INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®
SPATIAL ASSOCIATION MEASURES FOR AN ESDA-GIS FRAMEWORK: DEVELOPMENTS, SIGNIFICANCE TESTS, AND APPLICATIONS TO SPATIO-TEMPORAL INCOME DYNAMICS OF U.S. LABOR MARKET AREAS, 1969-1999

DISTRIBUTION

Presented in Partial Fulfillment of the Requirements for The Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Sang-Il Lee, M.A.

The Ohio State University
2001

Dissertation Committee:
Professor Lawrence A. Brown, Adviser
Professor Mei-Po Kwan
Professor Michael Tiefelsdorf

Approved by
Adviser
Department of Geography
ABSTRACT

This study is concerned with developing new spatial association measures (SAMs), elaborating generalized significance testing methods, proposing associated graphical and mapping techniques for an ESDA-GIS (Exploratory Spatial Data Analysis-Geographic Information Systems) framework, and applying those techniques to spatio-temporal income dynamics across U.S. labor market areas, 1969-1999. It is argued that SAMs play a central role in obtaining a seamless integration between ESDA and GIS where the cross-fertilization between them is highly achieved in such a way that ESDA takes advantage of GIS's data manipulation and visualization capabilities and a GIS utilizes ESDA's statistical integrity and computational efficiency.

Two sets of new SAMs are developed: global $S$ and local $S$, as univariate SAMs, and global $L$ and local $L$, as bivariate SAMs. Global $S$, spatial smoothing scalar, captures the degree of spatial smoothing when a geographical variable is transformed to its spatially smoothed vector in which each observation is re-computed in conjunction with its neighbors as defined in a spatial weights matrix. If a spatial pattern is more spatially clustered, it is given a higher value of $S$. Local $S$, defined as an observation's relative contribution to the corresponding global $S$, allows a researcher to detect spatial clusters.
Global $L$ and local $L_t$ are devised to conform to two concepts of association involved in comparing two spatial patterns in a simultaneous fashion: pairwise point-to-point association and univariate spatial association. Whereas aspatial bivariate association measure, such as Pearson's correlation coefficient, is dedicated solely to the first type of association, global $L$ captures numerical co-variances conditioned by topological relationships among observations to parameterize bivariate spatial dependence and to calibrate the degree of spatial co-patterning. Local $L_t$, a localized spatial correlation, captures the degree to which each location conforms to or deviates from the corresponding global $L$, and allows for exploring spatial heterogeneity in a bivariate relation.

Two sets of generalized significance testing methods are elaborated: one based on normality assumption and the other on randomization assumption. It is demonstrated that a transformation of SAMs to ratio of quadratic forms allows for a derivation of first four moments for global and local univariate SAMs including $S$ and $S_t$ under normality assumption. The Extended Mantel Test and the generalized vector randomization test are elaborated to compute first two moments of SAMs under randomization assumption. It is evidenced that the devised randomization test procedures can be applied to all the SAMs, whether global or local, whether univariate or bivariate, or whether a zero-diagonal in a spatial weights matrix or not.
A new set of ESDA techniques utilizing SAMs are proposed and its usefulness in geographical inquiries is illustrated with a hypothetical data set. For univariate situations, local-S significance map and Geary significance map are devised in comparison with the preexisting Moran significance map. For bivariate situations, local-L and local-r maps, local-L and local-r scatterplots, and local-L and local-r significance maps are proposed. When various significance levels are applied to those significance maps, probability maps can be created where higher $p$-value areas are expected to surround lower $p$-value areas, resulting in a probability surface. Those ESDA techniques are expected to accomplish various ESDA purposes.

The ESDA techniques are applied to an empirical study on spatio-temporal income dynamics across the U.S. labor market areas from 1969 to 1999. A series of local-S significance maps identify spatial clusters in regional per capital personal distribution, and show that spatial integration within the spatial clusters has been eroded. A notion of $\sigma$-convergence is not evidence. Rather, a trend toward income divergence is detected since the late 70s. Two distinctive trends towards regional income divergence observed in the late 1980s and the late 1990s seem to be associated with different spatial processes: the former with contagious spatial processes; the latter with sporadic spatial processes. A spatial autoregressive model concludes that there is no statistical evidence of $\beta$-convergence between 1969 income levels and income growth rate between 1969 and 1999. Various local-L scatterplot maps and significance maps report that spatial heterogeneity in $\beta$-convergence is evident during the entire period and varies sub-period to sub-period in terms of strength and locations.
Dedicated to my parents, Yong-Woo Lee and Chun-Ja Kwon,

and my sisters, In-Sun, Sung-Mi, and Hye-Sun
ACKNOWLEDGMENTS

I wish to thank my advisor, Dr. Lawrence A. Brown, for his support, encouragement, and endorsement. It has been an honor to work with him not only as a great scholar but also as a wonderful person. I have owed him everything during my stay at the Ohio State University.

I wish to express special thanks to Dr. Michael Tiefelsdorf for guiding me to the amazing world of spatial statistics and for consistently stimulating discussions. He has been a beacon for me to navigate the world without getting lost, although any errors found in this dissertation are utterly pertinent to me.

I am grateful to Dr. Mei-Po Kwan for allowing me to develop crucial ideas for the dissertation in her classes and for always encouraging me.

I also wish to thank faculty members and colleagues at the Ohio State University who have supported and encouraged me in pursuing the degree, among others, Dr. Kevin Cox, Dr. Nancy Ettlinger, Dr. Linda Lobao, Dr. Allan Murray, Dr. Bae-Gyoon Park, Dr. Joe Weber, Jon Moore, and Philip Hisnay.

I am grateful to faculty members in Department of Geography Education in Seoul National University, Dr. Jae-Ki Hwang, Dr. Ki-Suk Lee, Dr. Manik Hwang, Dr. Jong-
Wook Kim, and Dr. Jae-Myong Ryu, for having got me well prepared in pursuing for my final degree in this institution.

I am enormously indebted to a group of mentors in Korea who made my life here possible, among others, Dr. Sangcheol Kwon, Dr. Tae-Yeol Seo, Dr. Jung-Hwa Kwon, and Dr. Young-Min Lee. I won’t be able to repay them.

My special thanks go to my wife, Hyun-Mi Kim, who is the best colleague and friend as well as the loveliest woman on this planet.
VITA

November 27, 1966 .................................. Born – Pohang, Korea

1989 ........................................................... B.A., Geography Education, Seoul National University

1991 ........................................................... M.A., Geography Education, Seoul National University


1996 – present .......................................... Graduate Teaching and Research Associate, The Ohio State University

PUBLICATIONS

Research Publication


Major Field: Geography
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Abstract</strong></td>
<td>ii</td>
</tr>
<tr>
<td></td>
<td><strong>Dedication</strong></td>
<td>v</td>
</tr>
<tr>
<td></td>
<td><strong>Acknowledgments</strong></td>
<td>vi</td>
</tr>
<tr>
<td></td>
<td><strong>Vita</strong></td>
<td>viii</td>
</tr>
<tr>
<td></td>
<td><strong>List of Tables</strong></td>
<td>xiii</td>
</tr>
<tr>
<td></td>
<td><strong>List of Figures</strong></td>
<td>xv</td>
</tr>
<tr>
<td></td>
<td><strong>Chapters:</strong></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td><strong>Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.1 Problem specifications and research purposes</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.2 Structure of the dissertation</td>
<td>6</td>
</tr>
<tr>
<td>2.</td>
<td><strong>Spatial association measures and an ESDA-GIS framework</strong></td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>2.1 Nature of spatial data and spatial association measures</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>2.1.1 Spatial scale</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>2.1.2 Spatial structures and processes</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>2.1.3 Spatial dependence</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>2.1.4 Spatial heterogeneity</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>2.2 Significance testing methods for spatial association measures</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>2.2.1 Need for a generalized procedure and its role for ESDA</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>2.2.2 Different significance testing methods</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>2.3 An ESDA-GIS framework and spatial association measures</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>2.3.1 CSDA, ESDA, and GIS</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>2.3.2 Global and local spatial association measures</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>2.3.3 A SAM-based ESDA-GIS framework</td>
<td>40</td>
</tr>
</tbody>
</table>
7. ESDA techniques using local spatial association measures ........................................... 134

7.1 Univariate ESDA ........................................................................................................... 134
  7.1.1 ESDA using local Moran’s I ................................................................................. 134
  7.1.2 ESDA using local Geary’s c and Lee’s Si ............................................................ 137

7.2 Bivariate ESDA ........................................................................................................... 139
  7.2.1 Local-r and local-L maps ...................................................................................... 139
  7.2.2 Local-r and local-L scatterplots, and local-r and local-L scatterplot maps ........ 139
  7.2.3 Local-L significance and probability maps ............................................................ 142
  7.2.4 A discussion ........................................................................................................... 142


  8.1 Theoretical underpinnings of regional income convergence ................................ 146
  8.2 A critical review of empirical studies on regional income dynamics:
    recovery of spatiality ................................................................................................. 149
    8.2.1 σ-convergence: numerical variance vs. spatial clustering ............................... 149
    8.2.2 β-convergence: numerical correlation vs. spatial co-patterning .................... 154
  8.3 Spatio-temporal income dynamics across the US labor market areas ............... 160
    8.3.1 Research design ................................................................................................. 160
    8.3.2 Regional income distribution and identification of spatial clusters ............... 163
    8.3.3 Spatial co-patterning and bivariate spatial clusters in regional income growth ... 169
    8.3.4 σ-convergence and spatial dependence of income distribution .................... 175
    8.3.5 β-convergence and spatially autocorrelated errors ............................................. 180
    8.3.6 Spatial heterogeneity in β-convergence .............................................................. 187
    8.3.7 A geographically weighted regression (GWR): spatially drifting β-coefficients .. 192

9. Conclusions .................................................................................................................. 196

  9.1 Summary of findings ................................................................................................. 196
  9.2 Future research agenda ........................................................................................... 202

Bibliography ...................................................................................................................... 205
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Significance testing methods for spatial association measures</td>
<td>30</td>
</tr>
<tr>
<td>2.2</td>
<td>Studies utilizing local univariate spatial association measures</td>
<td>39</td>
</tr>
<tr>
<td>3.1</td>
<td>Univariate spatial association measures</td>
<td>49</td>
</tr>
<tr>
<td>3.2</td>
<td>Characteristics of local univariate spatial association measures</td>
<td>63</td>
</tr>
<tr>
<td>5.1</td>
<td>Definitions of global spatial proximity matrix T for global univariate spatial association measures</td>
<td>92</td>
</tr>
<tr>
<td>5.2</td>
<td>Definitions of local spatial proximity matrix T&lt;sub&gt;(l)&lt;/sub&gt; for local univariate spatial association measures</td>
<td>94</td>
</tr>
<tr>
<td>6.1</td>
<td>Quantities necessary for computing first two moments of spatial association measures</td>
<td>101</td>
</tr>
<tr>
<td>6.2</td>
<td>Five elements for computing mean and variance of spatial association measures with quantities</td>
<td>103</td>
</tr>
<tr>
<td>6.3</td>
<td>Significance testing methods for spatial association measures under the randomization assumption</td>
<td>106</td>
</tr>
<tr>
<td>6.4</td>
<td>Global spatial association measures in a matrix notation</td>
<td>108</td>
</tr>
<tr>
<td>6.5</td>
<td>Definitions of P and Q&lt;sub&gt;j&lt;/sub&gt;, and equations for expected values for spatial association measures</td>
<td>109</td>
</tr>
<tr>
<td>6.6</td>
<td>Local spatial association measures in a matrix notation</td>
<td>120</td>
</tr>
<tr>
<td>6.7</td>
<td>Definition of matrices and vectors</td>
<td>123</td>
</tr>
</tbody>
</table>
6.8 Expected values for local spatial association measures under the total randomization ................................................................. 126

6.9 Expected values for local spatial association measures under the conditional randomization ................................................................. 128

7.1 SAM-based ESDA techniques and tasks ................................................................. 143

8.1 Empirical studies on σ-convergence ................................................................. 150

8.2 Empirical studies on β-convergence ................................................................. 157

8.3 Top and bottom 10 LMAs in per capita personal income, 1969, 1979, 1989, and 1999 ................................................................. 165

8.4 Correlations between 1969 and 1999 per capita income distributions ............ 171

8.5 Correlations between 1969 income levels and income growth rates between 1969-1999 ................................................................. 182

8.6 OLS and SAR models for logarithmic initial per capita personal income levels and annual income growth rates ..................................... 186
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Structure of the dissertation</td>
</tr>
<tr>
<td>2.1</td>
<td>A SAM-based ESDA-GIS framework</td>
</tr>
<tr>
<td>3.1</td>
<td>Three spatial realizations of a hypothetical numeric vector</td>
</tr>
<tr>
<td>3.2</td>
<td>The relationship between $S$ and Moran’s $I$</td>
</tr>
<tr>
<td>3.3</td>
<td>A comparison of local univariate spatial association measures</td>
</tr>
<tr>
<td>4.1</td>
<td>Three bivariate spatial associations among three spatial patterns</td>
</tr>
<tr>
<td>4.2</td>
<td>Conceptualization of bivariate spatial association</td>
</tr>
<tr>
<td>4.3</td>
<td>$L$ with different bivariate spatial associations</td>
</tr>
<tr>
<td>4.4</td>
<td>Local $L_s$ in different local settings</td>
</tr>
<tr>
<td>7.1</td>
<td>Hypothetical spatial patterns</td>
</tr>
<tr>
<td>7.2</td>
<td>Univariate ESDA techniques</td>
</tr>
<tr>
<td>7.3</td>
<td>Bivariate ESDA techniques</td>
</tr>
<tr>
<td>8.1</td>
<td>Spatial distributions of per capital personal income across the US LMAs</td>
</tr>
<tr>
<td>8.2</td>
<td>Local-S significance maps: spatial clusters in per capita personal income across the US LMAs</td>
</tr>
<tr>
<td>8.3</td>
<td>Geary significance maps: local homogeneity in per capita personal income across the US LMAs</td>
</tr>
</tbody>
</table>

Page references: 8, 42, 47, 54, 60, 66, 68, 76, 82, 135, 138, 140, 164, 168, 170
8.4 Local-r scatterplot and scatterplot map of logarithmic per capita personal income across the US LMAs, 1969-1999

8.5 Local-L scatterplot map and significance map of logarithmic per capita personal income across the US LMAs, 1969-1999

8.6 Coefficients of variation of per capita personal income in the US, 1969-1999

8.7 Coefficients of variation and Moran's $I_s$ of per capita personal income across the US LMAs, 1969-1999

8.8 Moran significance maps: detection of spatial outliers in per capita personal income across the US LMAs, 1969-1999

8.9 OLS regression between 1969 logarithmic per capita personal income and annual income growth rate across the US LMAs, 1969-1999


8.12 Local-L significance maps: initial logarithmic per capita personal income levels and annual income growth rates, 1969-1999

8.13 A geographically weighted regression
CHAPTER 1

INTRODUCTION

1.1 Problem specifications and research purposes

Recent advances in spatial data analysis (SDA) or spatial statistics have attracted considerable attention from academic fields dealing with geographically referenced data (for overviews, see Fischer 1999; Getis 1999; Goodchild & Longley 1999; Fotheringham et al. 2000). SDA appreciates the particular nature of spatial data and attempts to spatialize general statistics by recognizing that uniform statistical assumptions seldom hold for spatial data. For example, observations in geographically referenced data are rarely independent from one another (spatial dependence), and spatial distributions often display significant local variations resulting in the presence of discrete spatial regimes within a study area (spatial heterogeneity).

Technological advances in computer and information sciences have dramatically changed the geocomputational environment for SDA (Openshaw & Clarke 1996; Openshaw and Alvanides 1999), and the development and maturation of GIS
spatial data have allowed researchers to manage and analyze massive spatial data with high-performing functionalities (Goodchild 1996; Goodchild and Longley 1999).

Need for an integration of GIS and SDA have increasingly been recognized, theoretical and technical issues regarding that matter have been discussed (among others, Goodchild 1987; Openshaw 1990; Fotheringham 1991; Anselin 1992; Fischer and Nijkamp 1992; Goodchild et al. 1992; Anselin and Getis 1993; Fotheringham and Rogerson 1993; Griffith 1993; Bailey 1994; Haining 1994), and a variety of analytical platforms (including commercial packages, e.g. SpaceStat for ArcView and S-Plus Extension for ArcView) have been produced. Further, a widespread emphasis on the visual in analytical sciences overall and related technological advances in visual sciences have precipitated SDA to adopt and develop tools for scientific visualization (Fotheringham 1999; Wise et al. 1999), cartographic visualization (Kraak and Ormeling 1996; Dykes 1996; 1997; 1998; Kraak 1999; Kraak and MacEachren 1999), or geovisualization (Kwan 2000). More importantly, the advent and development of EDA (Exploratory Data Analysis) (Tukey 1977; Good 1983; Cleveland 1993) as a new paradigm in statistics and its introduction to SDA (Jones 1984; Sibley 1987; Anselin 1988; Monmonier 1989; Bailey 1990; Haslett et al. 1990; Wartenberg 1990), and significant advances in local statistics in SDA (Getis and Ord 1996; Fotheringham 1997; Fotheringham and Brunsdon 1999) have collectively led to an enormous emphasis on ESDA (Exploratory Spatial Data Analysis) (Among others, Anselin & Getis 1992;
This dissertation underlines the importance of spatial association measures (SAMs) in an ESDA-GIS framework as a general research scheme for geographic information sciences. SAMs are defined as descriptive statistics for a spatial pattern or a set of spatial patterns, and is divided into two categories, global SAMs and local SAMs. The former summarizes the degree of spatial dependence in an overall spatial pattern or relations among spatial patterns, the latter gauges the extent to which a locale conforms to the overall global trend. For example, spatial autocorrelation as a univariate global SAM parameterizes the degree of univariate spatial dependence by capturing how (dis)similar a value in a location is to those in neighboring locations. Local spatial autocorrelation as univariate local SAMs, when mapped, reveals spatial heterogeneity indicating that spatial dependence varies locale to locale. In this respect, recent achievements in developing local univariate spatial association measures need to be appreciated. Getis and Ord (Getis 1991; Getis and Ord 1992; Ord and Getis 1995) developed a set of local measures ($G_l$ and $G_{l*}$). Anselin (1995; 1996) decomposed two global measures into respectively corresponding local measures, local Moran’s $I_l$ and local Geary’s $c_l$, and devised related graphical tools, such as local scatterplot and map. This class of univariate local SAMs (LISA: Local Indicators of Spatial Association) has played a crucial role in the integration of ESDA and GIS.

These endeavors, however, remain far from completed. First, with few exceptions, the general trend has focused on univariate spatial associations; techniques
attention (Fotheringham and Charlton 1994:322). Accordingly, a full-fledged ESDA-GIS framework should be built on a hybrid research environment where different research dimensions (univariate, bivariate, and multivariate) of analyses are cross-fertilized and fused into an integrative analytical framework.

Second, more efforts should be dedicated to devising local SAMs and to decomposing existing global techniques. Although it is obviously discouraging to face a historical fact (Unwin 1996:392) that almost a half-century has passed from global Moran's $I$ (1948) to its local version (Anselin 1995), rapid developments in geocomputation are expected to provide a better embryo sac for local statistics. This should be initiated by making a clear distinction between truly spatial measures and pseudo- or quasi-spatial ones, such as z-scores, OLS residuals, and factor scores: they may be still spatial in a certain sense, but not be exhaustively spatial because they are computed without any information on spatial relationships among observations. Further, ways of spatializing those aspatial measures should be devised.

Third, lack of generality in significance testing for SAMs has been a problem: most literature regarding significance testing has been confined to particular SAMs without providing generalized testing procedures. When SAMs are defined in a general form, a generalized significance testing procedure can be applied to the measures, whether univariate or bivariate or whether global or local.

Fourth, it should be recognized that, albeit the development of LISA, its applications have been limited at a large extent especially in geography. As Brown
(2000) correctly pointed out, more substantive research based on ESDA using local statistics should be conducted; thus conveying a 'demonstration effect'.

Thus, main research purposes are threefold.

First, this research devises two sets of SAMs: one set consisting of a univariate global SAM $S$ and local SAM $S_l$; the other set composed of a bivariate global SAM $L$ and local SAM $L_l$. The development of the former is motivated by the fact that current univariate SAMs, such as Moran's $I$ and Geary's $c$, are largely determined by reference areas so that a new measure needs to be introduced to capture the degree of spatial dependence pronounced over entire locales. The latter is expected to provide a new insight into bivariate spatial dependence and heterogeneity since similar local correlations are often spatially clustered and the degree of the spatial correspondence could vary locale to locale. For each measure, some ESDA techniques including graphical and mapping procedures are illustrated.

Second, the research provides two sets of generalized significance testing methods for SAMs: one set based on the normality assumption and the other on the randomization assumption that is composed of two different procedures, the Extended Mantel Test and a generalized vector randomization test. Whereas the former is confined to local SAMs, the latter is extended to bivariate SAMs. It is demonstrated that those generalized statistical tests are applicable to any spatial weights matrices.

Third, measures and associated ESDA techniques along with other spatial statistical techniques such as spatial autoregressive models and geographically weighted regression methods, are applied to an empirical study on spatio-temporal income
dynamics across the U.S. labor market areas (LMAs), 1969-1999. Theoretical underpinnings and empirical findings resulting from previous studies are critically reviewed and how SAM-based ESDA-GIS framework can make positive contributions to the topic.

1.2 Structure of the dissertation

Figure 1-1 summarizes the structure of this dissertation. The dissertation is largely divided into three sections: the first section with Chapter 1 and 2; the second section from Chapter 3 to Chapter 6; the third section with Chapter 7, 8, and 9.

In the first section, Chapter 2 provides a conceptual overview on roles of SAMs for an ESDA-GIS framework. The nature of spatial data is formulated in terms of four associated concepts, spatial scale, spatial structure and process, spatial dependence, and spatial heterogeneity. These concepts underlie the rationales for SAMs. Finally, a SAM-based ESDA-GIS framework is proposed.

In the second section, Chapter 3 and 4 address the development of new SAMs, $S$ and $S_h$, and $L$ and $L_i$. Chapter 5 and 6 are dedicated to the development of generalized significance testing methods based on the normality assumption and the randomization assumption. For each chapter, a general procedure is first provided and then is applied to various SAMs.

The third section illustrates ESDA techniques using local SAMs (Chapter 7) and their applications, with other spatial statistical techniques, to spatio-temporal income
dynamics across the U.S. 391 labor market areas, 1969-1999 (Chapter 8). The final chapter discusses related issues and future research themes (Chapter 9).
CHAPTER 1
Introduction

CHAPTER 2
SAMs and ESDA-GIS Framework

CHAPTER 3
Univariate SAMs: $S$ and $S_i$

CHAPTER 4
Bivariate SAMs: $L$ and $L_i$

CHAPTER 5
A Generalized Significance Test
I: Normality Assumption

CHAPTER 6
A Generalized Significance Test
II: Randomization Assumption

CHAPTER 7
ESDA Techniques Using SAMs

CHAPTER 8
An Application

CHAPTER 9
Conclusions

Figure 1.1: Structure of the dissertation
CHAPTER 2

SPATIAL ASSOCIATION MEASURES AND AN ESDA-GIS FRAMEWORK

This chapter reviews the current status of spatial association measures (SAMs) with respect to overall spatial data analysis (SDA) and formulates a SAM-based ESDA-GIS framework. I first clarify that the nature of spatial data necessitates the development of SAMs. Second, I demonstrate that generalized statistical tests are needed for SAMs and are necessary for ESDA. Third, I formulate a SAM-based ESDA-GIS framework.

2.1 Nature of spatial data and spatial association measures

Spatial statistics or spatial data analysis is not simply a bundle of methods or statistical techniques, but is a new perspective on geographically referenced data with a theoretical integrity centered on space. Applications of general statistical techniques to spatial data, from ANOVA to multivariate regression, may be flawed because they are based on aspatial statistical assumptions on underlying processes. For example, properties for hypothesis testing in regression analyses, \( t \)-test for parameter estimators
and $F$-test for overall goodness-of-fit of a regression equation, may not be dependable due to the very nature of spatial data. The relationship between general statistics and spatial statistics is similar to that between aspatial political economy and spatially informed critical social sciences (e.g. Lobao et al. (1999) show a good example of spatially informed Social Structures of Accumulation (SSA) approach). Further, a ‘global’ or ‘average’ estimation (e.g. a single set of regression parameters) does not apply equally to all parts of the whole study area, because statistical relationships vary sub-region to sub-region (Brown and Jones 1985; Brown 1991). For example, a positive relationship between two variables may be reversed in certain sub-regions. Estimating localized parameters has increasingly been a crucial part of spatial statistics.

Although spatial statistics has been introduced to social science under the banner of spatial data analysis or spatial econometrics (Anselin 1988), its scope is much broader to include geostatistics, environmetrics, biostatistics, and so on. I give a brief introduction to spatial statistics, focusing on quantitative geography and its evolution to spatial statistics. Rationales for spatial statistics are associated with the appreciation of ‘spatial effects’ differentiating spatially referenced data from general data, such as scale effect, spatial dependence, and spatial heterogeneity, and the recognition of those spatial effects in the discipline of statistics was responsible for much of early spatial statistics (Gehlke and Biehl 1934; Moran 1948; Geary 1954; Krishna Iyer 1949).

According to Anselin and Griffith (1988), the real exposure of regional science and geography to spatial statistics was achieved by the work of Cliff and Ord (1973; revised in 1981), which Getis (1999:241) regards as having opened “the door to a new era
in spatial statistics." What makes the book really important is the fact that it enhanced spatial data analysis or quantitative geography qualitatively from a simple application of general statistical techniques to spatial data (this include some 'for geographers'-type books; e.g. Ebdon 1977; Johnston 1978; Clark and Hosking 1985; Rogerson 2001) or a descriptive or semi-inferential level (for a review on this level of spatial data analysis, Unwin 1981; Gatrell 1983) to a full-fledged inferential level. The trend was fully adopted in a bible of quantitative geography (Haggett et al 1977), and gave birth to several seminal books for spatial statistics (Upton and Fingleton 1985; 1989; Anselin 1988; Griffith 1988; Haining 1990; Cressie 1993; Bailey and Gatrell 1995; Tiefelsdorf 2000; Fotheringham et al. 2000), with being inspired by statisticians (Ripley 1981; 1988; Silverman 1986; Cressie 1993) and geostatisticians (Isaaks and Srivastava 1989). I will try to demonstrate the importance of spatial statistics by examining each of its conceptual elements; (i) spatial scale; (ii) spatial structures and processes; (iii) spatial dependence; (iv) spatial heterogeneity.

2.1.1 Spatial scale

Spatial scale implies three different things: geographical scale means the spatial extent of the whole study area; observational scale denotes the physical size of the spatial unit of a study; operational scale refers to the spatial extent to which a phenomenon underlain by a process operates. The relationship between spatial unit and operational scale is crucial because a mismatch between them may generate flawed results. For example, many of state-based regional analyses on economic performance variables in
interaction, spatial externalities, neighborhood effects, and other spillover effects, underlying the spatial variations in economic performance, may occur in intra-state spatial scale and, in some parts over the study region, across the state boundary. Accordingly, attempts to delineate commuting zones or functional regions more accurately (Cromartie and Swanson 1996; Morrill et al 1999; compare them with a Ghelfi and Parker 1997) obviously conform to the fact that commuting as a spatial process operates at a finer level than a county. Numerous methods have been proposed to identify the spatial scale of a process (for a review, Cliff and Ord 1981:123-127).

Study region, spatial unit, and operational spatial scale collectively dictate a spatial setting for a research in an interactive manner. What is more important, however, is the fact that different spatial settings may draw different results from the same data. The concept of scale dependency captures the situation that numerical or statistical properties of a spatial phenomenon vary with the geographical scale at which it is represented and analyzed. A significant spatial pattern at one geographical scale may evaporate or even disappear at another. Further, statistical relationships among variables in one spatial scale could become greater, lesser, or even reversed at other spatial scales. All these spatial effects fall under the ‘modifiable areal unit problems (MAUP)’.

Although MAUP was coined by Openshaw and Taylor (1979), its effects have long been recognized (Gehlke and Biehl 1934). The crux of MAUP is contained in its name. As Holt et al. (1996:181) point out, “the spatial areas are termed modifiable because the choice of area boundaries and the number of areas used to cover the
MAUP can be seen as pointing to "the sensitivity of analytical results to the definition of units for which data are collected" (Fotheringham and Wong 1991:1025). This, further, means that research results at a given spatial configuration are not decisive, but somewhat provisional, because the results could vary with the observational level (the scale effect) and with the configuration of the zoning system (the zoning effect or aggregation effect) (Openshaw 1984; Fotheringham and Wong 1991). Although the scale effect is a most obvious manifestation of MAUP, the zoning effect also makes a significant contribution to MAUP, especially when raw spatial units are aggregated into the same or similar number of higher-order aggregates.

