07	-0.035	0.984
LNWAGE90	0.83	0.725	1.15	0.27	-0.736	2.397
UNEMPL90	0.11	0.070	1.56	0.14	-0.042	0.260
PSAMECNT90**	-2.67	1.066	-2.51	0.03	-4.972	-0.367
SCALE	0.05	0.049	1	0.34	-0.057	0.154
METRO93	0.24	0.163	1.48	0.16	-0.111	0.594
_cons	-5.96	6.854	-0.87	0.40	-20.767	8.846

Appendix F: OLS Regression in Complete Form

GeoDa output:

```
REGRESSION SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set         : newrun
Dependent Variable :  CMPT90_00 Number of Observations: 410
Mean dependent var :  -0.159379 Number of Variables  :  34
S.D. dependent var :  0.253312 Degrees of Freedom   : 376

R-squared         :  0.477808 F-statistic        :   10.4255
Adjusted R-squared :  0.431977 Prob(F-statistic)    :6.96024e-036
Sum squared residual:   13.7381 Log likelihood      :   114.412
Sigma-square      :  0.0365374 Akaike info criterion :  -160.824
S.E. of regression :  0.191148 Schwarz criterion    :  -24.2751
Sigma-square ML    :  0.0335075
S.E of regression ML:  0.183051
----------------------------------------------------------------------
  Variable  Coefficient   Std.Error  t-Statistic  Probability
----------------------------------------------------------------------
  CONSTANT   0.8524948   0.9398896   0.9070158  0.3649790
   METRO93   2.234332    2.009016    1.112153  0.2667836
   M_MIN90   -1.067555   0.9546939   -1.118217  0.2641868
  M_MANFC90   -1.353293   0.3424731   -3.951529  0.0000928
  M_TRNSP90   -1.722105   0.6815674   -2.526683  0.0119241
  M_WHTRD90   1.077476   0.8794343   1.225193  0.2212695
  M_SERV90   -2.588241   0.5122081   -5.053105  0.0000007
   M_GOV90   -1.684338    0.403335   -4.176028  0.0000369
 M_BEAGINI9   -0.3838108   0.2848899   -1.347225  0.1787201
  M_SGINI90  0.01433731   0.1092197   0.1312704  0.8956220
  M_SIM_M90  0.05345272   0.08257265   0.6473417  0.5178099
 M_LNWAGE90  -0.04814726   0.2024615   -0.2378094  0.8121601
 M_UNEMPL90  0.008640487   0.01575967   0.5482656  0.5838337
 M_PHSGRAD9  0.03764387   0.5273733   0.07137992  0.9431760
 M_PSAMECNT   -1.73917    0.3944176   -4.409464  0.0000135
   M_SCALE  0.01059255   0.02310846   0.4583842  0.6469447
  M_REGIONN  -0.08077091   0.09206242   -0.8773494  0.3808565
  M_REGIONS   0.0112211   0.08984956   0.1248877  0.9006728
   N_MIN90   -1.226374   0.1985062   -6.178014  0.0000000
  N_MANFC90  -0.7314531   0.1611426   -4.539166  0.0000076
  N_TRNSP90  -0.6193191   0.3369019   -1.838277  0.0668098
  N_WHTRD90  -0.2339414    0.68404   -0.3419996  0.7325504
  N_SERV90   -1.447641   0.2065534   -7.008557  0.0000000
   N_GOV90   -1.091843   0.2317113   -4.712082  0.0000035
 N_BEAGINI9   -0.356131   0.1765538   -2.017124  0.0443926
  N_SGINI90  0.04152135   0.07117227   0.5833922  0.5599741
  N_SIM_M90  0.001161411   0.0305043   0.03807369  0.9695745
 N_LNWAGE90   0.1922334   0.1000293    1.92177  0.0553898
 N_UNEMPL90  0.01107982   0.005379754    2.059541  0.0401304
 N_PHSGRAD9   -1.111692   0.2431831   -4.571421  0.0000066
 N_PSAMECNT   -1.751223   0.2633499   -6.649794  0.0000000
   N_SCALE  0.01616135   0.01057841   1.527767  0.1274119
  N_REGIONN   0.112766   0.04710633   2.393861  0.0171614
  N_REGIONS  0.04472594   0.03615515   1.237056  0.2168382
----------------------------------------------------------------------
```

Appendix G: Ordinary Least Square (OLS), Spatial Lag (SL) and Spatial Error (SE) Regressions in Reduced Form

G-1: OLS Regression vs. SL and SE Regressions using Queen Contiguity Weight

GeoDa output:

```
REGRESSION SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set       : newrunp
Dependent Variable :  CMPT90_00 Number of Observations: 410
Mean dependent var :  -0.159379 Number of Variables   :  16
S.D. dependent var :  0.253312 Degrees of Freedom   : 394

R-squared       :  0.442638 F-statistic       :   20.8601
Adjusted R-squared :  0.421419 Prob(F-statistic)    :2.52087e-041
Sum squared residual:   14.6633 Log likelihood    :   101.05
Sigma-square    :  0.0372166 Akaike info criterion :  -170.101
S.E. of regression :  0.192916 Schwarz criterion    :  -105.842
Sigma-square ML   :  0.0357643
S.E of regression ML:  0.189114
-----------------------------------------------------------------------
  Variable  Coefficient   Std.Error  t-Statistic  Probability
-----------------------------------------------------------------------
  CONSTANT   2.538337   0.2867759    8.85129  0.0000000
  M_MANFC90   -1.528766   0.2385629   -6.408233  0.0000000
  M_TRNSP90   -1.962514   0.6170154   -3.180656  0.0015857
  M_SERV90   -2.90363   0.3949812   -7.351312  0.0000000
   M_GOV90   -1.889287   0.3418206   -5.527128  0.0000001
 M_PSAMECNT   -1.867064   0.2848492   -6.554569  0.0000000
   N_MIN90  -0.9543176   0.152975   -6.238391  0.0000000
  N_MANFC90  -0.5622583   0.1394945   -4.030683  0.0000668
  N_SERV90   -1.439891   0.2046518   -7.035813  0.0000000
   N_GOV90   -1.027953   0.2124619   -4.838293  0.0000019
  N_BEAGINI9  -0.3277234   0.1753818   -1.868629  0.0624163
  N_UNEMPL90  0.007875714  0.005058027   1.557072  0.1202562
  N_PHSGRAD9  -0.7563575   0.1980215   -3.819574  0.0001553
  N_PSAMECNT   -1.838453   0.2595618   -7.082911  0.0000000
  N_REGIONN  0.05076975   0.03661835   1.386456  0.1663913
   METRO93  0.2666227   0.3829368   0.6962576  0.4866801
-----------------------------------------------------------------------


REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER  108.4966
TEST ON NORMALITY OF ERRORS
TEST          DF     VALUE       PROB
Jarque-Bera       2      1191.297    0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST          DF     VALUE       PROB
Breusch-Pagan test  15      184.3425    0.0000000
Koenker-Bassett test 15      37.37089    0.0011168
SPECIFICATION ROBUST TEST
TEST          DF     VALUE       PROB
White         135      N/A       N/A
```

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : **newrunp_410_queen.GAL** (row-standardized weights)

TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.154182	5.3556721	0.0000001
Lagrange Multiplier (lag)	1	38.2774088	0.0000000
Robust LM (lag)	1	14.2147440	0.0001631
Lagrange Multiplier (error)	1	24.0949116	0.0000009
Robust LM (error)	1	0.0322469	0.8574870
Lagrange Multiplier (SARMA)	2	38.3096556	0.0000000

======================== END OF REPORT ==============================

REGRESSION
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set : **newrunp**
Spatial Weight : **newrunp_410_queen.GAL**
Dependent Variable : **CMPT90_00** Number of Observations: 410
Mean dependent var : -0.159379 Number of Variables : 17
S.D. dependent var : 0.253312 Degrees of Freedom : 393
Lag coeff. (Rho) : 0.312194

R-squared : 0.495286 Log likelihood : 117.193
Sq. Correlation : - Akaike info criterion : -200.387
Sigma-square : 0.032386 Schwarz criterion : -132.112
S.E of regression : 0.179961