Generally known MAUP effects are: as spatial aggregation proceeds from smaller spatial units to larger spatial units, (i) variance of a variable decreases (Fotheringham and Wong 1991; Wong 1996); (ii) correlation between two variables increases (Gehlke and Biehl 1934; Openshaw and Taylor 1979; Openshaw 1984; Fotheringham and Wong 1991; Amrhein 1994; Wong 1996); (iii) coefficient of determination increases (Amrhein 1995; Wong 1996); and (iv) spatial autocorrelation decreases (Chou 1991; 1995; Anselin and Getis 1992).

Researchers working on a regional analysis on a country, e.g. U.S., are often faced with questions in establishing a spatial setting for their research: what spatial unit should be used; at what spatial scale data are available; what kinds of aggregation schemes have been proposed. The phenomenon having been set at regional economic performance, a viable spatial unit could be a regional labor market area where a vast
majority of people live and work, and an intra-regional functional integration is
distinctive at a large degree. Using county as a raw spatial unit to construct a labor
market area or functional region may be acceptable in the context of U.S. Since it would
be another research question to construct a new regionalization scheme, the next question
might be what regionalization scheme should be ‘chosen’ among preexisting different
functional regions.

Lee (1999) investigated some aspects of the MAUP with county-based functional
regions in the US. He identified 17 functional regions available for research. His results
show that there are substantial differences in a variety of statistics among different
regionalization schemes. Given the situation that a spatial unit is dictated by data
availability, how could we possible report research results with a considerable
confidence? Some efforts have been done to provide a way of finding the most adequate
spatial aggregation level or spatial scale (Moellering and Tobler 1972; Openshaw 1977;
Wrigley 1995). However, a better way might be to apply statistical techniques less
sensitive to MAUP to spatial data such that results are less variant among different spatial
scales and configurations. The alternative may be spatial statistics. Green and
Flowerdew (1996) demonstrate that a certain form of a spatial autoregressive model is
resistant to the effect of MAUP.

What could cause MAUP? A simple and general statistical answer is that, as
spatial aggregation proceeds, a ‘smoothing effect’ happens so that the uniqueness of each
area and dissimilarity among areas are reduced, and variance for the entire research area
is suppressed (Fotheringham and Wong 1991; Wong 1996). Reduced variance, then,
leads to increases in correlation and goodness-of-fit of a regression equation. However, a spatial statistical answer is different. The effects of the MAUP occur because spatial data are usually characterized by spatial dependence that can be defined as the propensity for nearby locations to influence each other and possess similar attributes (Anselin 1988; Anselin and Griffith 1988; Anselin 1990; Haining 1990; Goodchild 1992; Anselin and Getis 1992). If values are randomly distributed across space (spatial randomness), variance will never significantly change with any level or way of spatial aggregation (Arbia 1989). Different spatial aggregations may result in different level of spatial dependence, differently being subject to MAUP’s effects. Since spatial statistical methods are designed to deal effectively with the spatial dependence, they are much more resistant to the effects. All these things lead us to the second conceptual element: spatial dependence.

2.1.2 Spatial structures and processes

A study area \( \mathcal{R} \) is seen as a set of spatial objects or entities that are topologically related to one another. These topological relationships among observations in a given study area may be called spatial structure (Gatrell 1983). More formally, “spatial structure is regarded here as a linking substance and functional connection between interrelated spatial objects on which the spatial process evolves with a given strength (Tiefelsdorf 2000:1).” However, this concept needs to be elaborated further. In absolute space, relationships among spatial entities are reduced to their geometrical connectedness, whereas, in relative space, they depend upon a comparative evaluation of
selected attributes associated with spatial objects (Tiefelsdorf 2000:23) and can be underlain by spatial processes. Thus, an observed spatial pattern can be seen as an overlay of often multiple spatial structures each of which reflects a particular spatial process.

In the absolute view of spatial structure, a spatial structure is characterized by various forms of spatial relations that are defined on the Cartesian product of \( \mathcal{X} \), that is, ordered pairs obtained by taking a spatial object of \( \mathcal{X} \) for each member of the pair (Gatrell 1983:13; Tiefelsdorf 2000:25). Whatever types of spatial entities, lattice or point, are considered, spatial relations are formally defined in a spatial weights matrix. Whereas a spatial weights matrix takes a form of contiguity or connectivity matrix in lattice or areal data, it is represented by a distance matrix in point data (see Tiefelsdorf 2000:26, Fig.3.1). Spatial conversion makes it possible to transform a dimension of spatial entities to another. For example, areal data can be converted to point data by way of extractions of centroids; point data can be transformed to areal data by constructing Voronoi polygons around points (Boots 1986).

A general spatial weights matrix, \( V \), containing information on spatial proximity among spatial objects can be given:

\[
V = \begin{bmatrix}
    v_{11} & v_{12} & \cdots & v_{1n} \\
    v_{21} & v_{22} & \cdots & v_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    v_{n1} & v_{n2} & \cdots & v_{nn}
\end{bmatrix}
\]  

(2.1)
Various ways of defining $V$ have been proposed, each of which proposes a particular type of coding scheme. In the context of lattice data, two coding schemes have been most commonly used: $C$ and $W$. In a $C$ matrix, an entry is given a value of 1 when two spatial objects share a common boundary and 0 otherwise. A $W$ matrix is a row-standardized version of $C$ such that each element is divided by a row-sum. In addition to these coding schemes, Tiefelsdorf et al. (1999) proposes a $S$-coding scheme. In these schemes, diagonal elements are set to zeroes. Throughout this dissertation, I also utilize two derivatives of $C$ and $W$ matrices for supplementary purposes. A $C^*$ is a $C$ with 1s on its diagonal and a $W^*$ is a row-standardized version of $C^*$. These additional coding schemes are introduced to examine whether generalized significance testing methods that I will develop properly deal with a non-zero diagonal in a spatial weights matrix, and to evaluate behaviors of SAMs in terms of their performance in detecting spatial patterns. When $V$ is not symmetric, it can be made symmetric by a transformation function:

$$V^s = \frac{1}{2}(V + V^r)$$

(2.2)

Note that $C$ and $C^*$ are always symmetric but $W$ and $W^*$ are not. This symmetric transformation function should be utilized to compute distributional quantities in significance testing, which will be elaborated in Chapter 6.

It should be noted that a mathematical transformation of $V$ brings additional information on topological relationships among spatial entities. For example, powering of a spatial weights matrix is associated with deriving higher order spatial lags (Anselin
When \( V \) is symmetric, \( V^2 \) or \( V^TV \) contains information on a second-order spatial lag, that is, second nearest neighbors. This transformation will be seen in formulate new SAMs in Chapter 3 and 4.

While a matrix \( V \) represents a global spatial structure, each row in the matrix contains information on a local spatial structure that refers to spatial proximity between a spatial object and all the other objects. Particularly in \( C \) and \( W \) matrices, each row defines a local set that is composed of a reference object and its neighbors, and can be given:

\[
V_i = \begin{bmatrix}
0 \\
v_{ii} & \cdots & v_{ii} & \cdots & v_{im} \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & v_{mi} & \cdots & v_{mn}
\end{bmatrix}
\tag{2.3}
\]

This specification is different from Tiefelsdorf (1998)'s symmetric star-shaped form. I prefer this specification mainly because it works not only for univariate SAMs, but also for bivariate SAMs. This specification is crucial in elaborating significance testing methods for local SAMs, which will be seen in Chapter 5 and 6.

In the relative view of space, spatial structures are assumed to be underlain by spatial processes that can be seen as numerical or statistical abstractions of real socio-economic processes deploying through space. Those spatial processes may be categorized into spatial diffusion, spatial exchange (e.g., spillover and externalities),
spatial interaction, and spatial dispersal or relocation (Haining 1990:24-25). One may further classify those spatial processes into two categories: contagious and sporadic spatial processes. The former is associated with processes based on spatial proximity among regions; the latter based on other forms of spatial interaction, such as hierarchical diffusion and backwash effects, leading to spatial dispersion. It should be noted that contagious spatial processes do not necessarily induce spatial clustering. Some processes, such as contagious diffusion and spillover effects, may result in spatial cluster, but other processes, such as spatial repulsion and spatial competition, may induce spatial dispersion. Spatial patterns under investigation may be seen as particular interactions among those spatial processes, and should be analyzed based on proper assumptions on those processes, that is, certain spatial stochastic processes.

2.1.3 Spatial dependence

It may be interesting observation that, even though spatial dependence is more complicated (multi-dimensional) and prompting than serial dependence (one-dimensional), researchers dealing with spatial data seem less sensitive to the matter than ones with serial data (Anselin and Griffith 1988; Anselin 1990; Fotheringham 1993; Griffith 1993). Spatial dependence is what 'First Law of Geography' (Tobler 1970:236) implies: "everything is related to everything else, but near things are more related than distant things." This spatial dependence is not exceptional in the context of geographically referenced data and obviously contradicts the usual assumption of independent observations in general statistics. Spatial dependence results not only from
intrinsic spatial processes such as spatial interaction, spatial externalities, and other spillover effects, but from mismatch between the scale of the spatial unit (observational scale) and the phenomenon of interest (operational scale) (Anselin and Griffith 1988; Anselin 1990; Anselin and Getis 1992).

At the heart of the problem lies the loss of information that an observation carries. When spatial dependence dominates, information from observations is less than would have been obtained from independent observations, because a certain amount of the information carried by each observation is duplicated by other observations in the cluster (Haining 1990:40-41; Anselin 1990). This loss of information may invalidate some of statistical tests, because it lowers the effective number of degrees of freedom in a test (Goodchild 1996:244). In the context of OLS regression, the presence of spatial autocorrelation misleads significance tests and measures of fit (Anselin and Griffith 1988:16; Fotheringham and Rogerson 1993:11). A statistically significant spatial autocorrelation in OLS residuals could be caused by model misspecifications that one or more significant variables are missing (Getis 1999:241), but mostly it results from intrinsic properties of spatial data.

In the univariate context, spatial dependence of a mapped variable is usually captured by global univariate SAMs, more often known as spatial autocorrelation indices, such as Moran’s I and Geary’s c (Moran 1948; Geary 1954; Cliff and Ord 1981; Goodchild 1986; Griffith 1987; Odland 1988). Both measure how (dis)similar neighbors as defined in a spatial weights matrix are in terms of single variable, simply capturing the level of spatial clustering of a variable.
Researchers often use variance or the coefficient of variation to estimate the level of spatial disparity using concepts clearly invoking spatial clustering or dispersion, such as regional convergence and divergence. In essence, however, numeric variance has nothing to do with spatial pattern, because a numeric vector with \( n \) observations of different values can generate \( n! \) different spatial patterns which have the same mean and variance, but are different in the level of spatial autocorrelation (Lee 2001b).

The same problem occurs when the index of dissimilarity (ID) is used to measure spatial exclusion among socially defined groups. Since there is no spatial element in the equation for ID, totally different spatial association between two groups have the same ID. This problem has long been recognized, and efforts to devise spatial version of ID have been done (e.g., Morgan 1983; White 1983; Massey and Denton 1988; Morrill 1991; Wong 1993; Wardorf 1993; Chakravorty 1996; Lee and Culhane 1998). Accordingly, other univariate statistics should be modified to deal with spatial data, such as spatial ANOVA (Griffith 1978; 1992) and the spatial chi-square test (Rogerson 1998; 1999).

In bivariate research, the same problem occurs. Significance testing for Pearson’s correlation coefficients may be flawed if any of or both of the two variables are spatially autocorrelated (Bivand 1980; Richardson and Hemon 1981; Clifford and Richardson 1985; Haining 1991). It should be noted that the concept of ‘association’ means two things: (i) point-to-point association or pairwise association; (ii) spatial association among observations. Pearson’s correlation coefficient only measures the first association, with an assumption of no spatial association among observations. However,
It is often found that similar point-to-point associations are spatially clustered, that is, bivariate spatial dependence. Spatial dependence in association between two variables should reduce the degree of freedom, and thus the critical value for significance testing at any given confidence level should be adjusted higher in absolute values. Another way to demonstrate bivariate spatial dependence is to indicate that, from a pair of numeric vectors with \( n \) observation, \( n! \) different pairs with the same Pearson's \( r \) but different spatial associations can be drawn (Haining 1991). We need a new spatial correlation coefficient which differentiates the \( n! \) associations by measuring both point-to-point association and spatial association.

In the multivariate situation, spatial dependence in residuals invalidates much significance testing for parameter estimators and goodness-of-fit. A simple way to deal with multivariate spatial dependence might be to conduct a spatial filtering to eliminate the spatial dependence (Getis 1990; 1995; Griffith 2000). A better or more suitable way, however, is to fit a spatial autoregressive model (Upton and Fingleton 1985:347-366; Anselin 1988:32-39; Griffith 1988:17-19; Haining 1990:80-90; Bailey and Gatrell 1995:282-289; Tiefelsdorf 2000:43-47). This model cannot be directly fitted in matrix notions as in general OLS model, but requires more computationally demanding maximum likelihood estimation. When a spatial autoregressive model is fitted, the spatial autocorrelation in residuals is usually eliminated and slightly modified set of regression parameters are derived. For example, a commonly used spatial autoregressive model, simultaneous autoregressive (SAR) model or autocorrelated errors model, decomposes residuals into spatially autocorrelated errors and random disturbances with
no spatial autocorrelation, and the former error elements are autoregressed iteratively to estimate a new set of regression parameters. Even with a well-designed spatial autoregressive model, one might wonder if there is another way that allows regression parameters to vary spatially so as to explore spatial variations in multivariate relationships among variables under investigation. This leads to the third conceptual element: spatial heterogeneity.

2.1.4 Spatial heterogeneity

Spatial heterogeneity, or spatial non-stationarity, refers to geographical variations or differentiations of statistical properties of data, which results from intrinsic uniqueness of each point or sub-region (Anselin 1990; Fischer 1999). The presence of spatial non-stationarity invalidates the assumption that all observations in a sample are drawn randomly from the same population (Goodchild 1996). More specifically, complete spatial stationarity refers to a situation where means and standard deviations (univariate), covariances (bivariate), and multivariate parameters (multivariate) are constant from sub-region to sub-region (Haining 1990; Getis and Ord 1996; Fotheringham 1997; Anselin 1999). Complete spatial stationarity or homogeneity is as rare as complete spatial non-stationarity that indicates an absolute uniqueness of each observation. Geographers usually find the entire study area divided into several segments or 'spatial regimes', each of which has a certain level of internal homogeneity and, at the same time, a particular level of external heterogeneity. Detection of spatial regimes, therefore, appears similar to regional classification. Since internal homogeneity is, in turn, associated with spatial
dependence, spatial heterogeneity can be redefined as "a lack of spatial uniformity of the effects of spatial dependence and/or of the relationships between the variables under study" and can be thought of as "a special case of spatial dependence" (Anselin and Getis 1992:24).

With this respect, most spatial statistical techniques so far depend upon an assumption of spatial stationarity or spatial homogeneity in the sense that they focus on deriving global or average statistics, rather than local or deviant statistics. Recent advances in spatial statistics deal with spatial heterogeneity, resulting in a class of statistics called local statistics. Lee (2001a) shows that distance parameters in spatial interaction models vary place to place.

In the context of univariate spatial dependence, three local spatial autocorrelation indices have been articulated, Getis-Ord $G_i$ and $G_i^*$ (Getis and Ord 1992; Ord and Getis 1995), local Moran’s $I_i$, and Geary’s $c_i$ (Anselin 1995), and collectively construct a class of Local Indicators of Spatial Association (LISA) (Anselin 1995; Getis and Ord 1996). Tiefelsdorf (1998) extends local Moran’s $I_i$ further to embrace conditionality of the measure. When local measures are mapped, you can see not only spatial variations in spatial clustering, but also the spatial extent of the process under investigation, that is, operational scale. Further, a distinctive operational scale may help identify localities in terms of variables of interest (Unwin 1996). LISA has been applied to a variety of research topics. Among them, applications to regional economic analysis include: Bao et al 1995; Barkely et al 1995; Bernat 1996; Lopez-Bazo et al 1999; Rey and Montouri 1999; Ying 2000.
In bivariate spatial heterogeneity, a localized spatial correlation index can be suggested. A local spatial correlation index first indicates the relative contribution an individual area makes to a global spatial correlation index: it measures the degree to which an area conforms to a global trend in direction and magnitude of spatial association across two variables. Second, it involves the degree of similarity between an area with its neighbors: areas more similar to their neighbors will be given higher local values than others dissimilar to their neighbors, with point-to-point association being identical.

Methods for multivariate heterogeneity have a relatively long history in comparison to other local statistics. The spatial expansion method, allowing for spatially drifting regression parameters, was devised as early as the 1970s and applied to various research situations (Casetti 1972; Brown and Jones 1985; Brown 1991; Foster 1991; Jones and Casetti 1992; Jones and Hanham 1995; Casetti 1997; Casetti and Can 1999). Other approaches are considered in the same conceptual line, e.g., spatial adaptive filtering (Foster and Gorr 1986; Gorr and Olligschlaeger 1994), and spatial multilevel modeling (Jones 1991a; 1991b; Ward and Dale 1992).

However, a full-fledged spatially varying regression analysis is by the geographically weighted regression (GWR) (Brunsdon et al. 1996; 1998a; 1998b; Fotheringham 1997a; 1997b; 1998; Leung et al 1999). At a glance, this approach is similar either to the weighted least squares (WLS) regression or to the kernel regression. It is different from the former in the sense that weights matrices in WLS are constant across observations, and is different from the latter in the sense that weights matrices are based on spatial proximity (GWR), rather than numerical similarity (the kernel
regression). Also, GWR is different from spatial autoregressive models simply for producing localized parameter estimators. When localized parameters for an explanatory variable are mapped, one may see the relative explanatory power of the variable over space. When coefficients of determination are mapped, one can see that the goodness-of-fit of the specified model vary sub-region to sub-region.

Spatial statistics provides a solid foundation on which spatial data are explored, analyzed, and represented. Especially, local statistics have appeared in a revolutionary fashion. With a univariate local statistics, one can test which labor market areas are significantly prosperous or lagged. With a global spatial correlation index, one can gauge the extent to which socio-economic restructuring has been entailed by spatial restructuring.

With a local spatial correlation index, one can document with a certain level of statistical confidence on which labor market areas have experienced significant continuity or change in terms of economic performance during the regime changing period, 1970 to 1990. With a local spatial regression model, one can examine how a model designed to explain the spatial variation of economic performance spatially behaves over the whole study region in terms of different goodness-of-fit and different associations among variables. Further, this may allow us to explore and visualize the spatial division of causality.
2.2.1 Need for a generalized procedure and its role for ESDA

I contend there has been a lack of generality in conducting significance testing for SAMs. First, little attention has been dedicated to the connection between global and local measures with the few exceptions of Tiefelsdorf and Boots 1997; Tiefelsdorf 1998; Tiefelsdorf 2000; Boots and Tiefelsdorf 2000. Second, investigating distributional properties has never been undertaken for bivariate SAMs. Third, existing procedures are confined to a particular type of spatial weights matrix, that is, one with a zero-diagonal. Thus, we need a generalized significance testing procedure which different measures, whether univariate or bivariate or whether global or local, are commonly predicated on with different spatial settings, whether zero-diagonal spatial weights matrices or not.

Even though significance testing is confirmatory in nature, I argue that well-founded significance testing is necessary for ESDA. First, pattern detection using SAMs will be theoretically more meaningful and practically more efficient if it is guided by a statistical procedure. In the context of global SAMs, significance testing allows researchers to report the overall degree of spatial dependence in a probabilistic fashion. It is more crucial when different spatial patterns are compared to assess which ones are more spatially dependent with a statistical confidence. In the context of local SAMs, the task of exploring mapped local SAMs will be enhanced when they are in conjunction with their $p$-values. For example, a spatial pattern would be more obvious when only observations with a certain level of significance are displayed.
that, as far as local SAMs are concerned, it should be considered to be exploratory in
nature, because there are two crucial pitfalls. First, an alpha-level for a global SAM
should be lowered when it is applied to its local SAMs (Getis and Ord 1992; Anselin
1995). Even though some procedures such as a Bonferroni bounds procedure have been
proposed, this problem has never been solved. Second problem arises because
significance testing procedures for local SAMs are indifferent to the global level of
spatial dependence even though distributional properties change as levels of global
spatial dependence change as discussed (Anselin 1995; Ord and Getis 1995; Tiefelsdorf
1998; Ord and Getis 2001). To data, only the exact distribution approach can solve this
problem, but only for local Moran's $I$ (Tiefelsdorf 1998; 2000). These two problems
collectively dictate a restriction on the use of local SAMs that they should be used in an
exploratory manner, not in a confirmatory manner (Sokal et al. 1998).

2.2.2 Different significance testing methods

Significance testing for spatial association measures can be categorized into three
different approaches: (i) approximation (normality and randomization assumptions); (ii)
exact distribution; (iii) simulation. The approximation approach is further divided into
two classes in terms of whether a population distribution is assumed normal or not. If
observed sample values are assumed to be random independent drawings from one
normal population, the normality assumption applies to provide the first two moments
(Cliff and Ord, 1981). Even though Henshaw (1966; 1968), inspired by Durbin and
Watson 1950; 1951), provides a general procedure to compute the first four moments under the normality assumption, its use has been confined to global Moran's I (Hepple 1998) (see Table 2.1). In Chapter 5, I formulate a general procedure and apply to global Geary's c, local Geary's c, a new global univariate SAM S, and its local version S.

It should be noted that it is very often unsustainable to assume a normal distribution of a population that samples are drawn from. In addition, it would be more intuitive to regard observed sample values as a particular realization of all possible spatial patterns with the sample, than as one out of an infinite number of numerical vectors with the same mean and variance. This leads to the randomization assumption. Cliff and Ord (1981) contend that the randomization approach is preferable either (i) when we consider all possible permutations with a given data set, or (ii) for any non-normal population. The second issue is more crucial because variance computed under the set of random permutations provides an unbiased estimator for the variance of a statistic for any underlying distribution (Cliff and Ord 1981:42). As can be seen from Table 2.1, the randomization approach is further divided into two distinctive assumptions for local SAMs, total randomization and conditional randomization.

In Chapter 6, I propose two generalized randomization significance testing methods, the Extended Mantel Test and a generalized vector randomization test, and demonstrate that the two methods can be applied to any SAMs, univariate or bivariate, global or local, under any assumption, total or conditional randomization. It should be noted, however, that a fundamental notion in the randomization assumption that all
<table>
<thead>
<tr>
<th></th>
<th>Approximation Approach</th>
<th>Randomization Assumption</th>
<th>Exact Distribution Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normality Assumption</td>
<td>Total</td>
<td>Conditional</td>
</tr>
<tr>
<td>Univariate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moran</td>
<td>Global</td>
<td>Cliff and Ord (1981)</td>
<td>N/A</td>
</tr>
<tr>
<td>Geary</td>
<td>Global</td>
<td>Cliff and Ord (1981)</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Local</td>
<td>Sokal et al. (1998)</td>
<td>Sokal et al. (1998)</td>
</tr>
<tr>
<td>Lee's S</td>
<td>Global</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Local</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bivariate</td>
<td>Cross-Moran</td>
<td>Global</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Local</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lee's L</td>
<td>Global</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Local</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Shaded cells are covered by Chapter 5 and 6)

Table 2.1: Significance testing methods for spatial association measures
possible permutations of the regional values are equally likely does not always hold because “permutations with atypically high or low values in the periphery are more likely than permutations with atypically high or low values near the center (Rogerson 2001:171).” This is more obvious when spatial autocorrelation in regression residuals, because a test based on randomization assumption ignores autocorrelation among regressor variables so that random permutations do not constitute an appropriate reference set for testing regression residuals (Cliff and Ord 1981:200).

The exact distribution approach (Tiefelsdorf and Boots, 1995; Hepple, 1998) is superior to the approximation approach in the sense that it can deal with other aspects of a sampling distribution (i.e., skewness and kurtosis). The normal approximation with the first two moments on which the approximation approach is usually based often appears flawed even with a large sample size (Siemiatycki, 1978; Mielke, 1979). Moreover, even with higher moments, the approximation does not always yield accurate probability values (see Costanzo et al., 1983; Hepple 1998). The exact distribution approach is parametric along with the approximation approach based on normality assumption in the sense that they are built on a particular population distribution, that is, normal distribution. In contrast, the approximation approach based on randomization assumption is non-parametric simply it does not assume any population distribution such that it is a distribution-free testing method.

In spite of its superiority in an inferential test for SAMs, the exact distribution approach has some drawbacks. First of all, the assumption of the normal distribution is still required so that it may not work properly in situations where a normality assumption
is hardly sustainable. Secondly, it is computational more intensive in comparison with
the normal approximation, although other approximation methods could alleviate the
computational burden substantially (see Tiefelsdorf 2002).

The simulation approach, including a Monte Carlo test (Cliff and Ord 1981:63-65,
can be seen as supplementary to the approximation approach. Two different simulation
designs could conform to the two approximation assumptions above: if a number of
numeric vectors with the same mean and variance as a given sample are randomly
generated, a set of statistics will be obtained for the normality assumption; in contrast, if
a number of different orders of a given sample are permuted, a set of resulting statistics
conforms to the randomization assumption. Although a Monte Carlo test could provide
more accurate p-values than the normal approximation with first two moments, especially
when an abnormal skewness or kurtosis is present, it is supplementary to the
approximation approach, as long as a set of equations for distributional moments are
known.

2.3 An ESDA-GIS framework and spatial association measures

2.3.1 CSDA, ESDA, and GIS

SDA may be divided into three categories; exploratory spatial data analysis
(ESDA), confirmatory spatial data analysis (CSDA), and prescriptive spatial data analysis
(PSDA) (Unwin 1996:510). Tasks of SDA may include, according to Fischer
(1999:284): (i) detection of patterns in spatial data; (ii) exploration and modeling of
relationships between such patterns; (iii) enhanced understanding of the processes that might be responsible for the observed patterns; and (iv) improved ability to predict and control events arising in geographical space. It seems that (i) and a half of (ii) pertain to ESDA, the other half of (ii) and (iii) to CSDA, and (iv) to PSDA. The distinction between ESDA and CSDA has been based on a dichotomy between data-driven and model-driven (Anselin 1990; Openshaw 1990; Anselin and Getis 1992), and sometimes based on one between inductive and deductive (Openshaw 1990). According to Haining (Haining 1990; Haining et al. 1998; 2000a; 2000b), ESDA is the extension of exploratory data analysis (EDA), and its aims are descriptive, seeking to detect patterns in spatial data, to formulate hypotheses, and to assess statistical models for spatial data. In contrast, CSDA is the extension of confirmatory data analysis (CDA), and its aims include testing hypotheses and fitting models that are explicitly spatial in the sense that spatial dependence is incorporated in the model specification. It should be noted, however, that the distinction between ESDA and CSDA is often blurred (Anselin and Getis 1992; Bailey 1994). Especially, it would be more so if a distinction between pre-confirmatory ESDA (before hypothesis formulation) and post-confirmatory ESDA (after hypothesis formulation) is introduced (Fotheringham and Charlton 1994). Further, hypothesis testing on LISA, as an important source for ESDA, has always been an issue (Anselin 1995; Ord and Getis 1995; Bao and Henry 1996).

I suggest, nevertheless, that the distinction is still of value, and ESDA is more needed for SDA than CSDA is. There are two reasons. First, ESDA is more congruent with the nature of spatial data, i.e., spatial dependence, spatial heterogeneity, and spatial
outliers. These spatial effects are simply implicated in CSDA. Some CSDA techniques such as spatial autoregressive models (Anselin 1988) and spatial ANOVA (Griffith 1978; 1992) may alleviate the effects in model specifications, but do not provide a way of revealing and exploring them for further insights. Second, ESDA is more congruent with current research platform, i.e. GIS. Since one of the major aims of ESDA is to detect spatial patterns by using visualization techniques, ESDA can take more advantage of GIS’s capabilities in visualization and spatial data mining (Fotheringham and Charlton 1994).

According to Bailey (1994:21), the value of GIS to SDA is: (i) flexible ability to geographically visualize both raw and derived data; (ii) provision of flexible spatial functions for editing, transforming, aggregating and selecting both raw and derived data; and (iii) easy access to spatial relationships between entities in the study area. All these benefits from integration between GIS and SDA more pertain to ESDA. In a practical sense, the only CSDA needs from GIS is the spatial weights matrix. Here, discussions on which SDA functions are more relevant to GIS environments may provide a good foundation.

10 GISable SDA techniques proposed by Openshaw (1990) and advocated (Fischer and Nijkamp 1992; Bailey 1994; Openshaw and Clarke 1996; Unwin 1996) are more related to ESDA, rather than CSDA. Openshaw and Clarke (1996:32) contend that “future GISable spatial analysis methods will be essentially descriptive, exploratory, and probably not inferential in a traditional spatial hypothesis testing sense.” Further, it is also noteworthy that “it is not necessary to use a GIS to perform spatial analysis and that
 Integrating the two will not necessarily lead to any greater insights into geographical theory. However, “under certain circumstances, the integration of GIS and spatial analysis will have a reasonable high probability of producing insights that would otherwise be missed” (Fotheringham 1992:1675-6; Fotheringham and Charlton 1994:316). I suggest that the circumstances are more likely to happen to ESDA than to CSDA.