Variable	Coefficient	Std.Error	z-value	Probability
W_CMPT90_00	0.3121943	0.05449624	5.728731	0.0000000
CONSTANT	2.219607	0.2741982	8.094901	0.0000000
M_MANFC90	-1.337203	0.2241154	-5.966583	0.0000000
M_TRNSP90	-1.707882	0.5762689	-2.96369	0.0030399
M_SERV90	-2.61545	0.3706547	-7.056298	0.0000000
M_GOV90	-1.552745	0.3205595	-4.84386	0.0000013
M_PSAMECNT	-1.483906	0.2719014	-5.457515	0.0000000
N_MIN90	-0.8223876	0.1447188	-5.682661	0.0000000
N_MANFC90	-0.5249294	0.1304514	-4.023947	0.0000573
N_SERV90	-1.28791	0.1920266	-6.706936	0.0000000
N_GOV90	-0.9302916	0.1989137	-4.676861	0.0000029
N_BEAGINI9	-0.3123406	0.1637543	-1.907373	0.0564722
N_UNEMPL90	0.007409395	0.004718972	1.570129	0.1163851
N_PHSGRAD9	-0.6568795	0.186879	-3.514998	0.0004398
N_PSAMECNT	-1.550067	0.2462742	-6.294068	0.0000000
N_REGIONN	0.063347	0.03416331	1.854241	0.0637046
METRO93	0.1407561	0.3572225	0.3940291	0.6935596

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	15	173.2892	0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : **newrunp_410_queen.GAL**

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	32.28597	0.0000000

======================== END OF REPORT ==============================

REGRESSION

SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION

```
Data set          : newrunp
Spatial Weight      : newrunp_410_queen.GAL
Dependent Variable :  CMPT90_00 Number of Observations: 410
Mean dependent var :  -0.159379 Number of Variables  :  16
S.D. dependent var :  0.253312 Degree of Freedom   : 394
Lag coeff. (Lambda) :  0.363064

R-squared        :  0.489417 R-squared (BUSE)    : -
Sq. Correlation    : -       Log likelihood     : 113.232801
Sigma-square     :  0.032763 Akaike info criterion :  -194.466
S.E of regression  :  0.181005 Schwarz criterion    : -130.207087
```

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	2.396021	0.288686	8.299748	0.0000000
M_MANFC90	-1.385462	0.2263938	-6.119699	0.0000000
M_TRNSP90	-1.646778	0.5743741	-2.867083	0.0041429
M_SERV90	-2.64086	0.3629613	-7.275873	0.0000000
M_GOV90	-1.533934	0.3258308	-4.707764	0.0000025
M_PSAMECNT	-1.631927	0.2796057	-5.836528	0.0000000
N_MIN90	-0.9095739	0.1505324	-6.042379	0.0000000
N_MANFC90	-0.5577251	0.1323187	-4.215012	0.0000250
N_SERV90	-1.27996	0.1911317	-6.696741	0.0000000
N_GOV90	-0.9621373	0.1965179	-4.895927	0.0000010
N_BEAGINI9	-0.3502576	0.1652447	-2.119629	0.0340372
N_UNEMPL90	0.006198323	0.00505406	1.226405	0.2200466
N_PHSGRAD9	-0.8378153	0.2096158	-3.99691	0.0000642
N_PSAMECNT	-1.643429	0.2487978	-6.605484	0.0000000
N_REGIONN	0.0856656	0.04119442	2.079544	0.0375673
METRO93	0.04670385	0.3740145	0.1248718	0.9006250
LAMBDA	0.3630636	0.06426622	5.64937	0.0000000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	15	186.538	0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : **newrunp_410_queen.GAL**

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	24.36467	0.0000008

======================== END OF REPORT ============================

G-2: OLS Regression vs. SL and SE Regressions using Commuting Zones Contiguity Weight

GeoDa output:

```
REGRESSION SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            : newrunp
Dependent Variable :  CMPT90_00 Number of Observations: 410
Mean dependent var :  -0.159379 Number of Variables  :  16
S.D. dependent var :  0.253312 Degrees of Freedom   : 394

R-squared         :  0.442638 F-statistic       :   20.8601
Adjusted R-squared :  0.421419 Prob(F-statistic)   :2.52087e-041
Sum squared residual:   14.6633 Log likelihood     :   101.05
Sigma-square      :  0.0372166 Akaike info criterion :  -170.101
S.E. of regression :  0.192916 Schwarz criterion   :  -105.842
Sigma-square ML    :  0.0357643
S.E of regression ML:  0.189114
-----------------------------------------------------------------------
  Variable  Coefficient   Std.Error  t-Statistic  Probability
-----------------------------------------------------------------------
  CONSTANT    2.538337   0.2867759    8.85129  0.0000000
  M_MANFC90   -1.528766   0.2385629   -6.408233  0.0000000
  M_TRNSP90   -1.962514   0.6170154   -3.180656  0.0015857
  M_SERV90   -2.90363   0.3949812   -7.351312  0.0000000
   M_GOV90   -1.889287   0.3418206   -5.527128  0.0000001
 M_PSAMECNT   -1.867064   0.2848492   -6.554569  0.0000000
   N_MIN90  -0.9543176   0.152975   -6.238391  0.0000000
  N_MANFC90  -0.5622583   0.1394945   -4.030683  0.0000668
  N_SERV90   -1.439891   0.2046518   -7.035813  0.0000000
   N_GOV90   -1.027953   0.2124619   -4.838293  0.0000019
 N_BEAGINI9  -0.3277234   0.1753818   -1.868629  0.0624163
 N_UNEMPL90  0.007875714  0.005058027    1.557072  0.1202562
 N_PHSGRAD9  -0.7563575   0.1980215   -3.819574  0.0001553
 N_PSAMECNT   -1.838453   0.2595618   -7.082911  0.0000000
  N_REGIONN  0.05076975   0.03661835    1.386456  0.1663913
   METRO93   0.2666227   0.3829368   0.6962576  0.4866801
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER  108.4966
TEST ON NORMALITY OF ERRORS
TEST             DF      VALUE       PROB
Jarque-Bera       2       1191.297    0.0000000
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST             DF      VALUE       PROB
Breusch-Pagan test  15       184.3425    0.0000000
Koenker-Bassett test 15       37.37089    0.0011168
SPECIFICATION ROBUST TEST
TEST             DF      VALUE       PROB
White            135       N/A       N/A

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : newrunp_410_CZs.GAL (row-standardized weights)
TEST                 MI/DF    VALUE       PROB
Moran's I (error)       0.135734   3.7634593   0.0001676
Lagrange Multiplier (lag)     1    26.9892369   0.0000002
Robust LM (lag)           1    17.2032339   0.0000336
Lagrange Multiplier (error)   1    11.8482608   0.0005772
Robust LM (error)         1    2.0622579   0.1509863
Lagrange Multiplier (SARMA)   2    29.0514947   0.0000005
======================== END OF REPORT ==============================
```

REGRESSION

SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION

```
Data set         : newrunp
Spatial Weight     : newrunp_410_CZs.GAL
Dependent Variable :  CMPT90_00 Number of Observations: 410
Mean dependent var :  -0.159379 Number of Variables  :  17
S.D. dependent var :  0.253312 Degrees of Freedom   : 393
Lag coeff.  (Rho) :  0.449155

R-squared       :  0.488222 Log likelihood     :   119.584
Sq. Correlation   : -        Akaike info criterion :  -205.168
Sigma-square     :  0.0328393 Schwarz criterion    :  -136.893
S.E of regression  :  0.181216
```

Variable	Coefficient	Std.Error	z-value	Probability
W_CMPT90_00	0.4491548	0.03498205	12.83958	0.0000000
CONSTANT	2.216775	0.2735206	8.104599	0.0000000
M_MANFC90	-1.000854	0.226167	-4.425287	0.0000096
M_TRNSP90	-1.426486	0.5804745	-2.457448	0.0139928
M_SERV90	-2.213125	0.3734194	-5.926648	0.0000000
M_GOV90	-1.236601	0.3225014	-3.834404	0.0001259
M_PSAMECNT	-1.130424	0.2724926	-4.148455	0.0000335
N_MIN90	-0.7709955	0.1451371	-5.312187	0.0000001
N_MANFC90	-0.4875663	0.1315369	-3.706689	0.0002100
N_SERV90	-1.255102	0.1932484	-6.494757	0.0000000
N_GOV90	-0.899127	0.2002116	-4.490883	0.0000071
N_BEAGINI9	-0.336473	0.1648149	-2.04152	0.0411990
N_UNEMPL90	0.007678306	0.004752012	1.615801	0.1061373
N_PHSGRAD9	-0.6556901	0.1873016	-3.500717	0.0004641
N_PSAMECNT	-1.541098	0.2459486	-6.265936	0.0000000
N_REGIONN	0.06716161	0.03441961	1.95126	0.0510260
METRO93	-0.3849908	0.3599771	-1.069487	0.2848503