ESDA inherits many properties from EDA that can be defined as “detective work” (Tukey 1977:1) and “an intermediate or soft statistics between descriptive and inferential or hard statistics” (Good 1983:291), and a bundle of statistical and graphical techniques that enhance a researcher’s intuition into data by utilizing a variety of visual representations. EDA techniques require relatively few, and weaker, assumptions and are resistant to outliers or atypical observations (Tukey 1977; Good 1983; Hamilton 1992). Some major concepts of EDA, such as brushing, conditioning, and spinning have been translated into the context of spatial data. For example, brushing techniques are to make connections among graphs and data tables such that one selection of point(s) in a window should simultaneously induce a selection for the corresponding data point(s) in other windows (McDonald 1982; Becker and Cleveland 1987). This technique is translated into ‘geographical brushing’ (Monmonier 1989) or ‘spatial windowing’ (Fotheringham and Charlton 1994) where a map window is connected to graph and data windows such that any selection in the map window makes subsequent selections in other windows, and vice versa. This technique has played a central role in conceptualizing and implementing ESDA (Haslett et al. 1990; 1991; MacDougall 1992; Symanzik et al. 1994; 1996; Majure
The use of other graphical techniques, such as box plot, qq plot, trellis graph, Chernoff faces plot, Tukey's star diagram, scatterplot matrix, and biplot, has been advised for spatial data. As mentioned before, I more focus on ESDA techniques based on spatial data analysis or spatial statistics, because ESDA techniques are basically aspatial, and their translations to spatial data are far from a true 'spatial' EDA (Anselin and Getis 1992:25).

I define ESDA, following Anselin (1994; 1998), as "a collection of techniques to describe and visualize spatial distributions, identify atypical locations or spatial outliers, discover patterns of spatial association, clusters or hot spots, and suggest spatial regimes or other forms of spatial heterogeneity". Several ESDA frameworks for a GIS environment have been proposed (Openshaw 1990; Goodchild et al. 1992; Fotheringham and Charlton 1994; Openshaw and Clarke 1996; Anselin 1998; Wise et al. 1999). Among them, I choose Anselin's framework (Anselin 1998:81 Table 5.1). He divides tasks for ESDA into four categories (visualizing spatial distribution, visualizing spatial association, local spatial association, and multivariate spatial association) and allocates relevant ESDA techniques to each. These ESDA techniques include some geostatistical techniques such as variogram (Cressie 1993), variogram cloud (Cressie 1993; Majre et al. 1996; Majure and Cressie 1997), pocket plots (Cressie 1993), variogram boxplot (Majure and Cressie 1997; Kaluzny et al. 1998), spatial lag scatterplot (Cressie 1993; Majre & Cressie 1997), and some lattice techniques such as spatial lag scatterplot (Fotheringham and Charlton 1994), spatial lag pie/bar charts (Anselin et al. 1993; Anselin 1994; Anselin and Bao 1997), Moran scatterplot and scatterplot map (Anselin 1994; 1995; Anselin and
significance map (Anselin 1995; 2000). This study largely follows this Anselin’s framework with more focusing on ESDA techniques based on local SAMs.

2.3.2 Global and local spatial association measures

The importance of local statistics is straightforwardly derived from limitations of global measures, or parameters. Global spatial measures, from spatial autocorrelation coefficients to regression parameters, are based on an assumption of spatial stationarity (Anselin 1996; Unwin 1996; Anselin and Bao 1997; Fotheringham 1997). According to Fotheringham (2000:71), “the raison d’être for the development of local statistics is the low probability in many situations that the ‘average’ results obtained form the analysis of a spatial data set drawn from a broad region apply equally to all parts of that region, the assumption of traditional global statistics.” It is ironic that, albeit a strong tradition of areal differentiation, quantitative geography has focused on spatial similarities rather than spatial differences, global generalities rather than local exceptions, and ‘whole-map’ values rather than mappable statistics (Fotheringham 2000). In conjunction with ESDA, major objectives of local statistics include: (i) identifying atypical locations (spatial clusters); (ii) discovering significant local spatial association (spatial clusters or hot spots); (iii) detecting local pockets of non-stationarity (spatial regimes) (Anselin 1995; 1999; Getis and Ord 1996). These are correspondent to what Fischer (1999:285) refers to as ‘spatial dependence and heterogeneity descriptors’. In addition, the integration of GIS and ESDA obviously favors local statistics rather than global ones (Openshaw 1990;
Openshaw & Clarke 1996; Anselin 1996). Anselin (1996:113) points out that "the focus of ESDA techniques used in conjunction with a GIS should be on measuring and displaying local patterns of spatial association, on indicating local non-stationarity, on discovering islands of spatial heterogeneity and so on."

Trends toward local statistics are not confined to SAMs (see Fotheringham and Brunsdon 1999). For example, place-specific distance parameters in spatial interaction models can be seen as local statistics (Fotheringham 1981; Stillwell 1991; Tiefelsdorf and Braun, 1997; 2001; Lee 2001a). However, I focus on local SAMs in this dissertation. As mentioned before, three univariate local spatial association measures have been proposed, Getis-Ord G_i and G_i* (Getis and Ord 1992; Ord and Getis 1995), and local Moran's I_i, and Geary's c_i (Anselin 1995), and collectively construct a class of LISA (Local Indicators of Spatial Association) (Anselin 1995; Getis and Ord 1996). Anselin (1996) subsequently developed Moran scatterplot and related mapping techniques for local Moran's I_i (Anselin 1996). Some issues on distributional properties and hypothesis testing for local Moran's I_i have been extensively discussed (Anselin 1995; Bao and Henry 1996; Tiefelsdorf and Boots 1997; Sokal et al. 1998; Haining 2000a; Tiefelsdorf 2000). Those issues will be addressed later on. The LISA has been applied to a variety of research topics (Table 2.2).

Obviously, local statistics for ESDA is not confined to univariate SAMs. A global bivariate spatial association measure that can be decomposed was proposed (Wartenberg 1985) and advocated (Griffith 1993; 1995). Geographically weighted regression scheme as a multivariate local spatial measure was also proposed and
<table>
<thead>
<tr>
<th>Subjects</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural hazard analysis</td>
<td>Pereira et al. 1998</td>
</tr>
<tr>
<td>Spatial interaction</td>
<td>Berglund and Karlstrom 1999</td>
</tr>
<tr>
<td>Social and political studies</td>
<td>O'Loughlin et al 1994; Talen 1997; Talen and Anselin 1998; Gleditsch and Kristian 2000</td>
</tr>
<tr>
<td>Modifiable areal unit problem</td>
<td>Amrhein and Reynolds 1996; 1997</td>
</tr>
<tr>
<td>Spatial epidemiology</td>
<td>Ord and Getis 1995; Tiefelsdorf 1998; 2000</td>
</tr>
<tr>
<td>Image data processing</td>
<td>Getis 1994; Wulder 1999</td>
</tr>
<tr>
<td>Real estate</td>
<td>Can 1998; Paez, et al. 2001</td>
</tr>
<tr>
<td>Urban population</td>
<td>Wang and Meng 1999</td>
</tr>
</tbody>
</table>

Table 2.2: Studies utilizing local univariate spatial association measures
elaborated (Brunsdon et al. 1996; 1998a; 1998b; 1999; Fotheringham et al. 1997a; 1997b; 1998). Despite the presence of other local statistics such as local point pattern analysis, spatial expansion method, and adaptive filtering, this dissertation exclusively focuses on local SAMs for areal data. In order to devise local spatial measures, one may need to set up several criteria. Any kind of measure should have one or two or all the three properties; (i) it is a statistically processed or derived value from raw data; (ii) it is derived in a disaggregated fashion; thus, it is another variable; (iii) it contains information on topological relationships among observations.

From those criteria, some aspatial and global measures such as mean, Pearson's correlation coefficient, and regression coefficient satisfy only (i). (i) and (ii) collectively confine aspatial and local measures such as z-score vector, local Pearson's correlation coefficient vector, and regression residual or factor score vector. A combination of (i) and (ii) characterizes spatial and global measures such as Moran's $I$ and bivariate Moran's $I$ (Wartenberg 1985) or cross-Moran (Griffith 1993; 1995), and parameters of spatial autoregressive models. Only spatial and local measures meet all three criteria, e.g. local SAMs.

2.3.3 A SAM-based ESDA-GIS framework

An ESDA-GIS framework is defined as a GIS-based research platform equipped with ESDA techniques. Developments of the ESDA-GIS framework are strongly connected to emergence of GIS as a general purpose platform for SDA (Haining et al. 2000a), where geographically referenced data are stored, retrieved, managed, analyzed,
and visualized in a truly spatial way. An ESDA-GIS framework based on SAMs is largely characterized by a continuous interaction between GIS and ESDA techniques (Figure 2.1).

First, an SDMS (Spatial Database Management System) module in GIS provides information on topological relationships among observations. The information could take a form of either vectors or matrices. A vector format such as sparse contiguity matrices (GAL) or sparse general weights matrices (GWT) (Anselin and Bao 1996) can be transformed to a matrix format by way of a matrix conversion function in a statistics module of ESDA.

Second, attributes associated with spatial entities can be exported and imported between GIS and ESDA by way of a data transfer protocol. A DMS (Database Management System) in an object-oriented ESDA program such as S-Plus can store and manipulate matrices along with other forms of data such as multi-layered arrays and lists.

Third, a manipulation module in GIS accomplishes spatial aggregation and spatial conversion that transforms dimensions of spatial entities, e.g., creating centroids from polygons or constructing Voronoi polygons from points. These procedures restructure topological relationships among spatial objects and change their attributes.

Fourth, ESDA computes SAMs and carries out significance tests. Local SAMs can take various forms: (i) row values; (ii) nominal values by way of a classification scheme; (iii) probability values by way of a significance test; (iv) a combination of (ii) and (iii).
Figure 2.1: A SAM-based ESDA-GIS framework
Fifth, local SAMs are exported to a scientific visualization module in ESDA as well as a cartographic visualization module in GIS. In the former, various ESDA graphics such as scatterplot, boxplot, etc. and, in the latter, local SAMs are mapped to allow for an exploration of spatial patterns.

Recent efforts to integrate univariate SMAs with GIS platforms can be seen as good examples of an ESDA-GIS framework. How to integrate GIS and ESDA has been an issue. The first way is to make a module for local SAMs in aspatial statistical package using script languages packages provide. For example, Bivand and Gebhardt (2000) developed a module for spatial statistics, including local statistics in R language. Some approaches, although they do not include functions for local statistics, fall into this category (e.g., SPLANCS for S language (Rowlingson and Diggle 1993); spatial autoregressive library for MATLAB (LeSage 1999); special scripts of spatial autocorrelation for MINITAB and SAS (Griffith 1988), for SAS (Griffith 1993), for SPSS (Tiefelsdorf and Boots 1995), and SAS and SPSS (Griffith and Layne 1999)).

The second way is to use stand-alone ESDA or SDA programs. For example, stand-alone programs implement local SAMs (Data Desk (Wilhelm and Steck 1998), REGARD (Unwin 1996), CrimeStat (Levine 1999), and Tcl/Tk/cdv (Dykes 1998)).

The third way is to customize GIS programs by developing script codes for local statistics, without connection to statistical programs or languages. For example, Zhang and Griffith (1997) and Hansen (1997) develop Avenue scripts for local statistics. Similarly, Ding and Fotheringham (1992) and Bao et al. (1995) implemented some of local statistics in ARC/INFO by way of AML.
The fourth way is to construct an ESDA-GIS platform that connects GIS and some other programs, usually statistical, by means of RPC (Remote Procedure Calls for UNIX), DDE (Dynamic Data Exchange for Window), or ActiveX for Microsoft Windows environment. In the context of local SAMs, at least 6 programs have been developed. These include SpaceStat-ArcView (Anselin and Bao 1996; 1997; Anselin 1998; 2000), R-GRASS (Bivand 2000), SAGE-ARC/INFO (Haining et al., 1996; 1998; 2000a; 2000b), SPlus-ArcView (Kaluzny et al. 1998; Bao et al. 2000), AWK/GMT/GRASS (Bivand 1997), and MicroSoft Access-MapObjects (Zhang and Griffith 2000).

One crucial implementation issue is raised in the context of this study. In situations where a researcher develops a set of statistical and graphical techniques and wants to connect with a GIS program for visualizing and exploring the mappable results, which way could be most viable? This is extremely important, because “ESDA ought to concern itself with the implementation of algorithms, not just their elaboration and the purchase of products claiming to include them” (Bivand 1998:500). It seems necessary to regard GIS as a general purpose platform (Haining et al. 2000a) and connect ESDA with it in order to take advantage of its full functionalities (Anselin 1999). It seems untenable to completely depend on a GIS so that new algorithms are made available within it by way of a script language the GIS provides, mainly because the GIS script language is not effective for intensive computations and quality-graphics. It also seems cumbersome to build a completely new platform for the ESDA-GIS integration, not only because it is not a technically easy task, but because it may prevent researchers from
continuously updating functions and from taking advantage of other integrations which already contain a number of functions. It is also recognized that languages for developing statistical algorithms are *interpreted* such as *Java*, *S*, and *R*, rather than *compiled*, such as *C* and *Fortran*, because the former more allows researchers to interact with data and prototype new algorithms (Bivand 1996; 1997; 1998; Dykes 1998; MathSoft 1999).

The ESDA-GIS platform for this study is composed of (i) a preexisting excellent ESDA-GIS integration, *SPlus-ArcView*; (ii) a bundle of *S*-scripts that generate various types of spatial weights matrices, calculate SAMs, and draw statistical graphs. A similar approach can be found in CFGIS-NEA by Pierce et al. (2001) where *SPlus-ArcView* plays a central role with being assisted by Avenue scripts for additional analytical tools and *Microsoft Access* for an efficient database management. Another benefit of using *S* language is that it is excellently performing not only for statistical procedures but for scientific visualization. Quality of graphics in most of GIS programs has been questioned.
3.1 Rationales for univariate spatial association measures

The need for univariate SAMs, often known as spatial autocorrelation indices, has long been recognized. A numeric vector with \( n \) observations with different values can generate \( n! \) different permutations or arrangements, each of which has a distinct order of data points. When referenced by spatial locations, different orders of a numeric vector result in different spatial patterns with different degrees of the univariate spatial dependence or spatial clustering. To illustrate, I generate three different spatial patterns from a numeric vector on a hypothetical space consisting of 37 hexagons (Figure 3.1). The numeric vector has a mean of 1.838 and a variance of 0.514. Since the spatial patterns are three out of all possible \( 37!/(7!17!13!) \) geographical variables, they share the same numerical properties. Differences in the univariate spatial dependence among the three patterns can only be computed by global univariate SAMs.
Figure 3.1: Three spatial realizations of a hypothetical numeric vector

Mean: 1.838
Variance: 0.514
In the context of local SAMs, each locale or local set consisting of a reference area and its neighbors has different degree of local spatial dependence. In other words, each locale tends to be unique in terms of the degree to which the spatial dependence in the particular location conforms to the overall global spatial dependence that is captured by the corresponding global SAM. Further, some locales show a higher level of internal homogeneity (local homogeneity). Some locales can be characterized by a higher level of deviation from the overall mean (spatial clustering): either clusters of higher-than-average values (hot spots) or clusters of lower-than-average values (cold spots). In contrast, reference areas in some locales may be extremely dissimilar to its neighbors (spatial outliers). Note that spatial outliers are only defined by relationships between a reference area and its neighbors, not by overall heterogeneity within a locale. All these different kinds of variability among locales in terms of the spatial dependence can be conceptualized as the univariate spatial heterogeneity.

Two global univariate SAMs have long been devised: Moran's $I$ (Moran 1948) and Geary's $c$ (Geary 1954), and earlier works on SAMs have focused only on these two measures (Cliff and Ord 1981; Goodchild 1986; Griffith 1987; Odland 1988). Their corresponding local SAMs (Anselin 1995) and newly developed local measures, $G_{ij}$ and $G_{ij}^*$ (Getis and Ord 1992; Ord and Getis 1995), collectively constitute a general class of local indicators of spatial association (LISA). Table 3.1 summarizes the equations of these measures except for Getis-Ord statistics because they do not have corresponding global SAMs. The relationships between global SAMs and local SAMs satisfy a more restrictive additivity requirement that an average value of local SAMs equal to a
<table>
<thead>
<tr>
<th>Univariate SAMs</th>
<th>Equations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Moran</strong></td>
<td></td>
</tr>
<tr>
<td>Global</td>
<td>$I = \frac{n}{\sum_i \sum_j V_{ij} \sum (x_i - \bar{x})^2} \sum_i \sum_j V_{ij} (x_i - \bar{x})(x_j - \bar{x})$</td>
</tr>
<tr>
<td>Local</td>
<td>$I_i = \frac{n^2}{\sum_i \sum_j V_{ij} \sum (x_i - \bar{x})^2} (x_i - \bar{x}) \sum_j V_{ij} (x_j - \bar{x})$</td>
</tr>
<tr>
<td><strong>Geary</strong></td>
<td></td>
</tr>
<tr>
<td>Global</td>
<td>$c = \frac{n-1}{2 \sum_i \sum_j V_{ij} \sum (x_i - \bar{x})^2} \sum_i \sum_j V_{ij} (x_i - x_j)^2$</td>
</tr>
<tr>
<td>Local</td>
<td>$c_i = \frac{n(n-1)}{2 \sum_i \sum_j V_{ij} \sum (x_i - \bar{x})^2} \sum_j V_{ij} (x_i - x_j)^2$</td>
</tr>
</tbody>
</table>

Table 3.1: Univariate spatial association measures
corresponding global SAM. Note that the equation for local Geary's $c_t$ is modified from Anselin's original specification (1995) in order to meet the requirement.

The two measures gauge the univariate spatial dependence differently. A higher local Moran's $I_t$ results more from how a reference area deviates from mean rather than from how similar a reference area is to its neighbors. This means that, even though a locale is characterized by a high level of local homogeneity, a lower local Moran's $I_t$ would be yielded if values in the locale are close to overall mean. In contrast, local Geary's $c_t$ exclusively tackles the local homogeneity in a locale by calculating an averaged difference between a reference area and its neighbors. Thus, Moran's $I_t$ is more suitable to identify spatial clusters and spatial outliers, whereas Geary's $c_t$ is more effective in measuring the local homogeneity within a local set. This brings a fundamental conceptual issue: by what should the univariate spatial dependence be defined, local homogeneity or spatial clustering.

Another issue can be seen from Table 3.1. Both measures are heavily dependent upon a reference area. In local Moran's $I_t$, overall direction is predominantly determined by direction in a reference area. For example, a reference area with a lower-than-average value is given a negative local Moran's $I_t$ even though it is surrounded by neighbors with extremely higher-than-average values, which may imply a positive spatial cluster in the overall local set. In local Geary's $c_t$, a local variance is seen as sum of squared deviations of values in neighbors from a value in a reference area, rather than a local mean. Even though a local SAM is assigned to a reference area, it would be worthwhile for a local
SAM to assess an overall locale in terms of local spatial dependence. All these issues will be discussed later on.

3.2 A global univariate spatial association measure, $S$

Lee (2001) formulates a concept of a spatial smoothing scalar in relation to Moran’s $I$ and suggests that it can be used as a univariate spatial association measure.

Now, a spatial smoothing scalar denoted by $S$ is given:

$$S_x = \frac{n}{\sum_i \left( \sum_j v_{ij} (x_j - \bar{x}) \right)^2} \cdot \frac{\sum_i \left( \sum_j v_{ij} (x_j - \bar{x}) \right)^2}{\sum_i (x_i - \bar{x})^2}$$  \hspace{1cm} (3.1)

where $v_{ij}$ is an entry in a general spatial weights matrix $V$. When a row-standardized spatial weights matrix $W$ is applied, equation (3.1) is simplified to:

$$S_x = \frac{\sum_i (x_i - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}$$  \hspace{1cm} (3.2)
where $x_i$ denotes a spatially smoothed value at the $i$th location that can be calculated by

$$\sum_j w_{ij} x_j$$

(see Lee 2001b). Conceptually, it corresponds what has been termed as a spatial lag (SL) or spatial moving average (SMA) that computes a localized mean value.

As Lee (2001) shows, equation (3.2) is approximately equivalent to:

$$\frac{\sum_i (x_i - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}$$

(3.3)

where $\bar{x}$ is a mean value of a SL or SMA vector that is denoted by $\bar{x}$. Thus, (3.3) is defined as a ratio of variance of a SL or SMA vector to variance of the original variable $X$.

Further, a decomposition of Moran's $I$ shows (Lee 2001b):

$$I_X = \frac{\sum_i (x_i - \bar{x})^2}{\sum_i (x_i - \bar{x})^2} \cdot r_{x,\bar{x}}$$

(3.4)

From (3.4), Moran's $I$ is seen as a Pearson's $r$ between a variable and its SL or SMA scaled by the square root of the ratio of the SL's or SMA's variance to the original variable's variance (or the ratio of SL’s or SMA’s standard deviation to the original variable’s standard deviation). The derivation corresponds to a well-known finding that Moran's $I$ is a regression coefficient when a variable's SL is regressed on the original variable.
variable (Anselin 1995; Griffith and Amrhein 1997). By utilizing the general relationship between a regression coefficient in a bivariate regression and Pearson's $r$ between two variables, (3.4) is easily proved.

Equation (3.4) is further decomposed to:

$$I_x = \sqrt{\frac{\sum (\bar{x}_i - \bar{x})^2}{\sum (x_i - \bar{x})^2}} \cdot \sqrt{\frac{\sum (\bar{x}_i - \bar{x})^2}{\sum (x_i - \bar{x})^2}} \cdot r_{x, \bar{x}}$$

(3.5)

Now, Moran's $I$ is seen as a product of square root of $S$ and Pearson's $r$ between a variable and its SL or SMA. From (3.5), two things should be noted.

First, $S$ reveals substantive information about the spatial clustering of a variable. If a variable is more spatially clustered, its $S$ is larger, because variance of the original vector is less reduced when it is transformed to its SL or SMA. For example, $S$s for three spatial patterns, A, B, and C in Figure 3.2, are respectively 0.649, 0.418, and 0.175. The value of 0.649 for pattern A indicates that the variance of A's SL is approximately 64.5% of that of the original A.

Second, $S$ is a crucial element in the Moran's $I$ equation. The other element, Pearson's $r$ between a variable and its SL, remains a measure of point-to-point association in the sense that very different associations between an area and its neighbors could result in very similar or even identical contributions. For example, if two
Figure 3.2: The relationship between $S$ and Moran's $I$

\[ S_A = 0.649 \]
\[ r_{A,A} = 0.848 \]
\[ I_A = 0.681 \equiv \sqrt{0.649 \cdot 0.848} \]

\[ S_B = 0.418 \]
\[ r_{B,B} = 0.597 \]
\[ I_B = 0.386 \equiv \sqrt{0.418 \cdot 0.597} \]

\[ S_C = 0.175 \]
\[ r_{C,C} = -0.453 \]
\[ I_C = -0.186 \equiv \sqrt{0.175 \cdot (-0.453)} \]
observations have the same value and their neighbors’ means are the same, their spatial
lag elements will be identical; thus their contributions to Pearson’s r between the variable
and its SL are identical. However, a neighbors’ mean does not take variance among
neighbors into account: one observation could be surrounded by homogeneous neighbors;
the other could be connected to neighbors which are very different from one another.

In summary, a spatial smoothing scalar S is a direction-free univariate spatial
association measure like Geary’s c. If a spatial pattern is more spatially clustered, it is
given a higher value of S.

3.3 A local univariate spatial association measure, $S_i$

A local spatial smoothing scalar ($S_i$) is given:

$$S_i = \frac{n^2}{\sum_i \left( \sum_j v_{ij} \right)^2} \cdot \frac{\left( \sum_j v_{ij} (x_j - \bar{x}) \right)^2}{\sum_i (x_i - \bar{x})^2}$$ (3.6)

when a row-standardized spatial weights matrix W is applied, equation (3.6) is simplified
to:

$$S_i = n \cdot \frac{(\bar{x}_i - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}$$ (3.7)
Equation (3.7) can be rewritten as:

$$S_i = \left( \frac{x_i - \bar{x}}{\sigma_x} \right)^2$$  \hspace{1cm} (3.8)

Now, a vector of local $S$s is seen as a vector of square of a normalized form of $\bar{X}$ (subtracted by the mean and divided by the standard deviation of the original variable $X$).

Here, non-squared portion in (3.8) can be conceptualized as a spatially smoothed z-score (SSZ), which can be given in a matrix form:

$$\bar{z}_x = W \cdot z_x$$  \hspace{1cm} (3.9)

where $\bar{z}_x$ is a vector of spatially smoothed z-scores and $z_x$ is a vector of normal z-scores, which is given:

$$z_i = \frac{x_i - \bar{x}}{\sigma_x} = \frac{x_i - \bar{x}}{\sqrt{\sum_{j=1}^{n} (x_j - \bar{x})^2 / n}}$$  \hspace{1cm} (3.10)

A spatially smoothed z-score is nothing but an weighted mean of z-scores of neighbors such that the relationship between an original vector and its SL or SMA is similar to one between a z-score vector and its spatially smoothed version. From (3.9) and (3.10), one may recognize:
where \( s_x \) is a vector of local \( S_i \).

Equation (3.11) indicates that a higher \( S_i \) results when not only a high z-score in an area itself but a high average of z-scores in its neighbors are present. This suggests that an area's value itself should be included in calculation.

Equation (3.9) can be written in a more general form:

\[
\bar{z}_i = \frac{\sum_j \sum_{y} y_j (x_j - \bar{x})}{\sqrt{\sum_i (x_i - \bar{x})^2 / n}}
\]  (3.12)

From (3.12), one may notice that a spatially smoothed z-score is very similar to Getis-Ord's statistics which is given (Ord and Getis 1995:289):

\[
G_i^* = \frac{\sum_j w_{ij} x_j - W_{i}^* \bar{x}}{s \cdot \sqrt{(n S_{ii}^* - W_{i}^{*2}) / (n-1)}}
\]  (3.13)

where \( W_{i}^* = \sum_j w_{ij} \), \( S_{ii}^* = \sum_j w_{ij}^2 \), and \( s \) is the sample standard deviation. When \( W \) matrix is applied to (3.9) or (3.12), a spatially smoothed z-score becomes similar to \( G_i \); when \( W^* \) is considered, it conceptually corresponds to \( G_i^* \).
Behaviors of $S_i$ can become more evident when it is compared to local Moran's $I_i$ that is given:

$$I_x = z_x \circ \overline{z}_x$$

(3.14)

where $I_x$ is a vector of local Moran's $I_i$'s and $\circ$ symbol denotes a pairwise dot product between two vectors. It is certain that a local Moran's $I_i$ is largely determined by a reference area. A negative value is assigned to a local set where a $z$-score in a reference location and a spatially smoothed $z$-score of its neighbors have different signs. It may be problematic in some cases: for example, when a reference area has a value of slightly smaller-than-average and its neighbors have values of highly larger-than-average, the overall locale should be detected as a spatial clustering of high values. In this case, a local Moran's $I_i$ in the reference area will be negative, whereas a local $S_i$ will be high.

In order to clarify all these issues revolving around local SAMs, it is necessary to clearly define such concepts as spatial clusters, hot spots, cold spots, spatial outliers, and local homogeneity. My definition here may be different from others, e.g. Wartenberg and Greenberg (1990).

A spatial cluster is a group of adjacent areas showing a significantly high level of spatial dependence. If spatial dependence is defined as a tendency of similar values being spatially clustered, it is immune to values at location themselves. In other words, clusters of higher values, lower values, or even average values should be detected as displaying spatial dependence. Anselin (1995) focuses on clusters of higher values and...
regards a spatial cluster as synonymous with a hot spot. However, following Sokal et al. (1998), I make a distinction between hot spots as spatial clusters with significantly high values and cold spots as spatial clusters with significantly low values. All the other forms of spatial dependence, e.g. spatial cluster of average values, can be conceptualized as local homogeneity. Hot or cold spots are not necessarily clusters of significantly higher level of local homogeneity, because local variance within spots is often high.

A spatial outlier refers to an area that is significantly different from its neighbors (Wartenberg 1990:140; Fotheringham and Charlton 1994:318; Anselin 1998:83). In other words, spatial outliers are ones with negative spatial autocorrelation in a univariate situation. This concept is extremely useful in postulating sporadic spatial processes that are defined in Chapter 2. The presence of spatial outliers indicates that certain negative spatial interaction, such as backwash effects, spatial repulsion, and spatial competition, occurs in the locales.

Figure 3.3 addresses some of those issues. Nine local spatial patterns are categorized into three groups in terms of reference areas: A for positively-centered; B for neutrally-centered; C for negatively-centered. Each group is associated with three different specifications of neighbors; 1 for positively-satellized; 2 for neutrally-satellized; 3 for negatively-satellized. As mentioned before, columns of b and c behave in a very similar way as Getis-Ord statistics. When values in hexagons are assumed to be z-scores, a product of a value in the center and an average of satellite hexagons (column b) yields a local Moran's $I$ (column d), whereas the square of an average of all the values (column c) is a local $S_i$ (column f).
<table>
<thead>
<tr>
<th>Local Spatial Patterns</th>
<th>$z_i^a$</th>
<th>$\bar{z}_i$</th>
<th>Local Moran's $I_i^d$</th>
<th>Geary's $c_i^e$</th>
<th>Local $S_i^f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A2</td>
<td>1</td>
<td>0</td>
<td>0.14</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A3</td>
<td>1</td>
<td>-1</td>
<td>-0.71</td>
<td>-1</td>
<td>4</td>
</tr>
<tr>
<td>B1</td>
<td>0</td>
<td>1</td>
<td>0.86</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>B2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B3</td>
<td>0</td>
<td>-1</td>
<td>-0.86</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>C1</td>
<td>-1</td>
<td>1</td>
<td>0.71</td>
<td>-1</td>
<td>4</td>
</tr>
<tr>
<td>C2</td>
<td>-1</td>
<td>0</td>
<td>-0.14</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>C3</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

a: see (3.10)  
b: see (3.12) with W  
c: see (3.12) with W*  
d: a·b  
e: $\sum_j (z_i - z_j)^2 / 6$  
f: $c^2$

Figure 3.3: A comparison of local univariate spatial association measures
First of all, local Moran's $I_s$ performs well in detecting spatial clusters and spatial outliers (A3 and C1). However, it does not make a distinction between hot and cold spots as can be seen from a comparison between A1 and C3. This limitation equally applies to local Lee's $S_i$. However, it should be noted that the distinction can be made internally in computing the measures. In contrast, Getis-Ord's statistics represented by columns of $b$ and $c$ in Figure 3.3 performs best in differentiating hot spots and cold spots among spatial clusters. However, it does not detect spatial outliers, which is parallel to local Lee's $S_i$.

Second, in terms of gauging the degree of spatial dependence, local Geary $c_i$ behaves in a completely different way. Maximum spatial dependence denoted by a value of 0 in the column of $e$ is detected in A1, B2, and C3. A striking case is B2, because none of the other measures identifies it as a spatial cluster. Thus, local Geary's $c_i$ can be said to capture local homogeneity.

Third, it appears that local Moran's is strongly influenced by reference areas, whereas Getis-Ord's statistics and local Lee's $S_i$ are relatively independent of them, which can be best seen from local patterns belonging to B-class: whatever characteristics neighbors have, the former is constant at zero; in contrast, the latter measures seem to properly capture the degree of spatial dependence within an overall locale. Local Geary's $c_i$ also seems to be strongly influenced by a reference area. Three local patterns, A1, B1, and C1 may be recognized as a spatial cluster of high values, even though the degree decreases from pattern A to pattern C. Local Geary's $c_i$ identifies them as totally different: a positive spatial cluster for A1; no spatial autocorrelation for B1; a negative spatial cluster for C1.
Forth, a comparison between b and c columns in Figure 3.3 may conclude that local $S_i$ and Getis-Ord statistic perform better with $C^*$ or $W^*$ rather than with $C$ or $W$ in terms of pattern detection, which is echoed by Getis and Ord (1996:273). When A3, B3, and C3 are compared, one may notice that the column of c captures different levels of spatial dependence more effectively than the column of b.

Table 3.2 summarizes each measure' performance for each task. Local Moran’s $I_i$ is the only measure that can detect spatial outliers, and is also effective in identifying spatial clusters. Local Geary’s $c_i$ performs best for gauging the level of local homogeneity, but is not effective for other tasks. Getis-Ord statistics performs best in differentiating hot spots and cold spots and is suitable for capturing the overall level of spatial clustering in a local set, not heavily depending on a reference area. Local Lee’s $S_i$ behaves similarly to Getis-Ord statistics, but requires extra efforts in distinguishing hot spots from cold spots.
<table>
<thead>
<tr>
<th></th>
<th>Local Moran's $I_i$</th>
<th>Local Geary's $c_i$</th>
<th>Getis-Ord's Statistics</th>
<th>Local Lee's $S_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial clusters</td>
<td>Effective</td>
<td>Effective</td>
<td></td>
<td>Not Effective</td>
</tr>
<tr>
<td>Hot spots vs. Cold spots</td>
<td>Neutral</td>
<td>Not Effective</td>
<td>Effective</td>
<td>Neutral</td>
</tr>
<tr>
<td>Local homogeneity</td>
<td>Not Effective</td>
<td>Effective</td>
<td>Not Effective</td>
<td>Effective</td>
</tr>
<tr>
<td>Spatial outliers</td>
<td>Effective</td>
<td>Not Effective</td>
<td>Effective</td>
<td>Not Effective</td>
</tr>
<tr>
<td>Dependency on reference area</td>
<td>Strong</td>
<td>Strong</td>
<td></td>
<td>Neutral</td>
</tr>
</tbody>
</table>

Table 3.2: Characteristics of local univariate spatial association measures
CHAPTER 4

BIVARIATE SPATIAL ASSOCIATION MEASURES, $L$ AND $L_1$

4.1 Parameterization of bivariate spatial dependence

The concept of spatial dependence points to the propensity for nearby locations to influence each other and to possess similar attributes (Anselin 1988; Anselin and Griffith 1988; Anselin and Getis 1992). The univariate spatial dependence, or spatial autocorrelation, has well been conceptualized and extended to deal with spatially autocorrelated errors. However, spatial dependence in bivariate situations has attracted little attention. The bivariate spatial dependence is defined here as spatial clustering of similar bivariate associations across two variables. Even though negative spatial autocorrelation that a bivariate association at a location is very different from ones at its neighbors is often as important as positive autocorrelation in terms of bivariate associations, more focus in this dissertation is placed on positive bivariate spatial association. In the presence of the bivariate spatial dependence, the significance testing for Pearson's correlation coefficient may be flawed because the degree of freedom cannot
be calibrated by $n - 2$ (Bivand 1980; Richardson and Hémon 1981; Clifford and Richardson 1985; Clifford et al. 1989; Haining 1991; Dutilleul 1993).

From two geographical variables, $n!$ different pairs can be drawn, when elements in each variable are all different (note that the corresponding data points are bound in a permutation process). The $n!$ different pairs are identical to one another in terms of the point-to-point association, e.g. Pearson's $r$. Since data points are spatially indexed, however, different pairs are characterized by different degrees of the bivariate spatial dependence, thus different levels of spatial co-patterning are revealed. To illustrate, three patterns in Figure 4.1 are now seen as different variables, and the three pairs, A-B, B-C, and C-A, show identical relationships in terms of Pearson's $r$ (0.422). The association of A-B, however, shows a higher level of bivariate spatial dependence or spatial co-patterning than those of B-C and C-A: the association of A and B displays the highest level of spatial clustering of hexagons sharing the same values between the two maps.

Having realized that a pair of variables under investigation represents only a particular case of all possible bivariate spatial associations, one may wish to devise a measure that effectively differentiates the associations by integrating the two concepts of 'association'. The importance of a conceptual disintegration and computational reintegration of 'point-to-point association' and 'spatial association' can be clarified by two conceptual illustrations. In the first example, a bivariate spatial association is seen as a Pearson’s $r$ between two sets of local Moran’s $I$s. This illustrates that locations with an
Figure 4.1: Three bivariate spatial associations among three spatial patterns
identical value might be differently recognized in measuring a bivariate association if their relations with neighboring locations are different.

The second conceptual illustration is provided by the global Moran's $I$ of what can be termed local Pearson's $r_i$. A local Pearson's $r_i$ captures the degree of numerical correspondence between two values at a location, and is simply calculated by multiplying two z-scores of the values, each of which is standardized by the mean and standard deviation of each variable. The mean value of local Pearson's $r_i$s is nothing but a *global* Pearson's $r$. When local Pearson's $r_i$s are mapped, a global Moran's $I$ captures the degree of spatial dependence of point-to-point associations across locations. This provides a conceptual foundation for a bivariate spatial association measure and suggests that an integration of Moran's $I$ as a univariate *spatial association* measure and Pearson's $r$ as an aspatial *bivariate association* measure may lead to a feasible measure. As can be seen from Figure 4.1, the level of bivariate spatial dependence is determined by the level of univariate spatial dependence of variables involved when the point-to-point association is held constant. This further suggests that a bivariate spatial association measure should be a composite of three elements: *univariate spatial associations of two variables and their point-to-point association in a certain form*. Figure 4.2 illustrates all these conceptual considerations.

Although the need for a bivariate spatial association measure has long been recognized, the only comprehensive attempt to devise a parametric bivariate spatial association measure is Wartenberg’s work (1985), which proposed a matrix algebraic form for the *bivariate* Moran’s $I$ intended to provide an alternative correlation matrix for
Figure 4.2: Conceptualization of bivariate spatial association

Modified from Tiefelsdorf (2001)
a spatial principal components analysis. His measure has drawbacks, however, which will be discussed subsequently.

4.2 A global bivariate spatial association measure, $L$

4.2.1 Criteria for a bivariate spatial association measure and critiques on Wartenberg's formulation

By reference to findings in the previous section, two criteria can be suggested for developing a bivariate spatial association measure. First, the measure should conform to Pearson's $r$ between two variables in terms of direction and magnitude to a certain extent. Although the measure has an exclusive interest in the spatial association among observations, it should retain the direction and magnitude of a point-to-point association between two variables, which requires the inclusion of a certain form of Pearson's correlation between two variables. Second, a bivariate spatial association measure should reflect the degrees of spatial autocorrelation for both variables under investigation. In other words, it should respond to the collective effect of $S$s of the variables.

The most comprehensive attempt to develop a bivariate spatial association measure by extending Moran's $I$ is Wartenberg's work (1985). He developed a bivariate Moran's $I$ following Mantel's formulations.

$$ I = \frac{Z^T CZ}{I^T CI} $$ (4.1)
where I is a variable-by-variable Moran correlation matrix, Z is a case-by-variable matrix whose elements are z-scored, C is a case-by-case binary connectivity matrix, and 1 is a case-by-1 column matrix with all elements being 1s. The diagonal values of I are Moran's I coefficients for the variables, with each off-diagonal element being a bivariate Moran's I (Griffith (1993; 1995) terms it Cross-MC (Moran Coefficient)), which is similar to the cross-correlation approach in geostatistics (Isaaks and Srivastava 1989).

By decomposing the matrix and applying a row-standardized spatial weights matrix, one can write an equation for an off-diagonal element in the matrix I between two variables, X and Y:

\[
I_{X,Y} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}} \equiv \sqrt{S_Y} \cdot r_{x,y}
\]

(4.2)

A comparison of (4.2) and (3.5) reveals that Wartenberg's bivariate spatial association measure, or bivariate Moran's I, captures a bivariate association between X and the SL of Y, and the association is scaled by square root of S for Y.

Using Wartenberg's formula as a bivariate spatial association measure has two obvious disadvantages that violates the two criteria established before. First, it is conceptually untenable to allow a bivariate spatial association measure to be primarily calibrated by the relationship between a variable and the other variable's SL. Moreover, a bivariate spatial association measure should incorporate both Ss of two variables in the
equation, not just the $S$ of a variable. Equation (4.2) also implies that $I_{x,y}$ and $I_{y,x}$ may be different when a row-standardized spatial weights matrix is involved, which nullifies much of Wartenberg's attempt to formulate a spatial principal component analysis using an $I$ matrix in (4.1).

Second, Wartenberg's equation is vulnerable to a reverse of the direction of association. For example, when an area $i$ with a higher-than-average value for both $X$ and $Y$ are surrounded by lower-than-average values, the numerator value in (4.2) could be given a negative value, because the SL of $Y$ for the area is negative (the right part of the numerator), with the left part being necessarily positive. A simulation observed that most of the associations with negatively autocorrelated $Y$ vectors were assigned negative bivariate association indices. In conclusion, Cross-MC should not be used as a bivariate spatial association measure.

4.2.2 A global bivariate spatial association measure, $L$

A bivariate spatial association measure ($L$) is defined as:

$$L_{x,y} = \frac{n}{\sum_i \left( \sum_j y_j \right)^2} \frac{\sum_i \left( \sum_j y_j (x_j - \bar{x}) \right) \left( \sum_j y_j (y_j - \bar{y}) \right)}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}$$

(4.3)

and, when a row-standardized spatial weights matrix ($W$) is applied, (4.3) is simplified to:

[Continues]
Further, when the SL operation is introduced, (4.4) is transformed to:

\[
L_{x,y} = \frac{\sum_i \left[ \left( \sum_j w_j (x_j - \bar{x}) \right) \cdot \left( \sum_j w_j (y_j - \bar{y}) \right) \right]}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}
\]  

(4.5)

Note that \(x, y\) are elements for a location \(i\) in \(X\)'s and \(Y\)'s SL vectors (\(\bar{X}\) and \(\bar{Y}\)).

To decompose (4.5) as undertaken for the Moran's \(I\) equation, it is compared to an equation for Pearson's \(r\) between SLs, which is given as:

\[
r_{\bar{x}, \bar{y}} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}
\]  

(4.6)

Note that \(\bar{x}\) and \(\bar{y}\) are mean values of the SL vectors. By utilizing (4.6), (4.5) can be rewritten as:
The element of A is approximately 1, and the element of B will be zero when either variable’s mean is identical to one of its SL, which is very likely. Then $L$ is redefined as:

$$L_{x,y} = \sqrt{S_x} \cdot \sqrt{S_y} \cdot r_{x,y}$$  \hspace{1cm} (4.8)$$

Now, $L$ between two variables is calculated by multiplying Pearson’s correlation coefficient between their SL vectors by the square root of the product of their $S$s.

Further, a matrix algebraic form for $L$ is provided, when variables are z-transformed:

$$L = \frac{Z^T (V^T V) Z}{1^T (V^T V) 1}$$  \hspace{1cm} (4.9)$$

where $L$ is a variable-by-variable bivariate spatial association matrix, $Z$ is an area-by-variable (z-scored) data matrix, and $V$ is an area-by-area general spatial weight matrix.

Note that, when $W$ is applied, the denominator is reduced to $n$. A spatial correlation
matrix driven by (4.9) can be furthered to calibrate a spatial principal components analysis as seen from Wartenberg's attempt (1985).

In addition, it should be noted that the diagonal elements in matrix L have a particular meaning. From equation (14), a diagonal element can be written as:

\[ L_{x,x} = \frac{\sum_i (x_i - \bar{x})^2}{\sum_i (x_i - \bar{x})^2} \]  

(4.10)

where \( L_{x,x} \) is simply the S of X defined in Chapter 3. Equation (4.10) allows a transformation of (4.8):

\[ L_{x,y} = \sqrt{L_{x,x}} \cdot \sqrt{L_{y,y}} \cdot r_{x,y} \]  

(4.11)

A higher value in the diagonal of the matrix L implies a higher Moran's I for the variable, and results in a higher L index between the variable and other variables, all other conditions being constant.

In summary, the L index as a bivariate spatial association measure is largely determined by Pearson's r between two SL vectors, which generates a smoothed version of Pearson's correlation coefficient between the original variables. Pearson's r between SLs, then, is scaled by a product of univariate Ss of the variables, which suggests that L
captures not only the bivariate 'point-to-point association' between two variables, but also the univariate spatial autocorrelation.

4.2.3 An illustration with a hypothetical data set

For the purpose of illustration, the three different spatial patterns, A, B, and C in Figure 4.1 are utilized (Figure 4.3). A', B', and C' are spatially rotated versions of those patterns, such that the univariate spatial dependence of the original patterns remain unchanged in terms of $S$ and Moran's $I$. From Table 1, four things should be acknowledged.

First, the sign in Pearson's $r$ between two variables remains unchanged in $L$ as long as the sign of Pearson's $r$ between their SLs is given accordingly. The only exception is found in the association of A-C', where Pearson's $r$ between the two patterns is positive (0.107), but one between their SLs is negative (-0.240). One way of dealing with this problem may be to apply the spatial moving average operation where the weighted mean of neighbors for an area is computed with the area itself being included. This means that the spatial weights matrix $W$ as a row-standardized version of $C$ is replaced by a matrix of a row-standardized version of a modified $C$, where $c_{ij} = 1$.

Second, as seen in equation (19), $L$ between two identical patterns does not yield a value of 1, and the value changes between pairs of variables (compare A-A, B-B, and C-C in Figure 4.2). This provides a crucial insight into the comparison between two spatial patterns. That is, the bivariate spatial dependence between identical patterns is completely determined by the univariate spatial dependence of the pattern.
<table>
<thead>
<tr>
<th>Association</th>
<th>Pattern</th>
<th>SSS</th>
<th>Correlation</th>
<th>$L_{X,Y}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$X$</td>
<td>$Y$</td>
<td>$X^a$</td>
</tr>
<tr>
<td>A-A</td>
<td>![Pattern]</td>
<td>0.649</td>
<td>0.649</td>
<td>1.000</td>
</tr>
<tr>
<td>B-B</td>
<td>![Pattern]</td>
<td>0.418</td>
<td>0.418</td>
<td>1.000</td>
</tr>
<tr>
<td>C-C</td>
<td>![Pattern]</td>
<td>0.175</td>
<td>0.175</td>
<td>1.000</td>
</tr>
<tr>
<td>A-B</td>
<td>![Pattern]</td>
<td>0.649</td>
<td>0.418</td>
<td>0.628</td>
</tr>
<tr>
<td>B-C</td>
<td>![Pattern]</td>
<td>0.418</td>
<td>0.175</td>
<td>0.577</td>
</tr>
<tr>
<td>C-A</td>
<td>![Pattern]</td>
<td>0.175</td>
<td>0.649</td>
<td>0.634</td>
</tr>
<tr>
<td>A-A'</td>
<td>![Pattern]</td>
<td>0.649</td>
<td>0.649</td>
<td>-0.800</td>
</tr>
<tr>
<td>B-B'</td>
<td>![Pattern]</td>
<td>0.418</td>
<td>0.418</td>
<td>-0.388</td>
</tr>
<tr>
<td>C-C'</td>
<td>![Pattern]</td>
<td>0.175</td>
<td>0.175</td>
<td>-0.185</td>
</tr>
<tr>
<td>A-C'</td>
<td>![Pattern]</td>
<td>0.649</td>
<td>0.175</td>
<td>-0.240</td>
</tr>
</tbody>
</table>

\[ e = \sqrt{a \cdot b \cdot c}. \]

Figure 4.3: $L$ with different bivariate spatial associations
Third, $L$ differentiates different spatial associations with an identical Pearson correlation coefficient. A-B, B-C, and C-A are identical in terms of Pearson's $r$ (0.422); however, they have $L$s respectively of 0.327, 0.154, and 0.214 (Figure 4.2). This implies that $L$ is largely determined by the SSSs of the two variables involved when Pearson's $r$ is identical. Since a negative $L$ indicates a spatial discrepancy, a poorer spatial co-patterning should be given a negative value with a larger amount. This is well illustrated by a comparison between B-B' and C-C': Pearson's correlation coefficients are identical (-0.051), but the spatial discrepancy is much more obvious in B-B', which is reflected in $L$ values (-0.162 and -0.024).

Fourth, $L$ differentiates different spatial associations with identical Ss but different Pearson's $r$, which can easily be acknowledged by comparing A-A and A-A', B-B and B-B', and C-C and C-C' in Figure 4.2.

Comparing $L$ values among different spatial patterns, one may recognize that the $L$ effectively measures similarity/dissimilarity among variables in terms of bivariate associations and their spatial clustering. In computation, the numerical point-to-point association is calibrated largely by Pearson's $r$ between SL vectors, and the spatial association is recognized by the product of Ss. Thus, the two elements collectively capture the spatial co-patterning and parameterize the bivariate spatial dependence.
Albeit its obvious rationale, the concept of the local bivariate spatial association in the context of lattice data has never been extensively tackled, with few exceptions. A concept of a geographically weighted correlation or spatial moving correlation has been proposed (Brunsdon et al. 1999, Fotheringham and Brunsdon 1999, Fotheringham 2000: 75), which is very similar to what has been used to calibrate the level of correspondence between two raster or image layers. As can be seen from the previous section, a local Cross-Moran's $I_i$ can be drawn from the corresponding global measure (Wartenberg 1985). However, mainly due to its unreliable behaviours (see Lee 2001b, Tiefelsdorf 2001), it should be ruled out as a local association measure. Tiefelsdorf (2001) proposes a spatial cross-correlation coefficient in the context of the comparison between two sets of regression residuals. A local $L_i$ largely corresponds to his specification; practically, a spatial cross-correlation coefficient is a local $L_i$ without the denominator under a restriction that a spatial weights matrix should have a zero-diagonal.

4.3.1 Defining a local bivariate spatial association measure, $L_i$

By decomposing (4.3), a bivariate local spatial association measure is presented as (Lee 2001b):
When a row-standardized spatial weights matrix ($W$) applies, equation (4.12) is simplified to:

$$L_i = \frac{n^2}{\sum (\sum v_j)^2} \cdot \left( \frac{\sum v_j (x_j - \bar{x})}{\sqrt{\sum (x_i - \bar{x})^2}} \cdot \frac{\sum v_j (y_j - \bar{y})}{\sqrt{\sum (y_i - \bar{y})^2}} \right)$$  (4.13)

where $\bar{x}_i$ and $\bar{y}_i$ denote *spatially smoothed values* at $i$th location that can be calculated by $\sum_j w_{ij} x_j$ and $\sum_j w_{ij} y_j$ (see Lee 2001b). Now, a vector of local $L_i$s is seen as a vector resulting from a product of normalized forms of $\bar{x}$ and $\bar{y}$ (subtracted by means and divided by standard deviations of the original vectors), which can be written as:

$$l_{x,y} = \bar{z}_x \circ \bar{z}_y$$  (4.14)

where $l_{x,y}$ is a column vector of local Lee’s $L_i$s. $\bar{z}_x$ and $\bar{z}_y$ can be conceptualized as vectors of *spatially smoothed z-scores* defined in (3.9) and (3.12).
Similarly, a bivariate local pseudo- or quasi-spatial association measure, local Pearson's $r_n$, can be given by decomposing the Pearson's correlation coefficient (Lee 2001b):

$$r_i = n \cdot \frac{(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} = \frac{(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \cdot \sigma_y}$$

(4.15)

Now, a vector of local $r_s$ is seen simply as a vector consisting of products between $z$-transformed $X$ and $Y$, which is given:

$$r_{x,y} = z_x \circ z_y$$

(4.16)

By using the same notions, vectors of local Moran's $I_s$ and local Cross-Moran $I_s$ can be denoted respectively by $z_x \circ \bar{z}_x$ and $z_x \circ \bar{z}_y$. From equation (4.16), one may notice that a local $r_i$ gauges a local correlation between two values at $i$th location; in other words, the pair's relative contribution to the corresponding global correlation.

However, a local Pearson's $r_i$ is not a truly spatial statistic even though it is evidently local, because it does not have any information about the topological relationships among observations in a data set. As far as geographically referenced data are concerned, local $r_i$ only captures a point-to-point association or numerical co-varying, ignoring spatial association or spatial clustering (Hubert et al. 1985, Haining 1990, 1991, Lee 2001b, Tiefelsdorf 2001). Under the presence of the bivariate spatial
dependence (Lee 2001b), neighboring locations tend to retain similar covariances and thus a set of local \( r \)s should display a spatial clustering when it is mapped. In this sense, local \( L \)s can be seen as spatially smoothed local correlations. A value for a location is given not only in terms of the value at the location, but in terms of values in neighboring locations.

### 4.3.2 Behaviors of \( L_i \) in different local settings

In this section, it will be illustrated how the measure behaves in order to parameterize the bivariate spatial dependence under different local spatial configurations consisting of a reference location and its neighbors. This is conducted not only in comparison with local Pearson’s \( r_h \) but also in terms of the difference between spatial lag and spatial moving average operations. The transformation from (4.12) to (4.13) was made possible by defining spatially smoothed vectors, i.e. \( \tilde{x} \) and \( \tilde{y} \). Those vectors are spatial lag vectors when diagonal elements in a spatial weights matrix are zero; but they are spatial moving average vectors when one with a non-zero diagonal applies. In order to examine how those different operations impact on the behaviours of local \( L_i \), two different spatial weights matrices are utilized: \( W \) is a row-standardized version of \( C \) where a cell is assigned 1 if two areas are directly connected and zero otherwise; \( W^* \) is a row-standardized version of \( C^* \), a \( C \) matrix with 1s on its diagonal.

Figure 4.4 shows how \( L_i \) behaves in different local bivariate associations. A hypothetical local spatial configuration composed of seven hexagons with a central one being a reference location is introduced. Three different z-values, 1, 0, and -1, are
The approach presented here is adopted from Tiefelsdorf (2001)

Figure 4.4: Local $L_s$ in different local settings
assigned to reference locations such that we have three different groups of local patterns (positively-centered, neutrally-centered, and negatively-centered): e.g. local patterns starting with A (A1 to A6) have a z-value of 1 in the reference location. For each group, six different combinations of 1, 0, and -1 constitute sets of neighbors: e.g. local patterns ending with 1 (A1, B1, and C1) have the same set of neighbors consisting of six 1s. By cross-tabulating the 18 different local patterns, 6 large cross-sections (A-A, A-B, A-C, B-B, B-C, and C-C), each of which is composed of 36 cells, are generated (note that Figure 4.4 is symmetric). Since 3 diagonal sections are internally symmetric, 171 different local bivariate spatial associations are yielded. The figure is doubled because each cell has two entries respectively based on two spatial weights matrices (W and W'). Since values assigned to hexagons are assumed to be z-scores, a product of two spatially smoothed z-scores yield a local $L_i$ for the reference location.

This approach is adopted from Tiefelsdorf (2001). Major focus here, however, is somewhat different from his perspective. First of all, local Lee’s $L_i$ does not aim to assess spatial co-patterning between two vectors of regression residuals. Rather, it focuses more on what Tiefelsdorf (2001) terms plain observation, which are deviations from their global mean, ignoring the implications of projection matrix. Second, it does not take conditional probabilities of local measures into account, which means that it does not control univariate spatial dependence of each variable. Third, while his rules to assess spatial co-patterning revolve around an analytical distinction between a reference area and its neighbors, the evaluation scheme utilized here is overall correspondence.
between two variables at a local set, which implies a reference area and its neighbors are equally weighted in assess local spatial co-patterning.

From Figure 4.4, several observations are made.

First, each cross-section is represented by a single local Pearson's $r_l$. Since a local Pearson's $r_l$ is computed by two z-scores at a reference location, it is indifferent to neighbor settings. Thus, the sections of A-A, A-B, A-C, B-B, B-C, and C-C are given respectively 1, 0, -1, 0, 0, and 1.

Second, unlike local Pearson's $r_l$, local Lee's $L_l$ takes account of neighboring values so that it has a negative value in a section where a positive local Pearson's $r_l$ is expected. For example, the association of A1-A6 has a pair of negative $L_l$s in its cell with a local Pearson's $r_l$ of 1.

Third, there are some variations in diagonal cells that are denoted by associations of identical patterns. As Lee (2001b) points out, the measure differentiates among pairs of identical patterns in terms of variances among neighbors such that only two identical associations (A1-A1 and C3-C3) yield a maximum local $L_l$ (1 in our case) with both spatial weights matrices.

Fourth, no association results when a local pattern of no local autocorrelation (B2) is related to any other local patterns, which conforms to a basic rule in assessing spatial co-patterning (Tiefelsdorf 2001). All the cells associated with B2 are given a value of 0 in Figure 4.4.

Fifth, $W'$ yield a more reasonable set of local measures than $W$. Since $W$ does not take into account a reference location, upper values computed with $W$ in Figure 4.4
are identical among all the cross-sections. For example, Al-A1, A1-B1, and A1-C1 are all given a Local $L_i$ of 1 when $W$ is applied, even though the degrees of spatial co-patterning are obviously different: highest in A1-A1 and lowest in A1-C1. This echoes Tiefelsdorf’s (2001) observation that ignorance of reference cells leads to inability to distinguish spatial clusters (A1) from what he terms hot spots (C1). In this sense, the measure behaves more reasonably with $W^*$: it effectively differentiates those associations by assigning different values, 1, 0.86, and 0.71. In conclusion, unlike some local univariate SAMs such as Moran’s $I_i$ and Geary’s $c_i$, a spatial weights matrix with a non-zero diagonal should be utilized for local Lee’s $L_i$, which also applies to Lee’s $S_i$ and Getis-Ord’s $G^*_i$.