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	15	131.3859	0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : **newrunp_410_CZs.GAL**

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	37.06698	0.0000000

======================== END OF REPORT =============================

REGRESSION

SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION

```
Data set          : newrunp
Spatial Weight     : newrunp_410_CZs.GAL
Dependent Variable :  CMPT90_00 Number of Observations: 410
Mean dependent var :  -0.159379 Number of Variables  :  16
S.D. dependent var :  0.253312 Degree of Freedom    : 394
Lag coeff. (Lambda) :  0.707508

R-squared        :  0.479956 R-squared (BUSE)    : -
Sq. Correlation   : -       Log likelihood     : 116.667609
Sigma-square     :  0.033370 Akaike info criterion :  -201.335
S.E of regression  :  0.182674 Schwarz criterion    : -137.076704
```

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	2.290938	0.2954415	7.754287	0.0000000
M_MANFC90	-0.7727514	0.2343555	-3.297347	0.0009761
M_TRNSP90	-1.168803	0.5671539	-2.060822	0.0393199
M_SERV90	-1.861858	0.3605005	-5.164648	0.0000002
M_GOV90	-0.781218	0.3303541	-2.36479	0.0180403
M_PSAMECNT	-1.026475	0.2809606	-3.653447	0.0002588
N_MIN90	-0.8267698	0.150422	-5.496335	0.0000000
N_MANFC90	-0.547534	0.134175	-4.080746	0.0000449
N_SERV90	-1.16555	0.1827482	-6.377902	0.0000000
N_GOV90	-0.8611142	0.1907652	-4.513999	0.0000064
N_BEAGINI9	-0.3840766	0.1571496	-2.444018	0.0145247
N_UNEMPL90	0.002712699	0.005019287	0.5404551	0.5888832
N_PHSGRAD9	-0.8834116	0.2151411	-4.106196	0.0000402
N_PSAMECNT	-1.525791	0.2445147	-6.24008	0.0000000
N_REGIONN	0.1239154	0.04563657	2.715265	0.0066224
METRO93	-0.8635432	0.3854858	-2.240142	0.0250816
LAMBDA	0.7075077	0.02290749	30.88542	0.0000000

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	15	138.5611	0.0000000

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : **newrunp_410_CZs.GAL**

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	31.23429	0.0000000

======================== END OF REPORT ============================

References

Aldashev, Alisher; Joachim Möller. 2003. "Hysteresis in German Labor Market Participation?An East-West Comparison Using Regional Data." in *the Regensburg Econometric Re-search Seminar.*

Anselin, Luc. 1988a. *Spatial Econometrics: Methods and Models*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Anselin, Luc. 1988b. "Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity." *Geogrphical Analysis* 20:1-17.

Anselin, Luc. 1995. "Local Indicators of Spatial Association." *Geogrphical Analysis*:93-115.

Anselin, Luc. 1996. "The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association." Pp. 111-125 in *Spatial Analytical Perspectives on GIS in Environmental and Socioeconomic Sciences*, edited by M. Fischer, H. Scholten and D. Unwin. London: Taylor and Francis.

Anselin, Luc. 2001. "Spatial Econometrics." Pp. 310-330 in *Companion to Theoretical Econometrics*, edited by B. Baltagi. Oxford: Blackwell Scientific Publications.

Anselin, Luc. 2002. "Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models." *Agricultural Economics* 27:247-267.

Anselin, Luc. 2003b. "GeoDa™ 0.9 User's Guide."

Anselin, Luc, Bera, A. 1998. "Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics." Pp. 237-289 in *Handbook of Applied Economic Statistics*, edited by A. Ullah, Giles, D.E. New York: Marcel Dekker.

Anselin, Luc, Yong Wook Kim, Ibnu Syabri. 2003a. "Web-Based Analytical Tools for the Exploration of Spatial Data."

ARC, Appalachian Regional Commission. 1974. "The New Appalachian Subregions and Their Development Strategies." *Appalachia, A Journal of the Appalchian Regional Commission* 8:11-27.

Bao, Shuming. 2004. "Literature Review of Spatial Statistics and Models." China Data Center, University of Michigan.

BrandowCompany. 2001. "Analysis of Business Formation, Survival, and Attrition Rates of New and Existing Firms and Related Job Flows in Appalachia."