To compare a reference area with its neighbors in terms of bivariate association may be another research topic. On the one hand, bivariate association at a reference area is computed by aspatial local Pearson’s $r_i$. On the other hand, a local Lee’s $L_i$ with $C$ or $W$ provides a local spatial co-patterning around the reference area. A comparison of the two values may allow not only for exploring spatial clusters but also for identifying bivariate negative spatial autocorrelation or bivariate spatial outliers, reference areas that are significantly different from their neighbors in terms of bivariate associations. It is beyond the scope of this dissertation to integrate a task of assessing an overall level of spatial co-patterning at a local set and another task of examining the relationship between a reference area and its neighbors in terms of bivariate associations.
CHAPTER 5

A GENERALIZED SIGNIFICANCE TESTING METHOD I:
A NORMALITY ASSUMPTION

It has been well acknowledged that the use of Moran's I as a global spatial association measure to parameterize the spatial clustering in a geographical pattern is a special case of its more general use for assessing spatial autocorrelation among regression residuals with an assumption that unobservable disturbances are independent identically normal distributed (Cliff and Ord 1981; Upton and Fingleton 1985; Anselin 1988; Tiefelsdorf and Boots 1995). Distributional properties of the measure including higher moments under the assumption of spatial independence have been examined (Durban and Watson 1950; 1951; 1971; Henshaw 1966; 1968; Hepple 1998; Tiefelsdorf 2000). Further, an exact distribution approach has demonstrated its superiority over the approximation approach (Tiefelsdorf and Boots 1995; Hepple 1998) and its ability to embrace the conditional moments (Tiefelsdorf 1998; 2000).

This chapter is concerned with formulating a general procedure to generate first four moments of spatial association measure. This is based on a rationale that the
moment extraction procedure developed for Moran’s $I$ can be extended not only to other univariate spatial association measures as suggested for Geary’s $c$ (Cliff and Ord 1981; Hepple 1998), but also to local measures as applied to local Moran’s $I_l$ (Boots and Tiefelsdorf 2000), as far as a measure can be defined as scale invariant ratio of quadratic forms of residuals (Tiefelsdorf 2000). Since spatial association measures are seen as ratio of quadratic forms of deviants from an overall mean, resulting distributional moments correspond to those extracted under the normality assumption (Cliff and Ord 1981:21).

Subsequently, I first provide a generalized procedure to generate first four moments of spatial association measures. Second, I apply the generalized procedure to six different univariate spatial association measures such as global Moran’s $I$, local Moran’s $I_l$, global Geary’s $c$, local Geary’s $c_l$, global Lee’s $S$ and local Lee’s $S_l$. It will be demonstrated that all the measures are expressed as ratio of quadratic forms of deviants from an overall mean, and that only difference occurs in defining spatial proximity matrices.

### 5.1 A generalized procedure for univariate spatial association measures

A global spatial association measure should be defined as ratio of quadratic forms:

$$\Gamma = \frac{\delta^T \cdot T \cdot \delta}{\delta^T \cdot \delta} \quad (5.1)$$
where \( \delta \) is a vector of deviants of a variable denoted by a vector of \( y \), i.e., each value subtracted by mean, and \( T \) is a global spatial proximity matrix, a normalized form of a spatial weights matrix \( V \).

Equation (5.1) can be rewritten by utilizing a particular projection matrix that is defined as:

\[
M_{(n)} = I - \frac{1}{n} \cdot 1 \cdot 1^T = \begin{pmatrix}
1 - \frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\
-\frac{1}{n} & 1 - \frac{1}{n} & -\frac{1}{n} & \cdots \\
\vdots & \ddots & \ddots & \ddots \\
-\frac{1}{n} & -\frac{1}{n} & \cdots & 1 - \frac{1}{n}
\end{pmatrix}
\] (5.2)

This is a particular form of the projection matrix that projects a dependent variable and disturbances into a residual space that is orthogonal to a design matrix \( X \) consisting of independent variables (Tiefelsdorf 2000:16). That is,

\[
M = I - X(X^TX)^{-1}X^T
\] (5.3)

Since we focus on spatial association measures as pattern describers, the design matrix is solely composed of a vector of 1s resulting in (5.2). By utilizing (5.2), equation (5.1) is transformed to:
where $\delta = M_0 \cdot y$ and $M_0$ is an idempotent, symmetric matrix so that

$$M_0^2 = M_0 \cdot M_0.$$

Previous studies (Durbin and Watson 1950; 1951; 1971; Henshaw 1966; 1968) show that the distributional properties of a spatial association measure defined as in (5.4) are given by a matrix trace operation of $M_0 \cdot T \cdot M_0$ under an assumption of independence among observations. Since a trace operation of matrix products is indifferent to the order of the product, computationally $M_0 \cdot T \cdot M_0$ is reduced to $M_0 \cdot T$. When $M_0 \cdot T$ is denoted by $K$, first four central moments are given (Henshaw 1966; 1968; Hepple 1998; Tiefelsdorf 2000):

\[
\begin{align*}
\mu_1 &= E(\Gamma) = \frac{\text{tr}(K)}{n-1} \\
\mu_2 &= \text{Var}(\Gamma) = 2 \cdot \frac{(n-1) \cdot \text{tr}(K^2) - \text{tr}(K)^2}{(n-1)^2 \cdot (n+1)} \\
\mu_3 &= 8 \cdot \frac{(n-1)^2 \cdot \text{tr}(K^3) - 3 \cdot (n-1) \cdot \text{tr}(K) \cdot \text{tr}(K^2) + 2 \cdot \text{tr}(K)^3}{(n-1)^3 \cdot (n+1) \cdot (n+3)}
\end{align*}
\]
\[
\mu_4 = \frac{12}{(n-1)^4 \cdot (n+1) \cdot (n+3) \cdot (n+5)} \cdot \left[ (n-1)^4 \cdot \left( 4 \cdot \text{tr}(K^4) + \text{tr}(K^2)^2 \right) \right. \\
- 2 \cdot (n-1)^2 \cdot \left[ 8 \cdot \text{tr}(K) \cdot \text{tr}(K^3) + \text{tr}(K^2) \cdot \text{tr}(K) \cdot \text{tr}(K^2) \right] \\
+ (n-1) \cdot \left[ 24 \cdot \text{tr}(K^2) \cdot \text{tr}(K^2) + \text{tr}(K)^4 \right] \\
- 12 \cdot \text{tr}(K)^4 \right] 
\] (5.5d)

In order to use equations above, T matrix (thus V matrix) should be symmetric.

Skewness and kurtosis from the moments are given respectively by (Tiefelsdorf 2000:102):

\[
\beta_i = \frac{\mu_i}{(\mu_2)^{\frac{i}{2}}} \quad (5.6a)
\]

\[
\beta_2 = \frac{\mu_4}{(\mu_2)^2} \quad (5.6b)
\]

This procedure also holds for local spatial association measures as long as they are defined a scale invariant ratio of quadratic forms as demonstrated for local Moran’s \( I_i \) (Tiefelsdorf 1998; Boots and Tiefelsdorf 2000).

\[
\Gamma_i = \frac{y^T \cdot M_{(0)} \cdot T^{(i)} \cdot M_{(0)} \cdot y}{y^T \cdot M_{(0)} \cdot y} \quad (5.7)
\]

\( T^{(i)} \) is a particular form of a local spatial proximity matrix derived from a global proximity matrix. Again, the matrix should be symmetric. As will be seen in the next
section, spatial association measures are differentiated solely by the spatial proximity matrix.

5.2 Distributional moments of global univariate spatial association measures

Table 5.1 summarizes the definition of a global spatial proximity matrix $T$ for three global univariate spatial association measures, Moran’s $I$, Geary’s $c$, and Lee’s $S$. Practically, a matrix of $T$ is defined as a standardized form of a spatial weights matrix $V$. For example, when a row-standardized spatial weights matrix $W$ is applied, $T$ for Moran’s $I$ becomes identical to $V$. Since $T$ (thus $V$) should be symmetric in order to use equations (5-1)-(5-4), it may be necessary for some non-symmetric spatial weights matrices such as $W$ to be transformed according to an equation:

$$V^s = \frac{1}{2} (V + V^r)$$  \hspace{1cm} (5.8)

The $\Omega$ matrix for Geary’s $c$ should be further elaborated. According to Cliff and Ord (1981:167), it is defined:

$$\Omega = \frac{1}{2} \sum_j (v_j + v_j) = \sum_j v_j^s$$  \hspace{1cm} (5.9)
<table>
<thead>
<tr>
<th>( \Gamma )</th>
<th>Equations</th>
<th>( T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moran’s ( I )</td>
<td>( \frac{n}{I^T V 1} \cdot \frac{\delta^T \cdot V \cdot \delta}{\delta^T \cdot \delta} )</td>
<td>( n \cdot \frac{V}{I^T V 1} )</td>
</tr>
<tr>
<td>Geary’s ( c )</td>
<td>( \frac{n-1}{I^T V 1} \cdot \frac{\delta^T \cdot (\Omega - V) \cdot \delta}{\delta^T \cdot \delta} )</td>
<td>( (n-1) \cdot \frac{(\Omega - V)}{I^T V 1} )</td>
</tr>
<tr>
<td>Lee’s ( S )</td>
<td>( \frac{n}{I^T (V^r V) I} \cdot \frac{\delta^T \cdot (V^r V) \cdot \delta}{\delta^T \cdot \delta} )</td>
<td>( n \cdot \frac{(V^r V)}{I^T (V^r V) I} )</td>
</tr>
</tbody>
</table>

Table 5.1: Definitions of global spatial proximity matrix \( T \) for global univariate spatial association measures
Simply the matrix is a diagonal matrix with row-sums when a spatial weight matrix is transformed symmetric.

As mentioned in Chapter 2, $V^TV$ specification for Lee’s $S$ in Table 5.1 introduces an issue of higher-order spatial lag operation (Anselin and Smirnov 1996). In other words, a resulting value of $S$ contains certain information on a second-order spatial autocorrelation. This requires a further investigation, even though the original equation (3.1) does not seem to have any implications for higher-order spatial autocorrelation.

5.3 Distributional moments of local univariate spatial association measures

Table 5.2 summarizes the definition of a local spatial proximity matrix $T^{(l)}$ for three local univariate spatial association measures. The matrix notations for measures in Table 5.2 satisfy the additivity requirement that an average of local SAMs equals to the corresponding global SAM. A local spatial weights matrix $V_i$ is defined as a global spatial weights matrix whose elements are set to zeroes except for entries in $i$th row:

$$V_i = \begin{bmatrix} 0 \\ \vdots \\ v_{il} & \cdots & v_{il} & \cdots & v_{im} \\ 0 \\ \vdots \end{bmatrix}$$  (5.10)
<table>
<thead>
<tr>
<th>$\Gamma$</th>
<th>Matrix Notations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moran's $I_t$</td>
<td>$\frac{n^2}{1^T V 1} \cdot \frac{\delta^T \cdot V_t \cdot \delta}{\delta^T \cdot \delta}$</td>
</tr>
<tr>
<td></td>
<td>$T^{(0)}$ $\frac{n^2 \cdot V_t}{1^T V 1}$</td>
</tr>
<tr>
<td>Geary's $c_t$</td>
<td>$\frac{n(n-1)}{2 \cdot 1^T V 1} \cdot \frac{\delta^T \cdot {\Omega_t - (V_t^t + V_t^r) + \text{diag}(V_t^t)} \cdot \delta}{\delta^T \cdot \delta}$</td>
</tr>
<tr>
<td></td>
<td>$T^{(0)}$ $\frac{n(n-1)}{2} \cdot \frac{1^T V 1}{1^T V 1}$</td>
</tr>
<tr>
<td>Lee's $S_t$</td>
<td>$\frac{n^2}{1^T {V_t^r V_t^t}} \cdot \frac{\delta^T \cdot {V_t^r V_t^t} \cdot \delta}{\delta^T \cdot \delta}$</td>
</tr>
<tr>
<td></td>
<td>$T^{(0)}$ $\frac{n^2 \cdot {V_t^r V_t^t}}{1^T {V_t^r V_t^t} 1}$</td>
</tr>
</tbody>
</table>

Table 5.2: Definitions of local spatial proximity matrix $T^{(0)}$ for local univariate spatial association measures
Even though $V_l$ does not have to be symmetric in calculating a local measure, it should be transformed symmetric in order to calculate distributional moments. When $M_{(l)} \cdot T^{(l)}$ is denoted by $K$, the equations (5.6a)-(5.6d) compute first four moments at each location.

The local spatial weights matrix for local Geary's $c_i$ should be elaborated, because a local spatial proximity matrix cannot be directly derived from a global spatial proximity matrix $\Omega - V$. $\Omega_i$ is a matrix of zeroes except for $\omega_{ii}$ that is inherited from $\Omega$, $V_i$ is defined according to equation (5.10), and diag() operation transforms a vector to a diagonal matrix. However, it should be noted that $\omega_{ii}$ is $\sum_{j} v_{ij}$ rather than $\sum_{j} v_{ij}^s$ unlike in global Geary's $c$. Thus, a local spatial weights matrix for Geary's $c_i$ is given:

$$\Omega_i = (V_i + V_i^T) + \text{diag}(V_i) = \begin{bmatrix} v_{ii} & 0 & -v_{ii} \\ 0 & \ddots & \vdots \\ -v_{ii} & \cdots & \sum_{j} v_{ij} - v_{ii} & \cdots & -v_{nn} \\ 0 & \vdots & \ddots & 0 \\ -v_{nn} & 0 & v_{nn} \end{bmatrix}$$ (5.11)
This chapter formulates two generalized randomization methods for significance testing, the Extended Mantel Test and a generalized vector randomization test, and then demonstrates how they are unitized to compute first two moments of SAMs. The section for the Extended Mantel Test is based on Lee (2002). As mentioned in Chapter 2, randomization approach in general is a non-parametric or distribution-free significance testing method, and has some drawbacks, such as ignorance of the projection matrix and violating an assumption of random permutation due to spatially drifting variances (Rogerson 2001). Albeit some drawbacks of randomization approach, it can benefit researchers when they devise a new SAM in the sense that it could provide an initial way of conducting an inferential test for the measure. Most of all, a generalized procedure for randomization approach can be easily customized to obtain equations for first two moments of measures, and thus significantly reduce efforts dedicated to computationally intensive simulations.
6.1 Two randomization tests

6.1.1 The Extended Mantel Test

It has been recognized that both Moran's $I$ and Geary's $c$ are special cases of Mantel (1967)'s generalized cross-product association measure (Glick 1979; Hubert 1978; Sokal 1979; Cliff and Ord 1981; Hubert et al. 1981), and that the associated significance testing method can be used for deriving distributional properties of spatial autocorrelation indices (Cliff and Ord 1981; Upton and Fingleton 1985). Particularly, Hubert et al. (1981) strongly appreciate the benefits of generality that Mantel's generalized measure and the randomization inference associated with it may provide for the spatial association. Practically, when two matrices in Mantel's equation being properly defined for Moran's $I$ and Geary's $c$, equations for the expected value and variance of Mantel's statistic (Mantel 1967; Cliff and Ord 1981:23, Eq. 1.44 ~ 1.46) bring exactly the same set of values as one computed from usually used equations for both measures based on the randomization assumption (see Cliff and Ord 1981:21, Eq. 1.37, 1.39, and 1.42). This opens a possibility that the Mantel's notion of matrix comparison can be extended to bivariate situations, and their distributional properties can be derived accordingly as attempted (Hubert and Golledge 1982; Hubert et al. 1985).

It should be recognized, however, that the link between the Mantel Test and spatial association measures is ensured by an arbitrary restriction on defining a spatial weights matrix; the diagonal elements are set to zeroes. This restriction is often infeasible. On the one hand, there is no definite reason that prevents diagonal elements
from being non-zeroes; the expected value for Moran's $I$ cannot be given by the famous equation any longer, $-1/(n-1)$. On the other hand, it would be inevitable for a certain spatial association measure to have non-zero diagonal elements; Lee’s $L$ cannot be defined with a zero-diagonal spatial weights matrix to be a derivative of the Mantel statistic. Further, if diagonal elements in a spatial weights matrix are different from one another, resulting in a certain level of variance, the equation for the variance in any spatial association measure should be modified to include the variance among diagonal elements and their covariance with off-diagonal elements.

With respect to this, Heo and Gabriel (1998) succeed in extending the Mantel Test by devising a way of dealing with non-zero diagonal elements in any of two matrices involved. Thus, the main purpose of the present paper is to apply their formulations to spatial association measures to obtain an adequate set of distributional properties. Subsequently, I present the Extended Mantel Test based on Heo and Gabriel (1998) by utilizing a set of matrix quantities for the first two moments.

Mantel’s generalized cross-product association measure ($Z$) is originally proposed to explore spatio-temporal dependence among events, and is given by (Mantel 1967):

$$Z = \sum_i \sum_{j \neq i} x_{ij} y_{ij} = \sum_{i,j \neq i} (X \circ Y)$$  \hspace{1cm} (6.1)

where $x_{ij}$ is an element in the spatial (dis)similarity matrix $X$, and $y_{ij}$ is an element in the temporal (dis)similarity matrix $Y$, and $X \circ Y$ denotes a pairwise dot product between two matrices. By calculating the sum of pairwise dot products between two matrices, the
measure evaluates whether there is a certain relationship between the spatial distance and temporal distance between the members of all possible \( n(n-1) \) pairs (Mantel 1967).

Based on the randomization assumption, expected value and variance can be presented as

(6.2) \( E(Z) = \frac{S_0 T_0}{n(n-1)} \)

(6.3) \( \text{Var}(Z) = \frac{S_1 T_1}{2n(n-1)} + \frac{(S_2 - 2S_1)(T_2 - 2T_1)}{4n(n-1)(n-2)} + \frac{(S_0^2 + S_1^2 - 2S_0 S_1)(T_0^2 + T_1^2 + T_2^2)}{n(n-1)(n-2)(n-3)} - E(Z)^2 \)

Note that a restriction of \( i \neq j \) is required for all the elements in (6.2) and (6.3). In other words, the diagonal elements in at least one of two matrices should be set to zeroes.

By eliminating the restriction, a generalized spatial association measure \( (\Gamma) \) is defined as:

(6.4) \[ \Gamma = \sum_i \sum_j P_{ij} q_{ij} = \sum_{i,j} (P \cdot Q) = \text{tr}(P^r Q) = \text{tr}(PQ^r) \]

where \( P \) is a matrix of spatial proximity of locations, and \( Q \) is a matrix of numeric proximity of values on locations (Haining 1990:230). The measure is obtained by summing up all the pairwise dot products or all the diagonal elements of inner or outer
The overall expected value should be decomposed into two elements, one for off-diagonal elements and the other for on-diagonal elements:

\[ E(\Gamma) = E(\Gamma^{\text{off}}) + E(\Gamma^{\text{on}}) \]  

(6.5)

Now, the overall expected value is defined by a sum of expected values for off-diagonal elements and on-diagonal elements. Accordingly, the variance is decomposed into three elements (Heo and Gabriel 1998:847):

\[ \text{Var}(\Gamma) = \text{Var}(\Gamma^{\text{off}}) + \text{Var}(\Gamma^{\text{on}}) + 2 \cdot \text{Cov}(\Gamma^{\text{off}}, \Gamma^{\text{on}}) \]  

(6.6)

The equations that will be presented for all the elements in (6.5) and (6.6) are based on what Mantel calls a “finite population approach” (Mantel 1967:213). This is basically identical to what has been called the randomization approach. The rows and columns of the Q matrix are permuted while the P matrix arbitrarily being kept. He makes a requirement in the permutation process that “if any 2 rows are permuted, the corresponding 2 columns are also permuted so that, for each i, the ith row and ith column will correspond to the same case.” (Mantel 1967:215)

In order to define the original equations ((6.2) and (6.3)), only six quantities are needed (Mantel and Valand 1970). However, since those quantities are relevant only to off-diagonal elements, some additional quantities should be defined, making at least 12 being necessary (Table 6.1). One restriction is that both matrices should be symmetric to
<table>
<thead>
<tr>
<th>Quantities</th>
<th>Notations</th>
<th>P matrix</th>
<th>Q matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{0}^{\text{off}}$</td>
<td>$\sum_{i,j} p_{ij} = 1^T P 1 - \text{tr}(P)$</td>
<td></td>
<td>$G_{0}^{\text{off}}$</td>
</tr>
<tr>
<td>$F_{0}^{\text{on}}$</td>
<td>$\sum_{i} p_{ii} = \text{tr}(P)$</td>
<td></td>
<td>$G_{0}^{\text{on}}$</td>
</tr>
<tr>
<td>$F_{1}^{\text{off}}$</td>
<td>$\sum_{i,j} p_{ij}^2 = \text{tr}(P^T P) - [\text{diag}(P)^T \cdot \text{diag}(P)]$</td>
<td></td>
<td>$G_{1}^{\text{off}}$</td>
</tr>
<tr>
<td>$F_{1}^{\text{on}}$</td>
<td>$\sum_{i} p_{ii}^2 = \text{diag}(P)^T \cdot \text{diag}(P)$</td>
<td></td>
<td>$G_{1}^{\text{on}}$</td>
</tr>
<tr>
<td>$F_{2}^{\text{off}}$</td>
<td>$\sum_{i} \left( \sum_{j \neq i} p_{ij} \right)^2 = [P 1 - \text{diag}(P)]^T \cdot [P 1 - \text{diag}(P)]$</td>
<td></td>
<td>$G_{2}^{\text{off}}$</td>
</tr>
<tr>
<td>$F_{2}^{\text{all}}$</td>
<td>$\sum_{i} \left( \sum_{j \neq i} p_{ij} \right)^2 = (P 1)^T \cdot (P 1) = 1^T (P^T P) 1$</td>
<td></td>
<td>$G_{2}^{\text{all}}$</td>
</tr>
</tbody>
</table>

Table 6.1: Quantities necessary for computing first two moments of spatial association measures
compute those quantities. An asymmetric matrix (e.g., row-standardized spatial weights matrices) can be rendered symmetric by an equation, \( \frac{1}{2} (P + P^T) \).

From (6.5) and (6.6), one may notice that five elements should be obtained in order to compute mean and variance; two for mean (\( E(\Gamma^{off}) \) and \( E(\Gamma^{on}) \)) and three for variance (\( \text{Var}(\Gamma^{off}), \text{Var}(\Gamma^{on}), \) and \( \text{Cov}(\Gamma^{off}, \Gamma^{on}) \)). Table 6.2 presents equations for the five elements with quantities defined in Table 6.1. By combining the first two rows in Table 6.1, we have an equation for the overall expected value:

\[
E(\Gamma) = E(\Gamma^{off}) + E(\Gamma^{on}) = \frac{[1^T P 1] - \text{tr}(P) \cdot [1^T Q 1] - \text{tr}(Q)}{n(n-1)} + \frac{\text{tr}(P) \cdot \text{tr}(Q)}{n} \tag{6.7}
\]

If diagonal elements in any of \( P \) and \( Q \) are zeroes, then the overall expected value reduces to one for off-diagonal elements, because \( E(\Gamma^{on}) \) will be equal to zero. If sums of all the elements in \( P \) and \( Q \) are defined as constant values, \( E(\Gamma^{off}) \) will be further simplified.

For variance, it should be noted that \( \text{Var}(\Gamma^{off}) \) and (6.3) are identical, when the quantities are defined in accordance to Cliff and Ord (1981). It also should be recognized that \( \text{Var}(\Gamma^{on}) \) reduces to zero with a binary connectivity matrix (\( C \)) or its row-standardized version (\( W \)), because all the quantities in the element are zeroes. Even with a spatial weights matrix with non-zero diagonal elements, the variance will remain zero if
<table>
<thead>
<tr>
<th>Moments</th>
<th>Elements</th>
<th>Notations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>$E(\Gamma_{\text{off}})$</td>
<td>$\frac{F_{0}\cdot G_{0}}{n(n-1)}$ = $\left(1^T P 1 - \text{tr}(P)\right) \cdot \left(1^T Q 1 - \text{tr}(Q)\right)$</td>
</tr>
<tr>
<td></td>
<td>$E(\Gamma_{\text{on}})$</td>
<td>$\frac{F_{0}\cdot G_{0}}{n}$ = $\text{tr}(P) \cdot \text{tr}(Q)$</td>
</tr>
<tr>
<td>Variance</td>
<td>$\text{Var}(\Gamma_{\text{off}})$</td>
<td>$\frac{2F_{0}\cdot G_{0}}{n(n-1)} + 4\frac{(F_{0} - F_{0}) (G_{0} - G_{0})}{n(n-1)(n-2)} + \frac{(F_{0})^2 + 2F_{0} - 4F_{0} + 2G_{0} - 4G_{0}}{n(n-1)(n-2)(n-3)}$ $- E(\Gamma_{\text{off}})^2$</td>
</tr>
<tr>
<td></td>
<td>$\text{Var}(\Gamma_{\text{on}})$</td>
<td>$\frac{F_{0}\cdot G_{0}}{n(n-1)}$ = $\text{tr}(P) \cdot \text{tr}(Q)$</td>
</tr>
<tr>
<td></td>
<td>$\text{Cov}(\Gamma_{\text{off}}, \Gamma_{\text{on}})$</td>
<td>$\frac{(F_{2} - F_{2} \cdot F_{2}) (G_{2} - G_{2} \cdot G_{2})}{2n(n-1)} + \frac{(F_{0} \cdot F_{0}) - (F_{2} - F_{2} \cdot F_{2}) (G_{0} \cdot G_{0}) - (G_{2} - G_{2} \cdot G_{2})}{n(n-1)(n-2)} + E(\Gamma_{\text{off}}) \cdot E(\Gamma_{\text{on}})$</td>
</tr>
</tbody>
</table>

Table 6.2: Five elements for computing mean and variance of spatial association measures with quantities
they have a constant value, which can be easily verified from the notation for $\text{Var}(\Gamma^0)$ in Table 6.2.

For covariance ($\text{Cov}(\Gamma^{off}, \Gamma^0)$), one can notice that, if there is no variance among diagonal elements in any of $P$ and $Q$ due to either zero-diagonal or constant diagonal, then the overall variance reduces to the variance only for off-diagonal elements, because the covariance will be zero.

6.1.2 A generalized vector randomization test

A generalized vector randomization test is a very simple way of deriving distributional moments especially for local SAMs that will be seen later on. A general statistic conforming to the test is given:

$$\Gamma_i = p_i^{(i)} \cdot q_i^{(i)} + \sum_i (p_i^{(i)} \cdot q_i^{(i)})$$

where $p_i^{(i)}$ and $q_i^{(i)}$ are $i$th entries in $p^{(i)}$ and $q^{(i)}$ that are respectively a local spatial proximity vector and a local numeric proximity vector, and $p^{(-i)}$ and $q^{(-i)}$ are $(n-1)$-by-1 vectors derived from $p^{(i)}$ and $q^{(i)}$ by eliminating the $i$th elements. In this specification, the sampling distribution of local measures is determined by a permutation between two vectors, $p^{(-i)}$ and $q^{(-i)}$. Hubert (1984:453; 1987:28) provides equations for the expected value and variance of a measure that is defined as a sum of pairwise products of two
vectors with \( n \) observations. By slightly modifying the equations, the expected value and variance for local measures are given respectively by:

\[
E(\Gamma_i) = p_i^{(i)}q_i^{(i)} + \frac{1}{n-1} \cdot \sum_i p_i^{(-i)} \cdot \sum_i q_i^{(-i)}
\]

\[
Var(\Gamma_i) = \frac{1}{n-2} \cdot \sum_i (p_i^{(-i)} - \overline{p}^{(-i)})^2 \cdot \sum_i (q_i^{(-i)} - \overline{q}^{(-i)})^2
\]

where \( p_i^{(-i)} \) and \( q_i^{(-i)} \) are entries in \( p^{(-i)} \) and \( q^{(-i)} \), and \( \overline{p}^{(-i)} \) and \( \overline{q}^{(-i)} \) are mean values of those vectors. I would call a procedure from (6.8) to (6.9) a \textit{generalized vector randomization test}.

6.2 Application of randomization tests to spatial association measures

Here I deal with ten SAMs, five global ones (Moran’s \( I \), Geary’s \( c \), Lee’s \( S \), Cross-Moran, and Lee’s \( L \) and their local versions. Table 6.3 summarizes which randomization testing methods can be applied to what SAMs. The Extended Mantel Test can be used for all the global SAMs and local SAMs when a total randomization is assumed. It is also applied to local \( S_i \) and local \( L_i \) when a conditional randomization assumption. The generalized vector randomization test is utilized for all the local SAMs under the conditional randomization except for local \( S_i \) and \( L_i \). Subsequently, I demonstrate how the general procedures presented in the previous section are applied to particular SAMs.
<table>
<thead>
<tr>
<th>SAMs</th>
<th>Total Randomization</th>
<th>Conditional Randomization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moran's I</td>
<td>Global M</td>
<td>N/A</td>
</tr>
<tr>
<td>Local M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geary's c</td>
<td>Global M</td>
<td>N/A</td>
</tr>
<tr>
<td>Local M</td>
<td></td>
<td>V</td>
</tr>
<tr>
<td>Lee's S</td>
<td>Global M</td>
<td>N/A</td>
</tr>
<tr>
<td>Local M</td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Bivariate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-Moran</td>
<td>Global M</td>
<td>N/A</td>
</tr>
<tr>
<td>Local M</td>
<td></td>
<td>V</td>
</tr>
<tr>
<td>Lee's L</td>
<td>Global M</td>
<td>N/A</td>
</tr>
<tr>
<td>Local M</td>
<td></td>
<td>M</td>
</tr>
</tbody>
</table>

(M: The Extended Mantel Test; V: A Generalized Vector Randomization Test; N/A: Not Applicable)

Table 6.3: Significance testing methods for spatial association measures under the randomization assumption
6.3 Significance testing for global spatial association measures

Table 6.4 lists the five global SAMs in a matrix notation. Without loss of generality for the spatial weights matrix, V allows for any way of defining topological relationships among observations. The matrix notation column is presented by utilizing the z-transformation of variable(s). The matrix notation for Geary’s c should be elaborated. Cliff and Ord (1981:167) demonstrate that Geary’s c can be presented in a quadratic form like Moran’s I. The matrix of \( \Omega \) is defined as a diagonal matrix with nonzero elements, each of which is defined as:

\[
\omega_{ii} = \sum_j v_{ij}^2 (v^s = \frac{1}{2} (V + V^T)) \tag{6.11}
\]

Note that the diagonal of \( \Omega \) becomes identical to a vector of row-sums of V, when V is symmetric.

By using the equations in the matrix notation, one can define the matrices of P and Q for each spatial association measure in order to conform to the general form in (6.4) (Table 6.5). For all the measures, the matrix of P is defined as a standardized form of the spatial weights matrix. P matrix for Geary’s c should be further elaborated. From (6.11) for a symmetric V, the matrix of \( \Omega - V \) is presented as:
<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Measures</th>
<th>Matrix Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate</td>
<td>Moran’s ( I )</td>
<td>[ I_x = \frac{(z_x)^T V z_x}{1^T V 1} ]</td>
</tr>
<tr>
<td></td>
<td>Geary’s ( c )</td>
<td>[ c_x = \frac{n-1}{n} \cdot \frac{(z_x)^T (\Omega - V) z_x}{1^T V 1} ]</td>
</tr>
<tr>
<td></td>
<td>Lee’s ( S )</td>
<td>[ S_x = \frac{(z_x)^T (V^T V) z_x}{1^T (V^T V) 1} ]</td>
</tr>
<tr>
<td>Bivariate</td>
<td>Cross-Moran ( I_{x,y} )</td>
<td>[ I_{x,y} = \frac{(z_x)^T V z_y}{1^T V 1} ]</td>
</tr>
<tr>
<td></td>
<td>Lee’s ( L )</td>
<td>[ L_{x,y} = \frac{(z_x)^T (V^T V) z_y}{1^T (V^T V) 1} ]</td>
</tr>
</tbody>
</table>

Table 6.4: Global spatial association measures in a matrix notation
**Table 6.5: Definitions of P and Q, and equations for expected values for spatial association measures**

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>Q</th>
<th>E((\Gamma))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Gamma)</td>
<td>Definition</td>
<td>1(\Gamma P)</td>
<td>tr((\Gamma P))</td>
</tr>
<tr>
<td>Moran's I</td>
<td>(\frac{V}{1^T V 1})</td>
<td>1</td>
<td>(\frac{tr(V)}{1^T V 1})</td>
</tr>
<tr>
<td>Geary's c</td>
<td>(\frac{n-1}{n} \left(\frac{\Omega - V}{1^T V 1}\right))</td>
<td>0</td>
<td>(\frac{n-1}{n} \left(\frac{1 - tr(V)}{1^T V 1}\right))</td>
</tr>
<tr>
<td>Lee's S</td>
<td>(\frac{V^T V}{1^T (V^T V) 1})</td>
<td>1</td>
<td>(\frac{tr(V^T V)}{1^T (V^T V) 1})</td>
</tr>
<tr>
<td>Cross-Moran</td>
<td>(\frac{V}{1^T V 1})</td>
<td>1</td>
<td>(\frac{tr(V)}{1^T V 1})</td>
</tr>
<tr>
<td>Lee's L</td>
<td>(\frac{V^T V}{1^T (V^T V) 1})</td>
<td>1</td>
<td>(\frac{tr(V^T V)}{1^T (V^T V) 1})</td>
</tr>
</tbody>
</table>
Since the sum of each row is equal to zero, the sum of all the elements in the matrix is also equal to zero. The sum of diagonal elements is given:

$$\text{tr}(\Omega - V) = \sum_{j} \sum_{j} v_{ij} - \sum_{i} v_{ii}$$

$$= 1^T V 1 - \text{tr}(V)$$

(6.13)

Thus, the sum of diagonal elements in $P$ is given as seen in Table 6.5:

$$\text{tr}(P) = \frac{n-1}{n} \left( 1 - \frac{\text{tr}(V)}{1^T V 1} \right)$$

(6.14)

The matrix of $Q$ is identical between Moran’s $I$, Geary’s $c$, and Lee’s $S$ and between Cross-Moran and Lee’s $L$. The outer product of z-transformed variable $X$ is presented:
The sum of \( i \)th row without its diagonal element is given by:

\[
\sum_{j=1}^{n} q_{ij} = (x_i - \bar{x}) \cdot \left[ (x_1 - \bar{x}) + \ldots + (x_{i-1} - \bar{x}) + (x_{i+1} - \bar{x}) + \ldots + (x_n - \bar{x}) \right] = -\frac{(x_i - \bar{x})^2}{s_x^2}
\]  

(6.16)

Therefore, the sum of all the off-diagonal elements in \( Q \) can be computed by summing up all the row-sums:

\[
\sum_{i=1}^{n} \sum_{j \neq i} q_{ij} = -\sum_{i=1}^{n} (x_i - \bar{x})^2 = -n
\]  

(6.17)

The sum of on-diagonal elements is given by:
\[
\text{tr}(Q) = \sum_{i=1}^{n} q_{ii} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{s_x^2} = n
\]  

(6.18)

Thus, the sum of all the elements in \(Q\) (\(1^TQ1\)) is zero (see Table 6.5).

The outer product of z-transformed variables \(X\) and \(Y\) defines the matrix of \(Q\) for Cross-Moran and Lee's \(L\), and is given:

\[
Q = \begin{bmatrix}
\frac{(x_1 - \bar{x})(y_1 - \bar{y})}{s_xs_y} & \frac{(x_1 - \bar{x})(y_2 - \bar{y})}{s_xs_y} & \cdots & \frac{(x_1 - \bar{x})(y_n - \bar{y})}{s_xs_y} \\
\frac{(x_2 - \bar{x})(y_1 - \bar{y})}{s_xs_y} & \frac{(x_2 - \bar{x})(y_2 - \bar{y})}{s_xs_y} & \cdots & \frac{(x_2 - \bar{x})(y_n - \bar{y})}{s_xs_y} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{(x_n - \bar{x})(y_1 - \bar{y})}{s_xs_y} & \frac{(x_n - \bar{x})(y_2 - \bar{y})}{s_xs_y} & \cdots & \frac{(x_n - \bar{x})(y_n - \bar{y})}{s_xs_y}
\end{bmatrix}
\]  

(6.19)

The sum of \(i\)th row without its diagonal element is given by:

\[
\sum_{j=1, j \neq i}^{n} q_{ij} = \frac{(x_i - \bar{x})\left[\sum_{i=1}^{n} (y_i - \bar{y})\right] + (y_i - \bar{y}) + (y_{i+1} - \bar{y}) + \cdots + (y_n - \bar{y})}{s_x s_y}
\]

(6.20)

\[
\begin{align*}
\sum_{j=1, j \neq i}^{n} q_{ij} &= \frac{(x_i - \bar{x})\left[\sum_{i=1}^{n} (y_i - \bar{y})\right] - (y_i - \bar{y})}{s_x s_y} \\
&= \frac{-(x_i - \bar{x})(y_i - \bar{y})}{s_x s_y}
\end{align*}
\]

Therefore, the sum of all the row-sums in \(Q\) is given by:

112
The sum of diagonal elements is given by:

\[
\sum_{i} \sum_{j=i} q_{ij} = \frac{-\sum (x_i - \bar{x})(y_j - \bar{y})}{s_x s_y} \quad \text{(6.21)}
\]

\[= -n \cdot r_{x,y}\]

The sum of diagonal elements is given by:

\[
\text{tr}(Q) = \sum_{i} q_{ii} = \frac{-\sum (x_i - \bar{x})(y_i - \bar{y})}{s_x s_y} \quad \text{(6.22)}
\]

\[= n \cdot r_{x,y}\]

From (6.21) and (6.22), it is acknowledged that the sum of off-diagonal elements is identical to Pearson's correlation coefficient between two variables, multiplied by \(-n\), and that the sum of on-diagonal elements is the same as Pearson's correlation coefficient, multiplied by \(n\). In addition, the sum of all the elements in \(Q\) is equal to zero, as in \(Q\) for univariate SAMs (see Table 6.5).

When all these definitions are applied to (6.7), one can formulate an equation for the expected value specific to each spatial association measure as seen in the last column in Table 6.5. With a binary connectivity matrix (C) or its row-standardized version (W), the equation reduces to the well-known equation, \(-1/(n-1)\) for Moran's \(I\) and 1 for Geary's \(c\), since \(\text{tr}(V)\) is equal to zero. Those familiar expected values, however, do not hold for a spatial weights matrix with non-zero diagonal elements. The equations
The expected value for Cross-Moran is approximated by \(- r_{x,r} / (n - 1)\) with \(C\) or \(W\), which proves Griffith's findings (Griffith, 1993, page 111; Griffith and Amrhein, 1997, page 48). As Tiefelsdorf (2001) and Lee (2001) point out, however, Cross-Moran should not be used as a bivariate spatial association measure. The expected values for Lee's \(S\) and Lee's \(L\) hardly reduces to a simpler equation, because \(V^T V\) always has non-zero diagonal elements regardless of \(V\). When a row-standardized matrix \(W^*\) is applied, the equations are simplified respectively to:

\[
\frac{\text{tr}(W^T W) - 1}{n - 1}
\]

(6.23)

\[
\frac{\text{tr}(W^T W) - 1}{n - 1} \cdot r_{x,y}
\]

(6.24)

### 6.4 Significance testing for local spatial association measures

#### 6.4.1 Defining a general form of local spatial association measures and different types of randomization

A general form of global spatial association measures has been defined as (Anselin 1995; Lee 2001b):
\[ \Gamma = \sum_i \sum_j p_{ij} q_{ij} = \sum_i (P \cdot Q) = \text{tr}(P^T Q) = \text{tr}(Q^T P) \]  
\hspace{1cm} (6.25) 

where \( p_{ij} \) and \( q_{ij} \) are elements respectively in a spatial proximity matrix \( P \) and a numeric proximity matrix \( Q \). Accordingly, Anselin (1995:98) formulates a general form of local spatial association measures as:

\[ \Gamma_i = \sum_j p_{ij} q_{ij} \]  
\hspace{1cm} (6.26)

Now, a local spatial association measure at \( i \)th location is computed by summing up the casewise products of two \( i \)th rows in \( P \) and \( Q \).

However, this specification is somewhat misleading. First, equation (6.26) obviously leads to a conclusion that all the local measures sum to their corresponding global measure, but this in turn contradicts to his statement on the relationships between a global Moran's \( I \) and its local ones: i.e., "the average of the \( I_i \) will equal to the global \( I \)." (Anselin 1995:100) Contemplated on the more restrictive *additivity* requirement, equation (6.26) should be replaced by:

\[ \Gamma_i = n \cdot \sum_j p_{ij} \cdot q_{ij} \]  
\hspace{1cm} (6.27)

Second, in the context of the significance testing based on randomization assumptions, Anselin's specification, as Sokal et al. (1998) correctly point out, leads to
different general forms for the local spatial association. Equation (6.26) could be restated by:

\[
\sum_{i,j} (P^{(i)} \cdot Q), \text{ or } \tag{6.28}
\]

\[
\sum_{i,j} (P \cdot Q^{(i)}), \text{ or } \tag{6.29}
\]

\[
\sum_{i,j} (P^{(i)} \cdot Q^{(i)}), \text{ or } \tag{6.30}
\]

where \( P^{(i)} \) and \( Q^{(i)} \) are certain forms respectively of a local spatial proximity matrix and a local numeric proximity matrix. These three different definitions of the local spatial association yield an identical set of local measures but with different sampling distributions. From this observation, Sokal et al. (1998:335) incorrectly conclude that "when we permit only conditional permutations, all three Mantel versions of the LISA will give rise to the same (conditional) reference distribution" and "there is not a unique total reference distribution for a LISA, hence no unique set of total moments."

Conceptually, the three different specifications correspond to three different randomization assumptions, and thus it is not unusual to observe that they yield different sets of distributional properties.

First, equation (6.28) corresponds to what I call 'location-based total randomization' that is identical to what Sokal et al. (1998) calls 'total randomization', where a local spatial configuration consisting of \( i \)th location and its neighbors is fixed and different sets of values are permuted over there. A value being set to \( i \)th location, all
the other values are permuted to define a set of neighbors with resulting in a set of local statistics, and then another value is set to the \( i \)th location and the same procedure is undertaken with resulting in another set of local statistics. In this randomization, the expected value is the average value of all possible local measures that can be given to the \( i \)th location. Subsequently, the term of total randomization will be used for this location-based total randomization.

Second, equation (6.29) corresponds to what may be called 'value-based total randomization', where a numeric vector consisting of \( n \) observations are permuted over the given spatial configuration with \( i \)th value always being positioned at a reference spot. An \( i \)th value being set to a location, all the other values are permuted to define a set of neighbors, and then the \( i \)th value moves on to another area and the same procedure is undertaken. In this type of randomization, the expected value is the average value of all possible local measures that can be given to the \( i \)th value. Obviously, this randomization is not relevant to spatial statistics.

Third, in a conceptual sense, equation (6.30) conforms to what has been called the 'conditional randomization' (Anselin 1995; Sokal et al. 1998), where an \( i \)th value is set on the corresponding \( i \)th location and all the other values are permuted to constitute a set of its neighbors over the given local spatial configuration. This more restrictive randomization scheme gives rise to another issue. The local measure does not have to be defined by matrices. Rather, it is better seen as the sum of the casewise products between two vectors that are drawn from the corresponding global matrices (Sokal et al. 1998). In other words, a local measure can be given as \( \sum_i p_i q_i \) rather then \( \sum_{i,j} p_i q_j \) (where \( p_i \)
and are elements in certain forms of vectors derived from $P^{(i)}$ and $Q^{(i)}$. This implies that the Extended Mantel Test cannot be applied to the conditional randomization. Rather, a new set of equations that deals with permutations between two vectors should be used to compute a set of distributional moments. Further, the new set of equations for the first two moments should embrace the fact that an $i$th case in both vectors should not be involved in a permutation procedure. This will be discussed in the next section.

A bivariate situation requires a further clarification, because two elements in pairs between two variables should be specifically defined along with their relationships with spatial configurations. There might be three ways of defining the relationships among the three elements (two variables and their spatial settings). One way is first to assign a value in $X$ to a location and permute other values to define its neighbors. For each local setting in $X$, values in $Y$ are randomly permuted to define a local spatial setting in the other side. As in the univariate situation, then, another value in $X$ is set to the location and all the permutation procedure is repeated. In this case, the link between $x_i$ and $y_i$ in the original variables does not have to be maintained. Second, $i$th values in $X$ and $Y$ are bound to each other and are assigned to the $i$th location in both $X$ and $Y$. Their neighbors, then, are defined by permuting all the other values on each side. In this case, elements in a pair of neighbors, $x_j$ and $y_j$, need not follow the geographical reference in the original vectors. In other words, once pivot values at a location are determined from the original order, permutations for neighbors are independently conducted between two variables. Third way is to undertake the permutation procedure with all the values in $X$ and $Y$ being bound to each other according to the original order. For example, if $x_j$ is chosen as $x_i$'s
neighbor in a given location, the corresponding $y_j$ should be placed on that location as a $y_i$'s neighbor. Once the permutation is done for the $i$th pair, another pair of values is assigned to that location and all the permutation procedure is repeated for a total randomization.

In the context of the randomization approach, only the third way is relevant as can be seen from a Monte Carlo simulation conducted for Lee's $L$ (Lee 2001b). A conditional randomization for bivariate situations can be easily conceptualized. A permutation is conducted only on a location, not furthering onto other locations. That is, $i$th pair, whose elements are bound to each other, is assigned to $i$th location, and all the other pairs, whose elements are also bound to each other, are permuted to define two sets of neighbors.

Table 6.6 lists the five local spatial association measures for which I will attempt to derive the first two moments based on a generalized randomization approach. The equation for local Geary's $c_i$ is modified from Anselin's original one in order to make the average value of local measures equal to the global measure. Thus, all the four local measures satisfy the more restrictive additivity requirement that can be expressed by:

$$
\Gamma = \frac{\sum \Gamma_i}{n} \quad (6.31)
$$

The local Cross-Moran measure was derived from Wartenberg (1985)'s global measure, and local Lee's $L_i$ follows Lee's definition (Lee 2001b). Getis-Ord's $G$-statistics are not
<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Measures</th>
<th>Matrix Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate</td>
<td>Local Moran's $I_i$</td>
<td>$I_i = n \cdot \frac{(z_x)^\top V_i z_x}{1^\top V 1}$</td>
</tr>
<tr>
<td></td>
<td>Local Geary's $c_i$</td>
<td>$c_i = \frac{n-1}{2} \cdot \frac{\text{tr} \left( (z_x 1^\top)^2 - (z_x)^\top V_i^2 \right)}{1^\top V 1}$</td>
</tr>
<tr>
<td></td>
<td>Lee's $S_i$</td>
<td>$S_i = n \cdot \frac{(z_x)^\top (V_i 1^\top) z_x}{1^\top (V^2 1)}$</td>
</tr>
<tr>
<td>Bivariate</td>
<td>Local Cross-Moran</td>
<td>$I_i = n \cdot \frac{(z_x)^\top V_i z_y}{1^\top V 1}$</td>
</tr>
<tr>
<td></td>
<td>Lee's $L_i$</td>
<td>$L_i = n \cdot \frac{(z_x)^\top (V_i 1^\top) z_y}{1^\top (V^2 1)}$</td>
</tr>
</tbody>
</table>

(The $^{(2)}$ operation for local Geary's $c_i$ denotes the squaring of all the elements in a matrix)

Table 6.6: Local spatial association measures in a matrix notation
examined here mainly because it does not have a corresponding global measure satisfying
the additivity requirement. However, the significance testing method presented here will
easily apply to those measures.

6.4.2 Defining local matrices and vectors

It is necessary to define some local matrices and vectors in order subsequently to
define general forms of spatial association measures based on total and conditional
randomization assumptions. Tiefelsdorf (1998) defines a local spatial weights matrix in a
symmetric star-shaped form. However, I here define a local spatial weights matrix \( V_i \) by
assigning zeros to all the elements except for ones on \( i \)th row in a global spatial weights
matrix \( V \).

\[
V_i = \begin{bmatrix}
0 & & \\
\vdots & \ddots & \ddots \\
& \ddots & \ddots & \ddots \\
& & \ddots & \ddots & \ddots \\
0 & & & & \ddots \\
\end{bmatrix}
\] (6.32)

This non-symmetric form of a local spatial weights matrix is preferred not only
because the symmetricity required for the exact distribution approach can be preserved by
a transformation function, \( V_i^S = \frac{1}{2} \left( V_i + V_i^T \right) \), but because the symmetric form does not
work for bivariate measures. As a global spatial proximity matrix \( P \) is a normalized form
of a global spatial weights matrix $V$, a local spatial-proximity matrix $P^{(l)}$ is defined as a
normalized form of $V$, multiplied by $n$ (Table 6.7). Thus, $P^{(l)}$ for the first three local
statistics is derived from $P$ by replacing all entries with zeros except for ones on $i$th row
and multiplying it by $n$.

$$
P^{(l)} = n \cdot \begin{bmatrix}
0 \\
p_{in} & \cdots & p_{ii} & \cdots & p_{im} \\
0 & \cdots & 0 & \cdots & 0
\end{bmatrix}
$$

(6.33)

and $P^{(l)} = n \cdot (p_{in}, \ldots, p_{ii}, \ldots p_{im})^T$

(6.34)

As mentioned before, the number of observations ($n$) should be multiplied for local
measures in order to satisfy the more restrictive additivity requirement defined in (6.31).
However, $P^{(l)}$ for local Lee's $S_l$ and local Lee's $L_l$ is not defined as in (6.33); rather, it
takes a much more complicated form:
<table>
<thead>
<tr>
<th>Local Moran's $I_i$</th>
<th>Local Geary's $c_i$</th>
<th>Local Lee's $S$</th>
<th>Local Cross-MC</th>
<th>Local Lee's $L_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P$ [ \frac{V}{1^T V 1} ] [ \frac{n-1}{2n} \cdot \frac{V}{1^T V 1} ] [ V ] [ V ] [ V ] [ V ] [ V ] [ V ]</td>
<td>$P^{(i)}$ [ n \cdot \frac{V_i}{1^T V 1} ] [ \frac{n-1}{2} \cdot \frac{V_i}{1^T V 1} ] [ n \cdot \frac{V_i}{1^T V 1} ] [ n \cdot \frac{V_i}{1^T V 1} ] [ n \cdot \frac{V_i}{1^T V 1} ] [ n \cdot \frac{V_i}{1^T V 1} ] [ n \cdot \frac{V_i}{1^T V 1} ] [ n \cdot \frac{V_i}{1^T V 1} ]</td>
<td>Defined according to (6.33)</td>
<td>Not defined (6.33)</td>
<td>Not defined (6.33)</td>
</tr>
<tr>
<td>$Q$ [ z_x \cdot (z_x)^T ] [ z_x 1^T - 1 (z_x)^T ] [ z_x \cdot (z_x)^T ]</td>
<td>[ z_x \cdot (z_x)^T ]</td>
<td>Defined according to (6.36)</td>
<td>$z_x \cdot (z_x)^T$</td>
<td>$z_x \cdot (z_x)^T$</td>
</tr>
<tr>
<td>$Q^{(i)}$</td>
<td>[ z_{x_i} \cdot z_{x} ] [ z_{x_i} \cdot z_{x} ] [ z_{x_i} \cdot z_{x} ]</td>
<td>[ z_{x_i} \cdot z_{x} ] [ z_{x_i} \cdot z_{x} ] [ z_{x_i} \cdot z_{x} ]</td>
<td>[ z_{x_i} \cdot z_{x} ] [ z_{x_i} \cdot z_{x} ] [ z_{x_i} \cdot z_{x} ]</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.7: Definition of matrices and vectors
Obviously, rows other than the $i$th one should have non-zero entries, and all entries in the $i$th row and column will be zero if $V$ is a spatial weights matrix with a zero-diagonal.

In accordance to $p^{(i)}$, a local numeric proximity matrix $Q^{(i)}$ for all the four local measures is defined as (see Table 6.7):

$$Q^{(i)} = \begin{bmatrix}
0 \\
q_{ii} & \cdots & q_{ii} & \cdots & q_{in} \\
0 
\end{bmatrix} \quad (6.36)$$

and $q^{(i)} = (q_{ii}, \cdots, q_{ii}, \cdots, q_{in})^T \quad (6.37)$

where $q_{ij}$ is an element on an $i$th row in a global proximity matrix $Q$. Table 6.7 also shows how a column vector of $q^{(i)}$ can be specified for each of the local measures. For
example, a vector of \( q^{(i)} \) for local Moran's \( I_i \) is given by \( z_{xi} - \bar{z}_X \) where \( z_{xi} \) is an \( i \)th element in a standardized form of variable \( X \) (subtracted by a mean and divided by a standard deviation).

### 6.4.3 Total randomization and the Extended Mantel Test

A general local spatial association measure based on the total randomization assumption is given by:

\[
\Gamma_i = \sum_{i,j} (P^{(i)} - Q) = \text{tr}(P^{(i)T} Q) = \text{tr}(P^{(i)} Q) \tag{6.38}
\]

where \( P^{(i)} \) and \( Q \) for each local measure are seen in Table 6.7. Since a general form of local spatial association measure here is defined as relationships between two matrices, the Extended Mantel Test presented in previous section applies to compute the expected value and variance. The difference between global and local measures in terms of the sampling distribution result only from differences between \( P \) and \( P^{(i)} \).

Table 6.8 summarizes necessary quantities and an equation for the expected value for each measure. When necessary quantities are computed, the expected value is obtained by as in global measures:

\[
E(\Gamma_i) = \left[ \text{tr}(P^{(i)}) \right] \cdot \left[ (I^T Q 1) - \text{tr}(Q) \right] + \frac{\text{tr}(P^{(i)}) \cdot \text{tr}(Q)}{n(n-1)} \tag{6.39}
\]

125
<table>
<thead>
<tr>
<th>Measure</th>
<th>$\mathbf{P}^{(i)}$</th>
<th>$\mathbf{Q}$</th>
<th>$\mathbf{E}(\Gamma_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mathbf{1}^T \mathbf{P}^{(i)} \mathbf{1}$</td>
<td>$\mathbf{tr}(\mathbf{P}^{(i)})$</td>
<td>$\mathbf{1}^T \mathbf{Q} \mathbf{1}$</td>
</tr>
<tr>
<td>Local Moran's $I_i$</td>
<td>$\frac{n}{\mathbf{1}^T \mathbf{V} \mathbf{1}} \cdot v_i$</td>
<td>$\frac{n}{\mathbf{1}^T \mathbf{V} \mathbf{1}} \cdot v_u$</td>
<td>0</td>
</tr>
<tr>
<td>Local Geary's $c_i$</td>
<td>$\frac{n-1}{2 \cdot \mathbf{1}^T \mathbf{V} \mathbf{1}} \cdot v_i$</td>
<td>$\frac{n-1}{2 \cdot \mathbf{1}^T \mathbf{V} \mathbf{1}} \cdot v_u$</td>
<td>$2 \cdot n^2$</td>
</tr>
<tr>
<td>Local Lee's $S_i$</td>
<td>$\frac{n}{\mathbf{1}^T (\mathbf{V}^T \mathbf{V}) \mathbf{1}} \cdot v_i^2$</td>
<td>$\frac{n}{\mathbf{1}^T (\mathbf{V}^T \mathbf{V}) \mathbf{1}} \cdot v_i^{(2)}$</td>
<td>0</td>
</tr>
<tr>
<td>Local Cross-Moran</td>
<td>$\frac{n}{\mathbf{1}^T \mathbf{V} \mathbf{1}} \cdot v_i$</td>
<td>$\frac{n}{\mathbf{1}^T \mathbf{V} \mathbf{1}} \cdot v_u$</td>
<td>0</td>
</tr>
<tr>
<td>Local Lee's $L_i$</td>
<td>$\frac{n}{\mathbf{1}^T (\mathbf{V}^T \mathbf{V}) \mathbf{1}} \cdot v_i^2$</td>
<td>$\frac{n}{\mathbf{1}^T (\mathbf{V}^T \mathbf{V}) \mathbf{1}} \cdot v_i^{(2)}$</td>
<td>0</td>
</tr>
</tbody>
</table>

$v_i = \sum_j v_{ij}$ and $v_i^{(2)} = \sum_j v_{ij}^2$.