Brueckner, J.K. 2002. "Strategic Interaction among Governments: an Overview of Empirical Studies." *International Regional Science Review.*

Ciccone, A. and Hall R. 1996. "Productivity and the Density of Economic Activity." *American Economic Review*:57-70.

Dunn, Edgar S., Jr. 1960. "A Statistical and Analytical Technique for Regional Analysis." Pp. 97-112 in *Papers and Proceedings of the Regional Science Association.*

Garcia-Mila, Teresa; Therese J. McGuire. 1993. "Industrial Mix as a Factor in the Growth and Variability of States' Economies." *Regional Science and Urban Economics*:731-748.

Glaeser, Edward L., Hedi D., Kallal, Jose A. Scheinkman and Andrei Shleifer. 1992. "Growth in Cities." *Journal of Political Economy* 100:1126-1152.

Goodchild, M.F., Haining, R.P. and Wise, S. 1992. "Integrating GIS and Spatial Data Analysis: Problems and Possibilities." *International Journal of Geographical Information Systems* 6:407-423.

Heilbrum, James. 1987. *Urban Economics and Public Policy, 3rd edn*. New York: St. Martin's Press.
Henderson, Vernon; Kuncoro, Ari; Turner, Matt. 1995. "Industrial Development in Cities." *Journal of Political Economy* 103:1067-1090.

Hirschman, A. O. 1958. *The Strategy of Economic Development*. New Haven, CT: Yale University Press.
Isserman, Andrew. 1996. "Appalachia Then and Now: An Update of "The Realities of Deprivation" Reported to the President in 1964." in *Socioeconomic Review of Appalachia*.

Jacobs, Jane. 1969. *The Economy of Cities*: Random House.

Jensen, J. Bradford. 1998. "Birth and Death of Manufacturing Plants and Restructuring in Appalachia's Industrial Economy, 1963-1992: Evidence from the Longitudinal Research Database." H. John Heinz III School of Public Policy and Management,Carnegie Mellon Census Research Data Center.

Kasarda, John D.; Michael D. Irwin. 1991. "National Business Cycles and Community Competition for Jobs." *Social Forces* 69:773-761.

Killian, Molly Sizer, Tolbert Charles M. 1993. "Mapping Social and Economic Space: The Delineation of Local Labor Markets in the United States." in *Inequalities in Labor Market Areas*, edited by Joachim; Forrest A. Deseran Singelmann: Westview Press.

Kusimin, Lorin D. 1994. "Factors Associated with the Growth of Local and Regional Economies: A Review of Selected Empirical Literature." edited by Economic Research Service U.S. Department of Agriculture.

Kusimin, Lorin D.;John M. Redman; David W. Sears. 1996. "Factors Associated with Rural Economic Growth: Lessons from the 1980s." in *Technical Bulletin No. 1850*: U.S. Department of Agriculture, Economic Research Service.

Marshall, Alfred. 1890. *Principles of Economics*. London: Macmillan.

McGranahan, David. 1999. "Natural Amenities Drive Rural Population Change." in *Agricultural Economic Report No. (AER781)*: United State Department of Agriculture, Economic Research Service.

Nijkamp, Peter; Jacques Poot. 1998. "Spatial Perspectives on New Theories of Economic Growth." *The Annals of Regional Science* 32:7-37.

Paci, Raffaele; Usai Stefano. 1999. "The Role of Specialisation and Diversity Externalities in the Agglomeration of Innovative Activities." University of Cagliari and CRENoS and University of Sassari and CRENos, Italy.

Paelinck, J., L. Klaassen. 1979. *Spatial Econometrics*. Farnborough: Saxon House.

Pollard, Kelvin M. 2003. "Appalachia at the Millennium: an Overview of Results from Census 2000." Population Reference Bureau.

Pollard, Kelvin M. 2005. "Population Growth and Distribution in Appalachia: New Realities." in *Demographic and Socioeconomic Change in Appalachia*: Population Reference Bureau.

Quigley, John M. 1998. "Urban Diversity and Economic Growth." *Journal of Economic Perspectives* 12:127-38.

Romer, Paul M. 1986. "Increasing Returns and Long-Run Growth." *Journal of Political Economy*:1002-1037.

Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations (1937 Edition)*. New York: Random House.

Wilbur Smith Associates. 1998. "Appalachian Development Highways Economic Impact Studies."