Table 6.8: Expected values for local spatial association measures under the total randomization
Equation (6.39) can apply to any spatial weights matrices. When a row-standardized matrix (W) with a zero diagonal is concerned, the equations are more simplified. One can easily recognize that expected values are reduced to \(-1/(n-1)\), 1, and \(-r_{x,r}/(n-1)\) respectively for the first three local measures, and that they are identical to those for the corresponding global measures, which does not hold for other spatial weights matrices with non-zero diagonals. The expected values for local Lee's \(S_i\) and local Lee's \(L_i\) are also simplified to with W or \(W^*\) when the number of neighbors of \(i\)th observation is denoted by \(n_i\) (note that \(n_i\) is always one-degree larger in \(W^*\) than in W):

\[
E(S_i) = \frac{\left(\frac{n}{n_i}\right) - 1}{n - 1} \tag{6.40}
\]

\[
E(L_i) = \frac{\left(\frac{n}{n_i}\right) - 1}{n - 1} \cdot r_{x,r} \tag{6.41}
\]

6.4.4 Conditional randomization and a generalized vector randomization test

By following the generalized vector randomization test with necessary quantities, expected values for three local measures can be derived (Table 6.9). If a row-standardized spatial weights matrix with zeros in its diagonal (W) is concerned, the expected value for the measures are computed respectively by \(-z_{x,i}^2/(n-1)\), \((z_{x,i}^2 + 1)/2\), and \(-z_{x,i}^2 \cdot z_{r,i}^2/(n-1)\). Whereas expected values under the total randomization are
<table>
<thead>
<tr>
<th>$\Gamma_i$</th>
<th>$p_i^{(l)}$</th>
<th>$\sum_i p_i^{(l)}$</th>
<th>$q_i^{(l)}$</th>
<th>$\sum_i q_i^{(l)}$</th>
<th>$E(\Gamma_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local Moran's $I_i$</td>
<td>$\frac{n}{1^T V 1} \cdot v_k$</td>
<td>$\frac{n}{1^T V 1} \cdot (v_k - v_u)$</td>
<td>$z_{x_i}^2$</td>
<td>$-z_{x_i}^2$</td>
<td>$\frac{n}{1^T V 1} \cdot \frac{n \cdot v_k - v_u \cdot z_{x_i}^2}{n - 1}$</td>
</tr>
<tr>
<td>Local Geary’s $c_i$</td>
<td>$\frac{n-1}{2 \cdot 1^T V 1} \cdot v_k$</td>
<td>$\frac{n-1}{2 \cdot 1^T V 1} \cdot (v_k - v_u)$</td>
<td>0</td>
<td>$n \cdot (z_{x_i}^2 + 1)$</td>
<td>$\frac{n}{1^T V 1} \cdot \frac{v_k - v_u \cdot (z_{x_i}^2 + 1)}{2}$</td>
</tr>
<tr>
<td>Local Cross-Moran</td>
<td>$\frac{n}{1^T V 1} \cdot v_k$</td>
<td>$\frac{n}{1^T V 1} \cdot (v_k - v_u)$</td>
<td>$z_{x_i} \cdot z_y$</td>
<td>$-z_{x_i} \cdot z_y$</td>
<td>$\frac{n}{1^T V 1} \cdot \frac{n \cdot v_k - v_u \cdot z_{x_i} \cdot z_y}{n - 1}$</td>
</tr>
</tbody>
</table>

$v_k = \sum_j v_j$.  

Table 6.9: Expected values for local spatial association measures under the conditional randomization
identical for all the locations, ones under the conditional randomization are dependent upon values on each location. They are slightly different from ones based on the total randomization in Table 4. The equations for local Moran's $I_i$ and Geary's $c_i$ are identical to ones proposed by Sokal et al. (1998) (note that local Geary’s $c_i$ is differently defined in this paper). Variances are also identical to ones calculated by their equations.

An inferential test for local Lee’s $S_i$ and local Lee’s $L_i$ based on the conditional randomization is much more complicated. The equation for both measures is given by:

$$S_i \text{ or } L_i = p_{ui}^{(i)} \cdot q_{ui}^{(i)} + \sum_{i,j} (P^{(-i)} \cdot Q^{(-i)})$$

(6.42)

where $p_{ui}^{(i)}$ and $q_{ui}^{(i)}$ are entries in $P^{(i)}$ and $Q^{(i)}$, and $P^{(-i)}$ and $Q^{(-i)}$ are $(n-1)$-by-$(n-1)$ matrices derived from global matrices $P$ and $Q$. $P^{(-i)}$ is obtained by eliminating $i$th row and column from $P$. $Q^{(-i)}$ for Lee’s $S_i$ is given when a spatial weights matrix with a zero-diagonal:

$$Q^{(-i)} = z_{X}^{(-i)} \cdot (z_{X}^{(-i)})^T$$

(6.43)

where $z_{X}^{(-i)}$ is a column vector derived from $z_X$ by eliminating $i$th entry, and thus $Q^{(-i)}$ is a $(n-1)$-by-$(n-1)$ matrix obtained by eliminating $i$th row and column from $Q$. When a spatial weights matrix with a non-zero diagonal, however, the matrix takes a more complex form:
\[ Q^{(-i)} = z_x^{(-i)} \cdot (z_x^{(-i)})^T + 2 \cdot \text{diag}(z_x^{(-i)}) \]  \hspace{1cm} (6.44)

This specification is necessary to secure that \( i \)th element itself is not involved in a permutation process but its associations with other values resulting from non-zero diagonal elements in \( V \) are maintained in the permutation process.

Now local \( S_i \) is defined by relationships between two matrices, the Extended Mantel Test again applies to yields equations for the expected value and variance (not that now \( n \) should be replaced by \((n-1)\) in the computation of necessary quantities).

\[
E(S_i) = p_u^{(i)} \cdot q_u^{(i)} + \frac{[1^T \cdot P^{(-i)} \cdot 1] - \text{tr}(P^{(-i)}) \cdot [1^T \cdot Q^{(-i)} \cdot 1] - \text{tr}(Q^{(-i)})}{(n-1)(n-2)} + \frac{\text{tr}(P^{(-i)}) \cdot \text{tr}(Q^{(-i)})}{n-1} \hspace{1cm} (6.45)
\]

When a spatial weights matrix with a zero diagonal, the expected value is given:

\[
E(S_i) = \frac{1^T \cdot (V^T \cdot V) \cdot 1}{(n-1)(n-2)} \cdot \left\{ (n-4) \cdot (v_i^{(2)} - v_i^{(1)}) \cdot z_{x_i}^2 + (n-3) \cdot (v_i^{(2)} - v_i^{(1)}) \cdot n \right\} \hspace{1cm} (6.46)
\]

The expected value with a spatial weights matrix with a non-zero diagonal takes a much more complicated form:
\[ E(S_i) = \frac{n}{1^r (V^r V)^r} \left\{ \frac{1}{(n-1)(n-2)} \right. \]
\[ \left[ 2 \cdot v_i^2 - (3n-4) \cdot v_i^{(i)} + n^2 \cdot v_i^{(i)} - 4 \cdot v_i \cdot v_i^{(i)} \cdot z_{xi} \right] \]
\[ - \left[ v_i^2 - (n-1) \cdot v_i^{(i)} + n \cdot v_i^{(i)} - 2 \cdot v_i \cdot v_i^{(i)} \right] n \]  

(6.47)

When \( W \) is applied, (6.46) is simplified to:

\[ E(L_i) = \left[ \frac{\left( 2 - \frac{n}{n_i} \right) \cdot z_{xi}^2}{(n-1)(n-2)} \right] - \left[ \frac{\left( 1 - \frac{n-1}{n_i} \right) \cdot n}{(n-1)(n-2)} \right] \]  

(6.48)

When \( W^* \) is applied, (6.47) is simplified to:

\[ E(L_i) = \left[ \frac{\left( \frac{n-1}{n_i} \right) \cdot \left( \frac{n}{n_i} - 2 \right) \cdot z_{xi}^2}{(n-1)(n-2)} \right] - \left[ \frac{\left( 1 - \frac{n+1}{n_i} + \frac{n^2}{n_i^2} \right) \cdot n}{(n-1)(n-2)} \right] \]  

(6.49)

In the same way, \( Q^{(-i)} \) for Lee's \( L_i \) is given when a spatial weights matrix with a zero-diagonal:

\[ Q^{(-i)} = z_x^{(-i)} \cdot (z_y^{(-i)})^T \]  

(6.50)

where \( z_x^{(-i)} \) and \( z_y^{(-i)} \) are column vectors derived from \( z_x \) and \( z_y \) by eliminating \( i \)th entries, and thus \( Q^{(-i)} \) is a matrix obtained by eliminating \( i \)th row and column from \( Q \). 

131
When a spatial weights matrix with a non-zero diagonal, however, $Q^{(\cdot)}$ takes a more complex form:

$$ Q^{(\cdot)} = z_x^{(\cdot)} \cdot (z_y^{(\cdot)})^\top + \text{diag}(z_x \cdot z_y^{(\cdot)}) + \text{diag}(z_y \cdot z_x^{(\cdot)}) $$  \hspace{1cm} (6.51)$$

where the $\text{diag}()$ operation transforms a vector to a diagonal matrix. This specification is necessary to insure that the $i$th elements themselves in $X$ and $Y$ are not involved in a permutation process but their associations with other values resulting from non-zero diagonal elements in $V$ are maintained in the permutation process.

Since $Q^{(\cdot)}$ is differently defined between a spatial weights matrix with a zero-diagonal and one with a non-zero diagonal, it is necessary to present two different equations for the expected value. When a spatial weights matrix with a zero diagonal, the expected value is given:

$$ E(L) = \frac{n}{1^T (V^T V) 1} \cdot \frac{(2 \cdot \nu_k^2 - n \cdot \nu_i^{(2)}) \cdot z_x \cdot z_n - \left[ \nu_i^2 - (n-1) \cdot \nu_i^{(2)} \right] \cdot n \cdot r_{xy}}{(n-1)(n-2)} $$  \hspace{1cm} (6.52)$$

The expected value with a spatial weights matrix with a non-zero diagonal takes a much more complicated form:
When $W$ is applied, (6.52) is simplified to:

$$E(L_i) = \frac{1}{\binom{n}{3} \binom{n-2}{3}} \left\{ 
\left[ 2 \cdot v_i^2 - (3n-4) \cdot v_i^{(3)} + n^2 \cdot v_{i-2}^2 - 4 \cdot v_i \cdot v_{i-2} \right] \cdot z_{x_i} \cdot z_{y_i} 
- \left[ v_i^2 - (n-1) \cdot v_i^{(3)} + n^2 \cdot v_{i-2}^2 + 2 \cdot v_i \cdot v_{i-2} \right] n \cdot r_{x,y_i} 
\right\} \quad (6.53)$$

$$E(L_i) = \left[ \left( 2 - \frac{n}{n_i} \right) \cdot z_{x_i} \cdot z_{y_i} \right] - \left[ \left( 1 - \frac{n-1}{n_i} \right) \cdot n \cdot r_{x,y_i} \right] \quad (6.54)$$

When $W^*$ is applied, (6.53) is simplified to:

$$E(L_i) = \frac{\left[ \left( \frac{n}{n_i} - 1 \right) \left( \frac{n}{n_i} - 2 \right) \cdot z_{x_i} \cdot z_{y_i} \right] - \left[ \left( 1 - \frac{n+1}{n_i} \right) + \frac{n}{n_i^2} \right] \cdot n \cdot r_{x,y_i} }{(n-1)(n-2)} \quad (6.55)$$
CHAPTER 7

ESDA TECHNIQUES USING LOCAL SPATIAL ASSOCIATION MEASURES

This chapter proposes some ESDA techniques using SAMs and illustrate them with a hypothetical data set. I design two different spatial patterns sharing identical mean (1.838) and variance (0.514), and their Pearson’s correlation coefficient is 0.422 (Figure 7.1). Pattern A is used for univariate SAMs, and the relation between two patterns is utilized for bivariate SAMs. A row-standardized version (W) of a binary connectivity matrix (C) is used for Moran’s $I$, and Geary’s $c$, and a row-standardized version ($W^*$) of a binary connectivity matrix with 1s on its diagonal ($C^*$) is used for Lee’s $S_i$ and $L_i$.

7.1 Univariate ESDA

7.1.1 ESDA using local Moran’s $I_i$

If a research objective is to detect spatial clusters and spatial outliers, ESDA associated with local Moran’s $I_i$ can provide most effective techniques. A Moran scatterplot map is constructed by placing z-scores of a variable on the x-axis and
Figure 7.1: Hypothetical spatial patterns
z-scores of its spatial lag vector (note that elements in the spatial lag vector are
standardized by mean and standard deviation of the original vector, not by ones of the
spatial lag vector) on the y-axis. Four quadrants in Moran scatterplot denote different
spatial associations: upper right quadrant for high-high associations; lower left quadrant
for low-low associations; upper left quadrant for low-high associations; and the lower
right quadrant for high-low association (Figure 7.2-(a)). This graphical technique is
useful in exploring outliers or leverage points (Anselin 1996). Moran scatterplot map is a
nominal or categorical map corresponding to four quadrants in Moran scatterplot. A
Moran scatterplot map may give some ideas about possible spatial regimes. The most
useful mapping technique associated with local Moran's $I_i$ is the Moran significance map
(Anselin 2000). It displays observations with significantly high or low Moran's $I$s with
classes in a Moran scatterplot (Figure 7.2-(b)). Thus, significant spatial clusters, both
high and low, are easily detected, and spatial outliers, ones surrounded by dissimilar
neighbors, may be identified. It should be noted that the significance testing for local
SAMs is far from straightforward as mentioned in Chapter 2.

Among areas showing up in a Moran significance map, some can be identified as
spatial clusters if they belong to the first quadrant (hot spots) or the third quadrant (cold
spots). In contrast, areas belonging to the second or fourth quadrants can be recognized
as spatial outliers. In addition, when various significance levels apply to Moran
significance maps, as attempted by Tiefelsdorf (1998), Moran probability map can be
created.
If a research objective is to detect locales with a significant homogeneity, a *Geary significance map* can be used (Figure 7.2-(c)). A significance testing method selects only areas with a significantly high level of internal homogeneity. More correctly, an area that is significantly similar to its neighbors will be detected. This best conveys a concept of *spatially varying variances*. It should be noted that Geary significance map does not detect spatial clusters since it is indifferent to a value itself in a reference area. Significant areas may include not only spatial clusters (hot spots and cold spots), but also clusters of close-to-mean values. In Figure 7.2-(c), #8 area showing up significant is located in the center of a group of mean-values.

If a research objective is to detect spatial clusters without heavily depending upon reference areas, a *local-S significance map* can be utilized (Figure 7.2-(d)). Since, as discussed in Chapter 3, a local $S_i$ is computed by squaring a *spatially smoothed z-score* at a locale, it can avoid the tyranny of a reference area. Further, it can make a distinction between hot spots and cold spots by taking into account a sign of a z-score even though it is finally squared to yield a local $S_i$. Unlike in Moran and Geary significance maps, #30 showed up as a significant cold spot in Figure 7.2-(d).

When various significance levels apply to Geary significance map and local-S significance map, *Geary probability map* and *local-S probability map* can be created. In both maps, areas with higher $p$-values are expected to surround ones with lower $p$-values, if spatial processes leading to spatial clustering are dominant.
Figure 7.2: Univariate ESDA techniques
7.2 Bivariate ESDA

7.2.1 Local-r and local-L maps

A *local-L map* is a simple choropleth map with local Lee’s $L$,s (Figure 7.2-(b)).

Since the average of all the local $L$s equals the corresponding global $L$, a local $L_i$ can be interpreted as an area’s relative contribution to the global trend. A *local-r map* drawn with local Pearson’s $r$ according to equations (4.15) and (4.16) in Chapter 4 has the same properties (Figure 7.3-(a)). By comparing them, rationales of a local spatial correlation will be effectively illustrated.

From a local-r map and local-L map, one can assess the presence of bivariate spatial dependence that similar bivariate associations are spatially clustered. In both maps, areas with negative values indicate negative correlations when a global correlation is positive. Since a Lee’s $L_i$ captures not only the numeric similarity within a pair but also its relationships with neighboring pairs, a Local-L map appears to be a smoothed version of a Local-r map. Thus, the variance of a local-L map is much less than that of a local-r map. For exploration purposes, however, a local-L map provides a more distinctive spatial pattern. However, it should be noted that those maps do not distinguish either between high-high and low-low correlations or between high-low and low-high correlations.

7.2.2 Local-r and Local-L scatterplots, and Local-r and local-L scatterplot maps

In the same way as in local Moran’s $I$, a *local-r scatterplot* is constructed by setting $z_x$ on the x-axis and $z_y$ on the y-axis, and a *local-L scatterplot* is obtained by
Figure 7.3: Bivariate ESDA techniques
utilizing two vectors of spatially smoothed z-scores ($\bar{z}_x$ and $\bar{z}_y$) according to equation (4.14) in Chapter 4. While the slope line in Local-r scatterplot is equal to the global Pearson's $r$ as in a Moran scatterplot, the property does not hold for Local-L scatterplot. As expected, variances in both axes are much narrower in a local-L scatterplot than in a local-r scatterplot.

A more important aspect of those scatterplots is that they categorize local bivariate associations into four classes. On the scatterplots, the lower left and upper right quadrants indicate positive bivariate associations: the former for low-low associations and the latter for high-high associations. In contrast, the upper left and lower right quadrants indicate negative associations: the former for low-high associations and the latter for high-low associations. When areas are referenced by their quadrant locations, categorical maps can be created: *local-r scatterplot map* and *local-L scatterplot map* (Figure 7.3-(c)). Due to the presence of a highly positive global correlation, more areas are represented by low-low or high-high classes.

In the scatterplot map, areas belonging to the second and third quadrants correspond to areas with negative values in the local-L map (Figure 7.3-(b)). Again, a local-L scatterplot map can be seen as a spatially smoothed version of a local-r scatterplot map, with providing a more distinctive spatial pattern. Each class in a local-L map suggests a possible distinctive spatial regime.

This type of approach can be seen as a bivariate extension of an ESDA technique often called 'geographical brushing' (Monmonier 1989) or 'spatial windowing' (Fotheringham and Charlton 1994) where a map window is connected to a scatterplot.
such that any selection in the scatterplot makes subsequent indication on the map. Good examples of the technique in bivariate situations can be found in Haining et al. (2000a).

7.2.3 Local-L significance and probability maps

A local-L significance map is created when a local-L scatterplot map is combined to a significance testing (Figure 7.3-(d)). When an overall Pearson’s $r$ is high, most significant areas correspond to spatial clusters, either high-high associations or low-low associations. They may be called respectively bivariate hot spots and bivariate cold spots. Significant areas belonging to the second and fourth quadrants represent spatial clusters of observations that are against the general trend. If overall correlation is negative, however, those quadrants become representative of the general trend.

When various significance levels apply to a local-L significance map, a local-L probability map may be created. When bivariate spatial dependence is present, higher $p$-value areas are expected to surround lower $p$-value areas, resulting in a probability surface.

7.2.4 A discussion

Table 7.1 summarizes SAM-based ESDA techniques presented above and tasks. There are some issues unsolved about this approach. First, as Lee (2001c) points out, sampling distributions of Lee’s $L_i$ reveal significant levels of skewness and kurtosis. Thus, a more sophisticated inferential test should be devised by including higher moments, even though an intrinsic purpose of the measure is not confirmatory but
<table>
<thead>
<tr>
<th>Dimension</th>
<th>Measures</th>
<th>ESDA Techniques</th>
<th>Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate</td>
<td>Local Moran’s $I_t$</td>
<td>• Moran Scatterplot</td>
<td>• Spatial Outliers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Moran Scatterplot Map</td>
<td>• Spatial Clusters</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Moran Significance Map</td>
<td>• Spatial Regimes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Moran Probability Map</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Local Geary’s $c_t$</td>
<td>• Geary Significance Map</td>
<td>• Local Homogeneity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Geary Probability Map</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Local Lee’s $S_t$</td>
<td>• Local-S Significance Map</td>
<td>• Spatial Clusters</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Local-S Probability Map</td>
<td>• Spatial Regimes</td>
</tr>
<tr>
<td>Bivariate</td>
<td>Local Pearson’s $r_t$</td>
<td>• Local-r Scatterplot</td>
<td>• Local Bivariate Association</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Local-r Scatterplot Map</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Local Lee’s $L_t$</td>
<td>• Local-L Scatterplot</td>
<td>• Bivariate Spatial Clusters</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Local-L Scatterplot Map</td>
<td>• Bivariate Spatial Regimes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Local-L Significance Map</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Local-L Probability Map</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.1: SAM-based ESDA techniques and tasks
exploratory. Second, as Tiefelsdorf (2001) points out, the measure should be extended to
the spatial co-patterning between two sets of regression residuals by embracing a
projection matrix that projects a dependent variable and disturbances into a residual space
that is orthogonal to a set of independent variables (Tiefelsdorf 2000:16). Thus, the use
of the measure is currently confined to a pair of random variables. Third, it is unclear
how $L_i$ responds to different spatial weights matrices. For example, there is no statistical
interpretation for unequal weights given to reference areas as seen in $W^*$ where it is
hardly expected that reference locations are given an identical weight.

In spite of those unsolved pitfalls, the approach presented here may be benefited
from its applicability to a wide range of geographical inquiries. An instant application
can be undertaken to compute local spatial segregation indices. Since residential
segregation intrinsically involves a certain level of spatial clustering (univariate spatial
dependence), a comparison between spatial distributions of two different racial/ethnic
groups should embrace the bivariate spatial dependence. Although some global spatial
segregation indices have been proposed in order to amend the aspatial nature of regular
indices of segregation (among others, Morrill 1991, Wardorf 1993, Wong 1993,
Chakravorty 1996, Lee and Culhane 1998), researchers have suffered from inability to
explore local variations in the spatial segregation, with just few exceptions (Morrill 1995,
Wong 1996). While a global Lee’s $L$ can provide a reliable measure for overall spatial
exclusion between two racial/ethnic groups with a statistical test, Lee’s local $L_i$ can
capture a local degree of residential segregation. Another field that may effectively
utilize the approach could be a comparison between two raster-based layers such as
remotely sensed imageries (for an example of utilizing a local univariate spatial

144
association measure in image processing, see Wulder and Boots 1998). The utilization of local Lee’s $L_i$ could provide a feasible way of conducting the bivariate image generalization by generating a spatially smoothed layer of correspondence/discrepancy between two layers.

The conceptualization and parameterization of *spatially varying correlations* or *localized spatial correlations* could profoundly impact the research practices in a whole gamut of analytical geographies. This proposes a new research framework where researchers should be able not only to calibrate the relationship between two geographical variables in a *spatial* way, but also to assess how much each locale is deviated from the global trend, again in a *spatial* way. Furthermore, an advance from this bivariate dimension to *spatially varying multivariate associations* will open a gate to a concept of *spatially varying causalities*. This might lead to a recovery of the *areal differentiation* tradition at the heart of geographical information sciences.
CHAPTER 8

SPATIO-TEMPORAL INCOME DYNAMICS OF
U.S. LABOR MARKET AREAS, 1969-1999

8.1 Theoretical underpinnings of regional income convergence

Even though the theme of spatio-temporal dynamics in regional economic performance is not new, a rich body of literature has recently been devoted to the topic. It seems that at least two factors are responsible for this trend. First, a profound wave of socio-economic restructuring, occurring especially in advanced societies, has directed researchers to its implications for regional economic fortunes. Second, the advent of the European Union not only as an international integration but also as inter-regional integration has revived interest in the versatility of regional development (Arbia 2001). In this context, spatio-temporal dynamics of income distribution or regional income convergence/divergence across regions has a focal point, even though various academic camps provide different approaches to the issue.
At least three distinct academic threads are involved.

Firstly, the so-called 'new growth theory' based on a reformulation of neoclassical growth models (the inverted-U hypothesis by Kuznet 1955; Solow 1956; Williamson 1965), 'endogenous growth theory' (According to Armstrong (1995a), Evans and Karras (1996), and de la Fuente (1997), Romer 1986; Lucas 1988; Romer 1990), and post-Keynesian traditions (according to Pons-Novell and Viladecans-Marsal (1999), Verdoorn 1949; Kardor 1966; 1975) have stimulated empirical works on the growth convergence issue (among others Barro and Sala-i-Martin 1991; 1995; for a review European Commission 1997; Button and Pentecost 1999). With some exceptions, this theory tends to underline a general trend towards equilibrium, which is evidenced by $\sigma$-convergence (a decrease of overall level of regional income inequality) and $\beta$-convergence (a negative relation between initial regional income levels and regional income growth rates in a 'growth regression').

Secondly, the California or Los Angeles School (Dear and Dishman 2001) inspired by the French Regulation School postulates the nature of the transition from Fordism to post-Fordism and formulates conditions for 'New Industrial Spaces' (among others, Storper and Scott 1986; Scott 1988a; 1988b; Storper and Walker 1989; Storper and Scott 1992; Storper 1997; Scott 1998). Even though the main focus of the School is on 'successful regions', rather than an overall picture regarding spatial restructuring, some empirical works on regional disparities are based on notions of the School (e.g., Dunford and Perrons 1994; Rodriguez-Pose 1994; Matthews 1996; Dunford 1996; 1997; Rodriguez-Pose 1999; Dunford and Smith 2000;). The post-Fordist spatial economic landscape implied by this line of theorization seems to be divergent rather than
accommodating flexible specialization dominate over old Fordist industrial regions and small- and medium-sized central places. In short, Fordism induces employment growth and income convergence, whereas post-Fordism is characterized by growth slowdown and income divergence (Dunford and Perrons 1994).

Thirdly, ‘new economic geography’ referring to works by economists (among others, Krugman 1991; Arthur 1994; Krugman 1995; Fujita et al. 1999) has significantly contributed not only to the topic of regional income convergence but also to economic geography and regional sciences in general (for reviews, Martin 1999; Fingleton 2001). This camp provides a new insight into regional income dynamics. Reduction in transport and transaction costs associated with increased integration (by way of globalization or certain forms of economic supranationalism) fuels spatial agglomeration and localization externalities, leading to income divergence among regions in terms of the increased specialization (Martin 1999; Martin and Tyler 2000; Martin 2001).

In general, the new growth theory accentuates convergence over divergence, while the Los Angeles School and new economic geography lean towards divergence over convergence. However, evidence is far from consistent. Empirics, even from the same camp, often report different stories. In the context of the US, for example, Evans and Karras (1996) and Sala-i-Martin (1996) find a consistent trend of convergence, while Quah (1996b) and Tsionas (2000; 2001a; 2001b) obtain evidence in the opposite direction.

I argue that inconsistency in empirics on regional income convergence results not just from different data sources or different spatial units, but from commonly used
methodologies themselves. Most of all, regional analyses should take advantage of recent advances in spatial data analysis including spatial econometrics and spatial statistics as already suggested (Nijkamp and Poot 1998; Griffith 1999; Fingleton 2000; 2001). In the context of spatio-temporal regional income dynamics, only a few utilize spatial analytical devices (Lopez-Bazo et al. 1999; Rey and Montouri 1999; Rey 2001). Accordingly I first point to common methodological drawbacks of current empirical studies, and second demonstrate how SAM-based ESDA techniques could make a significant improvement.

8.2 A critical review of empirical studies on regional income dynamics: recovery of spatiality

8.2.1 $\sigma$-convergence: numerical variance vs. spatial clustering

$\sigma$-convergence refers to the reduction of dispersions or variances in per capita incomes across regions over time, usually measured by the standard deviation or coefficient of variation of the regional income distribution (Barro and Sala-i-Martin 1991; Rey 2001). Sometimes, this type of convergence is called ‘strong convergence’, as opposed to ‘weak convergence’ that refers to $\beta$-convergence (Nijkamp and Poot 1998). This notion of convergence is deeply rooted in neo-classical growth theory (Kuznets 1955; Williamson 1965), and has been applied to numerous countries, summarized in Table 8.1. The results show a trend of long-term convergence in regional income distribution with some discrepancies. In the context of the US, five studies listed in Table 8.1 are based on state-level data and utilize measures of standard deviation or
<table>
<thead>
<tr>
<th>Spatial Units</th>
<th>Studies</th>
<th>Years</th>
<th>Indices*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe EU regions</td>
<td>Barro and Sala-i-Martin (1991)</td>
<td>1950-1985</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>Armstrong (1995c)</td>
<td>1975-1992</td>
<td>CV</td>
</tr>
<tr>
<td></td>
<td>Quah (1996a)</td>
<td>1980-1989</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>European Commission (1997)</td>
<td>1975-1993</td>
<td>SD &amp; GC</td>
</tr>
<tr>
<td></td>
<td>Dunford and Perron (1994)</td>
<td>1960-1989</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>de la Fuente (1997)</td>
<td>1870-1990</td>
<td>CV</td>
</tr>
<tr>
<td>EU countries</td>
<td>Sala-i-Martin (1996)</td>
<td>1950-1990</td>
<td>SD</td>
</tr>
<tr>
<td>France regions</td>
<td>Sala-i-Martin (1996)</td>
<td>1950-1990</td>
<td>SD</td>
</tr>
<tr>
<td>Germany regions</td>
<td>Sala-i-Martin (1996)</td>
<td>1950-1990</td>
<td>SD</td>
</tr>
<tr>
<td>Italy regions</td>
<td>Sala-i-Martin (1996)</td>
<td>1950-1990</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>Sala-i-Martin (1996)</td>
<td>1950-1990</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>Cuadrado-Roura et al. (1999)</td>
<td>1955-1995</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>Sala-i-Martin (1996)</td>
<td>1880-1990</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>Sum and Fogg (1999)</td>
<td>1939-1996</td>
<td>SD &amp; CV</td>
</tr>
<tr>
<td></td>
<td>Rey and Montouri (1999)</td>
<td>1929-1994</td>
<td>CV</td>
</tr>
<tr>
<td>Others World countries</td>
<td>de la Fuente (1997)</td>
<td>1960-1985</td>
<td>SD</td>
</tr>
<tr>
<td>OECD countries</td>
<td>de la Fuente (1997)</td>
<td>1960-1985</td>
<td>SD</td>
</tr>
<tr>
<td>Japan prefectures</td>
<td>Sala-i-Martin (1996)</td>
<td>1955-1990</td>
<td>SD</td>
</tr>
</tbody>
</table>

* SD: Standard Deviation; CV: Coefficient of Variation; ID: Index of Dissimilarity; GC: Gini Coefficient

Table 8.1: Empirical studies on σ-convergence
income distribution has been characterized by a succession of a decrease, 1880-1920, an increase, 1920-1930, a decrease, 1930-the mid-1970s, and an increase, the mid-1970s-1988. Sala-i-Martin’s later study (1996), with additional years, indicates a decrease in the early 1990s. Fan and Casetti (1994) document similar results, that is, a decrease, 1950-1980 and an increase, 1980-1989. Rey and Montouri (1999) also show the identical picture; a decrease up until 1980, an increase during the 1980s, and a decrease during the early 1990s. In UK, while most studies report a constant trend towards convergence, Chatterji and Dewhurst (1996) found that per capita GDP distribution across the UK regions has become more divergent.

Albeit the intuitive simplicity, $\sigma$-convergence has serious drawbacks. It does not provide insights into processes that may be driving the narrowing (or widening) of regional incomes. No information is provided regarding the relative movements of individual economies within the income distribution (Rey 2001:196). In other words, a diminishing standard deviation of incomes does not tell whether some poorer economies catch up with the richer economies faster than some others (Sala-i-Martin 1996; Kangasharju 1999; Tsionas 2000). More serious problems that this approach bears, however, revolve around its lack of spatial perspectives.

First, studies based on $\sigma$-convergence should be enlightened by findings in the modifiable areal unit problem (MAUP). Such measures as standard deviation and coefficient of variation are strongly influenced by the level of spatial aggregation. In general, variance tends to decrease as the level of spatial aggregation escalates. In other words, a set of larger spatial unit is inclined to display a smaller variance due to a
smoothing effect that outliers lose their peculiarities as spatial aggregation proceeds (Fotheringham and Wong 1991; Wong 1996). The magnitude and temporal trend of regional income convergence could vary depending on the spatial configuration of a study. Further, it may be unsustainable to compare the spatio-temporal trend among different countries each of which has a particular regionalization scheme (see Figure 6 in Sala-i-Martin (1996)).

Second, the numeric variance that σ-convergence is predicated on is immune to spatial clustering (Arbia 2001). As illustrated in Figure 3.1, totally different spatial patterns can be generated from a numeric vector, and they cannot be differentiated by variance. What this implies is that σ-convergence does not measure spatial convergence that belies what is implied by ‘regional convergence’. It is necessary, thus, for researchers to look into the spatio-temporal trend of spatial dependence in income distribution if they are to obtain a substantive understanding of what has occurred in reality. In this sense, Wheeler (2001) reports from spatial correlogram analyses based on the US county level that spatial dependence in regional income growth is well pronounced and spatial autocorrelation drops off to zero over a distance of roughly 2000 miles, with a strong stability within 40 miles. More important aspects of spatial dependence include the presence of inter-regional interaction, co-dependence, or spillover effects in regional income distributions (Quah 1996a; Rey and Montouri 1999; Podriques-Pose 1999; Ying 2000; Martin 2001; Rey 2001). Quah (1996a:954) correctly contends that ‘physical location and geographical spillover matter more than do macro factors’. In this sense, a univariate SAM should be utilized to gauge spatial clustering not only for each year but also for growth rates between years. Surprisingly, only few papers
recognize the importance of spatial dependence in regional income distributions and utilize univariate SAMs such as Moran's $I$ and Geary's $c$ (European Commission 1997; Lopez-Bazo et al. 1999; Rey and Montouri 1999; Rodriguez-Pose 1999).

Third, global bivariate SAMs could help find breakpoints in the temporal trend. When global bivariate SAMs for consecutive years could reveal when significant spatial restructurings occur. This is expected to correspond to what differentials in global univariate SAMs between consecutive years may imply.

Fourth, $\sigma$-convergence is global in nature such that it focuses only on an average aspect, or trend, ignoring the possible spatial heterogeneity often pronounced in the spatial organization of income (Rey and Montouri 1999; Ying 2000). This point also applies to global SAMs. For each year, hot spots and cold spots can be identified. A series of spatial distributions of income over years could reveal a trajectory of spatial restructuring over time. In this sense, local univariate SAMs should be utilized as attempted (Lopez-Bazo et al. 1999; Rey and Montouri 1999; Ying 2000; Rey 2001).

Fifth, a well-designed spatial unit, other than arbitrary ones such as states and census regions, is needed. A viable spatial unit could be a regional labor market area where a vast majority of people live and work, and an intra-regional functional integration is distinctive to a certain degree. In the context of the US, Lee (1999) investigates existing regionalization schemes satisfying the notion of regional labor market and identifies 17 county-based functional regions. Among them, CZ (Commuting Zone), BTA (Basic Trading Area), LMA (Labor Market Area), and BEA (Bureau of Economic Analysis Economic Area) could be reliable candidates for research on US
regional income convergence. Use of regional labor market areas is expected to reveal the versatile nature of regional income disparities more efficiently and thoroughly.

8.2.2 β-convergence: numerical correlation vs. spatial co-patterning

β-convergence was introduced by Barro (1991), and Barro and Sala-i-Martin (1991) based on the neo-classical growth theory to capture the catch-up hypothesis that poorer regional economies grow faster than richer regional economies. There will be β-convergence if a negative relation is found between the initial level of income and the growth rate of per capita income (sometimes referred to as 'regression to the mean' or 'mean reversion') (Sala-i-Martin 1996:1327). This ‘weak convergence’ (Nijkamp and Poot 1998:26), as opposed to ‘strong convergence’ of α, often takes a regression form as:

$$\frac{1}{k} \cdot \ln \left( \frac{y_{i,t+k}}{y_{i,t}} \right) = \alpha - \beta \cdot \ln(y_{i,t}) + \varepsilon_{i,t}$$  \hspace{1cm} (8.1)

where $k$ is a year-differential, $y_{i,t}$ is income level in region $i$ at a starting year, $y_{i,t+k}$ is income level in region $i$ at an ending year. However, a more complicated form has been preferred in empirical studies that is given (Sala-i-Martin 1996; Nijkamp and Poot 1998):

$$\frac{1}{k} \cdot \ln \left( \frac{y_{i,t+k}}{y_{i,t}} \right) = a - \left( \frac{1-e^{-\beta k}}{k} \right) \ln(y_{i,t}) + \varepsilon_{i,t}$$  \hspace{1cm} (8.2)
From equation 8.1, an estimated value of $\beta$ is a slope coefficient in a regression of regional income growth rates on initial regional income level (Nijkamp and Poot 1998; Martin 2001), which constitutes a reason why work based on $\beta$-convergence has been called 'growth regression approach' (Martin 2001:62). More often, however, the parameter of $\beta$ is interpreted as the speed at which economies approach their own steady states or poorer regions catch up with richer ones (Sala-i-Martin 1996; Kangasharju 1999). When some other shock variables, such as industrial mix and regional dummies, are included in the equation, it calibrates 'conditional convergence' as opposed to 'absolute convergence' attributing to the original specification (Barro and Sala-i-Martin 1991; Armstrong 1995; Sala-i-Martin 1996; Dickey 2001). From both equations, a positive value of $\beta$ indicates economic convergence across regions.

$\beta$-convergence is usually preferred over $\sigma$-convergence in empirical studies, because the former conveys more information about regional income convergence than the latter. First, $\beta$-convergence is a necessary condition for $\sigma$-convergence (Nijkamp and Poot 1998). Without $\beta$-convergence, $\sigma$-convergence won’t happen. In other words, a substantial change in the ranking of regions in economic performance could happen without being captured by $\sigma$-convergence. Thus, $\beta$-convergence does not imply $\sigma$-convergence (Barro and Sala-i-Martin 1991; Sala-i-Martin 1996; Nijkamp and Poot 1998; Kangasharju 1999; Tsionas 2000). However, there are disagreements: the convergence rate does not mean that a poorer region catches up with a richer region at that rate (Tsionas 2000); the convergence may be unrelated to, or uninformative for, the dynamics of economic growth (Quah 1996).
convergence. Almost all studies listed report consistent convergence at a $\beta$ of about 0.02, which means that regions, wherever they are, tend to converge at a speed of approximately 2% per year (Barro and Sala-i-Martin 1991; Armstrong 1995; Sala-i-Martin 1996). For example, $\beta$-coefficient for 48 US states was estimated at 0.017 during 1880-1990 (Sala-i-Martin 1996). This striking coincidence across numerous countries or supranational entities such as the European Union arguably verifies the virtues of neo-classical growth theory. Although the original formulation is based on an assumption of closed Solow economies (Blanchard 1991:159), additional considerations such as labor mobility, capital mobility, and technology transfer adjust the theorem equally working across open economies such as countries and supranational regimes (Barro and Sala-i-Martin 1991; Blanchard 1991; Armstrong 1995).

The alleged myth of regional income convergence, however, has been challenged by theoretical arguments and empirics. Quah (1996b:1355) contends that “uniformity is due to something relatively uninteresting, namely, the statistical implications of a unit root in the time-series data”. The unit root problem refers to non-stationarity in residuals of a time-series OLS regression (Kennedy1998:268-269). In the same vein, Martin (2001:62) argues that “the growth regression approach has an inbuilt bias towards identifying convergence, so that the results may even over-estimate what little convergence has occurred.” Tsionas (2000) reports that there is a positive relationship between initial levels of regional income and income growth rates. A finding by most ardent advocates for the convergence thesis (Barro and Sala-i-Martin 1991) says that $\beta$-
### Table 8.2: Empirical studies on β-convergence

<table>
<thead>
<tr>
<th>Spatial Units</th>
<th>Studies</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Armstrong (1995c)</td>
<td>1975-1992</td>
</tr>
<tr>
<td></td>
<td>France regions</td>
<td>1950-1990</td>
</tr>
<tr>
<td></td>
<td>Germany regions</td>
<td>1950-1990</td>
</tr>
<tr>
<td></td>
<td>Italy regions</td>
<td>1950-1990</td>
</tr>
<tr>
<td></td>
<td>Spain regions</td>
<td>1955-1991</td>
</tr>
<tr>
<td></td>
<td>Finland regions</td>
<td>1974-1993</td>
</tr>
<tr>
<td></td>
<td>Cuadrado-Roura et al. (1999)</td>
<td>1955-1993</td>
</tr>
<tr>
<td></td>
<td>Sala-i-Martin (1996)</td>
<td>1880-1990</td>
</tr>
<tr>
<td></td>
<td>Sum and Fogg (1999)</td>
<td>1939-1996</td>
</tr>
<tr>
<td></td>
<td>Andres et al. (1996)</td>
<td>1960-1990</td>
</tr>
<tr>
<td></td>
<td>Sala-i-Martin (1996)</td>
<td>1939-1996</td>
</tr>
<tr>
<td>Others</td>
<td>OECD countries</td>
<td>1960-1990</td>
</tr>
<tr>
<td>Canada</td>
<td>provinces</td>
<td>1960-1990</td>
</tr>
<tr>
<td>Japan</td>
<td>prefectures</td>
<td>1960-1990</td>
</tr>
</tbody>
</table>
coefficient turns negative during the 1980s implying a possible divergence in recent years.

Again, I would argue that studies based on β-convergence should be edified by findings from spatial data analysis.

First, as far as the regression equation is calibrated based on the OLS algorithm, the approach is obviously subject to problems of spatially autocorrelated errors. Given various types of spatial interactions among regions, adjacent regions tend to show a similar trend in economic performance, which will be reflected in regression residuals. Quah (1996:954) correctly points out that “no region can be studied in isolation independently of others”. When a significant spatial autocorrelation is present in residuals, significance tests for coefficients may be flawed even though coefficients themselves are still unbiased (Anselin and Griffith 1988; Fotheringham and Rogerson 1993). Thus, the regression equation in 8.1 and 8.2 should be calibrated by a spatial autoregressive model (Armstrong 1995b; Molho 1995; Bernat 1996; European Commission 1997; Mencken 1998; Buettner 1999; Fingleton 1999; Rey and Montouri 1999; Pons-Novell and Viladecans-Marsal 1999; Martin 2001). When this problem is associated with other statistical symptoms such as non-normality, structural instability, and misspecification (Tsionas 2000), the whole research becomes unsustainable.

Second, as Martin (2001:62) correctly points out, the β-convergence approach is based on an unreliable assumption that “the underlying convergence process is identical across all regions, whereas in reality it is may well vary from region to region, or between different types or groups of regions”. This resonates with Quah’s argument (1996:954) that “regression-based approaches, averaging across either cross-section or time series
dimensions, are not useful...such methods construct a representative, and cannot provide a picture of how the entire cross-section distribution evolves”. This issue of spatial heterogeneity can be addressed by some other multivariate spatial statistical techniques such as expansion method (Casetti and Jones 1987), expanded rank-size function (Fan and Casetti 1994; Lopez-Bazo et al. 1999), Markov chain matrix (Quah 1993; Quah 1996b; Fingleton 1997; 1999; Rey 2001), and geographically weighted regression (GWR) (among others, Brunsdon et al. 1996; 1998a, 1998b; 1999a; 1999b; Fotheringham et al. 1997a; 1997b; 1998; 2000).

Third, as far as ‘absolute convergence’ without any other additional variables is concerned, β-convergence is nothing but a correlation between initial income levels and income growth rates. A β coefficient from an OLS regression is directly related to correlation between two variables. When a bivariate relation between initial income levels and income growth rates are spatially clustered, a global bivariate SAM should replace aspatial correlation measure such as Pearson’s r. Furthermore, there is a good reason to believe that the averaged correlation coefficient does not apply to the whole study area. Rather, local correlations may be highly heterogeneous. In the context of income convergence, some initially poor regions may have accomplished a certain level of catch-up, whereas some others may still fall behind. This spatial heterogeneity can only be tackled by local bivariate SAMs and related ESDA techniques illustrated in Chapter 7.
8.3 Spatio-temporal income dynamics across the US labor market areas

8.3.1 Research design

Data sources for regional incomes are dictated by spatial aggregation level to a large extent. In general, larger spatial units, such as census regions and states, provide more affluent data sources. Since those spatial units are often arbitrary regions, rather than functional regions, their use prevents researchers from obtaining a 'ground-level' reality. The county as a spatial unit is not a good choice either, simply because a substantive proportion of labor force commutes across county boundaries. Thus, a regionalization scheme aggregating counties into functional regions should be utilized.

In this study, main spatial units are labor market areas (LMAs) (Killian and Tolbert 1993; Tolbert and Sizer 1996). Their definition is first based on a commuting flow matrix among counties, and a hierarchical cluster analysis aggregates 3,141 counties into 741 commuting zones (CZs). The CZs are then aggregated into 394 LMAs in terms of a minimum population requirement (100,000) and inter-CZ commuting flows (Tolbert and Sizer 1996). States and another function regions, BEAs (Bureau of Economic Analysis Economic Areas), are utilized for supplementary purposes along with CZs. BEAs are also based on commuting flows among counties, but the metropolitan status of a county plays a crucial role (Johnson 1995). The BEA scheme divides the whole US into 172 units. Since the study focuses on the conterminous US, the number of spatial units in each scheme is reduced. Finally, 48 states plus the District of Columbia, 170 BEAs, 391 LMAs, and 722 CZs are utilized. CZs are completely nested in LMAs, and any of BEAs, LMAs, and CZs is not restricted by state boundaries.
Per capita personal income data are used for this empirical study. The data sets have been collected and maintained by the Bureau of Economic Analysis, and are available via REIS (Regional Economic Information System) at the county level from 1969 to 1999. Since both county population and personal income aggregate at a county are provided, per capita personal income at any spatial aggregates is easily computed as long as an applied regionalization scheme is conducted at the county level. An additional consideration has been given in each spatial aggregation process in order to deal with county boundary changes over time.

Even though most studies utilize per capita personal income, it should be pointed out that the data set in this research has some drawbacks. First, it does not take living costs into account. Thus, there might be a gap between nominal income levels and real income levels adjusted for purchasing power. Second, quality of data could be improved by controlling some confounding factors, such as population composition in terms of age and race. Third, regression model with the data set is inherently heteroscedastic because population sizes differ among LMAs.

Effective use of ESDA techniques presented in Chapter 7 is crucial for this application part. Each technique is utilized to reveal a particular aspect of spatio-temporal income dynamics in the US. The entire thirty years are divided into three sub-periods, 1969-1979, 1979-1989, and 1989-1999, and four years, 1969, 1979, 1989, and 1999, are utilized as benchmarks to provide particular snapshots, allowing for tracking spatio-temporal evolution.

This empirical study is largely divided into five parts.
First, spatial distributions of per capita income across LMAs are explored and significant spatial clusters are detected for the four different years. Quantile maps allow for an effective comparison among the four different spatial patterns. Local-S significance and Geary significance maps identify significant spatial clusters for each year and a comparison of different years is expected to reveal temporal heterogeneity in spatial dependence of regional income distributions.

Second, temporal trends in regional income distribution over time are examined. Global Pearson's $r$ as well as Lee's $L$ are utilized to capture the degree of continuity/change over time. Bivariate ESDA techniques, including local-r scatterplot, local-r and local-L scatterplot maps, and local-L significance map, reveal spatial heterogeneity in temporal change during 1969-1999 across LMAs.

Third, $\sigma$-convergence is examined in conjunction with spatial clustering. Temporal trend of coefficients of variation are computed for various regionalization schemes, including states, BEAs, LMAs, and CZs, to investigate effects of MAUP. Global Moran's $I$ is utilized to gauge the degree of spatial clustering. The relationship between coefficients of variation and Moran's $I$s is investigated for LMAs, and an attempt is undertaken to provide a feasible explanation of the relationship. Distributions of spatial outliers detected by Moran significance maps are expected to provide a new insight into the relationship between numerical variance and spatial clustering.

Four, $\beta$-convergence, a negative relationship between an initial income level and income growth rate between years, is critically evaluated. Global Pearson's $r$ and Lee's $L$ are computed for different sub-periods and for different regionalization schemes. Spatial autocorrelation in OLS residuals is assessed and SAR (simultaneous autoregressive)
models are introduced to alleviate the problem of spatially autocorrelated errors. Spatial patterns resulting from a decomposition of a SAL model effectively demonstrate necessities of using spatial autoregressive models when spatial autocorrelation in residuals is significant.

Five, spatial aspects of β-convergence are investigated. Local-r and local-L scatterplot maps and local-L significance maps are utilized to explore spatial heterogeneity in β-convergence. The general trend of a negative relationship between initial income levels and income growth rates is spatially evaluated; some following the trend; but others not. A geographically weighted regression (GWR) model will be used for a supplementary purpose.

8.3.2 Regional income distribution and identification of spatial clusters

Figure 8.1 shows spatial patterns of per capita personal incomes across LMAs for four benchmark years, 1969, 1979, 1989, and 1999. For comparison, a quantile classification scheme is applied to each map. One finding is that the spatial distribution of per capita income has not significantly changed over the 30 years. The persistent spatial structure involves higher income levels in the Megalopolis, the Midwest, and western coastal regions, southern Florida, and lower income levels in areas from the northwest Mountain region to southern Texas, most areas of the South, the northwestern part of the Midwest, and the Ohio River Valley (ORV) region (Brown et al. 1996; 1999; Brown 1999; Brown et al. 2001). Table 8.3 lists the top and bottom ten LMAs in terms of per capita income for four years, 1969, 1979, 1989, and 1999. Spatio-temporal continuation becomes more obvious from the list. Five out of top 10 LMAs in 1969
Figure 8.1: Spatial distributions of per capita personal income across the US LMAs
<table>
<thead>
<tr>
<th>Years</th>
<th>Top 10</th>
<th>Bottom 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1969</td>
<td>San Francisco, CA</td>
<td>Not distinguishable city, KY</td>
</tr>
<tr>
<td></td>
<td>New York, NY</td>
<td>Brownsville, TX</td>
</tr>
<tr>
<td></td>
<td>Bridgeport, CT</td>
<td>Laredo, TX</td>
</tr>
<tr>
<td></td>
<td>Chicago, IL</td>
<td>Greenville, MS</td>
</tr>
<tr>
<td></td>
<td>Newark, NJ</td>
<td>Not distinguishable city, KY</td>
</tr>
<tr>
<td></td>
<td>San Jose, CA</td>
<td>Clarksdale, MS</td>
</tr>
<tr>
<td></td>
<td>Wilmington, DE</td>
<td>Richmond, KY</td>
</tr>
<tr>
<td></td>
<td>Los Angeles, CA</td>
<td>McComb, MS</td>
</tr>
<tr>
<td></td>
<td>Detroit, MI</td>
<td>Somerset, KY</td>
</tr>
<tr>
<td></td>
<td>Reno, NV</td>
<td>Tuscaloosa, AL</td>
</tr>
<tr>
<td>1979</td>
<td>San Francisco, CA</td>
<td>Laredo, TX</td>
</tr>
<tr>
<td></td>
<td>Casper, WY</td>
<td>Brownsville, TX</td>
</tr>
<tr>
<td></td>
<td>San Jose, CA</td>
<td>Somerset, KY</td>
</tr>
<tr>
<td></td>
<td>Reno, NV</td>
<td>Gallup, NM</td>
</tr>
<tr>
<td></td>
<td>Chicago, IL</td>
<td>McComb, MS</td>
</tr>
<tr>
<td></td>
<td>Houston, TX</td>
<td>Richmond, KY</td>
</tr>
<tr>
<td></td>
<td>Newark, NJ</td>
<td>Hinesville, GA</td>
</tr>
<tr>
<td></td>
<td>Los Angeles, CA</td>
<td>Clarksdale, MS</td>
</tr>
<tr>
<td></td>
<td>Bridgeport, CT</td>
<td>Roanoke Rapids, NC</td>
</tr>
<tr>
<td></td>
<td>New York, NY</td>
<td>Not distinguishable city, MO</td>
</tr>
<tr>
<td>1989</td>
<td>West Palm Beach, FL</td>
<td>Laredo, TX</td>
</tr>
<tr>
<td></td>
<td>Bridgeport, CT</td>
<td>Brownsville, TX</td>
</tr>
<tr>
<td></td>
<td>San Francisco, CA</td>
<td>Gallup, NM</td>
</tr>
<tr>
<td></td>
<td>Newark, NJ</td>
<td>Not distinguishable city, KY</td>
</tr>
<tr>
<td></td>
<td>New York, NY</td>
<td>McComb, MS</td>
</tr>
<tr>
<td></td>
<td>Baltimore, MD</td>
<td>Greenville, MS</td>
</tr>
<tr>
<td></td>
<td>Boston, MA</td>
<td>Richmond, KY</td>
</tr>
<tr>
<td></td>
<td>San Jose, CA</td>
<td>Somerset, KY</td>
</tr>
<tr>
<td></td>
<td>Brick Township, NJ</td>
<td>West Memphis, AR</td>
</tr>
<tr>
<td></td>
<td>Sarasota, FL</td>
<td>Provo, UT</td>
</tr>
<tr>
<td>1999</td>
<td>San Jose, CA</td>
<td>Brownsville, TX</td>
</tr>
<tr>
<td></td>
<td>San Francisco, CA</td>
<td>Gallup, NM</td>
</tr>
<tr>
<td></td>
<td>Bridgeport, CT</td>
<td>Laredo, TX</td>
</tr>
<tr>
<td></td>
<td>New York, NY</td>
<td>Not distinguishable city, KY</td>
</tr>
<tr>
<td></td>
<td>Newark, NJ</td>
<td>Somerset, KY</td>
</tr>
<tr>
<td></td>
<td>West Palm Beach, FL</td>
<td>Greenville, MS</td>
</tr>
<tr>
<td></td>
<td>Boston, MA</td>
<td>El Paso, TX</td>
</tr>
<tr>
<td></td>
<td>Denver, CO</td>
<td>Not distinguishable city, KY</td>
</tr>
<tr>
<td></td>
<td>Baltimore, MD</td>
<td>McComb, MS</td>
</tr>
<tr>
<td></td>
<td>Minneapolis, MN</td>
<td>Yuma, AZ</td>
</tr>
</tbody>
</table>

(Cities are largest ones in LMAs)

Table 8.3: Top and bottom 10 LMAs in per capita personal income, 1969, 1979, 1989, and 1999
occupy the top-five spots in 1999, and 7 out of bottom 10 LMAs in 1969 have not lost their seats in the 1999 bottom 10 list. The top 10 list for 1999 shows that three LMAs centered on Boston, Denver, and Minneapolis emerge for the first time which have been regarded as cities successfully adjusting to new economic conditions in the post-Fordist era.

In spite of the continuation of the dominant spatial morphemics, several spatial shifts are also detected. First, the traditional industrial cores in the Midwest have been spatially disintegrated. Especially areas centered on Detroit have lost much of its internal integrity. Second, some areas in the South, particularly the Pediment, have experienced relatively higher income growth. Those areas include Winston-Salem, Charlotte, and Raleigh in North Carolina, Birmingham in Alabama, and Austin in Texas.

This finding well corresponds to Brown’s thesis of ‘continuity amidst restructuring’ (Brown 1999; Brown et al. 2001). He contends (Brown et al. 2001) that “while many types of change occurred through the Fordist/Post-Fordist transition, they are not necessarily manifest in terms of spatial variation over time ... all regions declined early in this transition and, apparently, more or less to the same degree ... yet, most of the formerly dominant regions rebounded, albeit with a different economic structure (e.g., service or high-technology industries rather than Fordist-type traditional industry)”. Even though theoretical underpinnings seeking to explain ‘spatial fixity’ over ‘spatial plasticity’ in economic performance have been proposed, empirical studies that might evidence the theoretical notions are very few. Among others, Melachroinos and Spence (1999) show how ‘sunk costs’ function as a change-inhibiting factor in regional economic performance across Greece prefectures from 1984 to 1993.
To identify spatial clusters for each pattern, I utilize the local-S significance map technique as illustrated in Figure 7.2. Since local $S_i$ is relatively liberated from the tyranny of reference areas, it works better than local Moran’s $I_i$ in identifying spatial clusters, as discussed in Table 3.2. Per capita income is first transformed by natural logarithm, and a one-tailed test at the 95% confidence level based on the conditional randomization presented in Chapter 6 is utilized. Figure 8.2 clearly shows the spatio-temporal dynamics in regional income distribution. Two crucial observations are made. First, the most dramatic spatial change had occurred in the 1980s. Second, internal integrity of spatial clusters has been substantively eroded over 30 years.

The 1969 map displays five distinctive spatial clusters: the Megalopolis, the Midwest industrial belt, the Pacific as richer regimes, and the ORV region and the South as poorer regimes. In 1999, the spatial regimes are still observed, but their internal integrity has significantly been eroded: the Midwest industrial belt has been largely disintegrated; the Pacific has shrunk to the San Jose-San Francisco area and the Seattle area; the poor parts of the South are now confined to the Lower Mississippi; spatial clustering is only found around Chicago area within the Midwest industrial belt. In contrast, the Megalopolis and the southern Florida have maintained regional homogeneity as higher income areas. Areas in Colorado centered on Denver, Colorado Springs, and Fort Collins have emerged as a new hot spot during the 1990s.

The 1979 map in Figure 8.2 show that northwestern mountain areas centered on Casper and Laramie in Wyoming appeared as significant higher income clusters. It also displays that the poor South had expanded to east and the Megalopolis had shrunk during 1970s. The trend, however, substantively reversed during the 1980s: the
Figure 8.2: Local-S significance maps: spatial clusters in per capita personal income across the US LMAs
northwestern high income centers disappeared; the Megalopolis had expanded; the poor South had been confined to the Lower Mississippi. The most notable change in during the 1990s seen from the 1999 map in Figure 8.2 is the shrinking California.

Figure 8.3 examines different aspects of spatial dependence in regional income distribution. As discussed in Chapter 3 and Chapter 7, local Geary's $c_i$ is better at assessing local homogeneity in comparison with local Lee's $S_i$ and Moran's $I_i$ that are better at detecting spatial clusters. Simply, local Geary's $c_i$ captures local variance. Spatial clusters do not necessarily mean that there is little variance within them; a high level of internal heterogeneity within a spatial cluster identified by local $S_i$ or $I_i$ is often observed. Geary significance maps in Figure 8.3 reveal that there are substantive internal variance within the Pacific region in 1969 and 1979 detected as spatial clusters in Figure 8.2, and suggest that rich clusters in 1989 and 1999 possess higher level of internal homogeneity than poor clusters. While the 1999 maps in Figure 8.2 and in Figure 8.3 are almost identical for high clusters, they are significantly different for poor clusters.

8.3.3 Spatial co-patterning and bivariate spatial clusters in regional income growth

The simplest way to examine the evolution of regional income distribution is to calculate correlations among spatial patterns for different years. Table 8.4 summarizes the computed bivariate associations between starting years and ending years according to Pearson's $r$ and Lee's $L$. For all regionalization schemes, simple correlations between 1969 and 1999 income distributions are extremely high (highest for states (.887) and
Figure 8.3: Geary significance maps: local homogeneity in per capita personal income across the US LMAs
## Table 8.4: Correlations between 1969 and 1999 per capita income distributions

<table>
<thead>
<tr>
<th>Regionalization Scheme</th>
<th>Years</th>
<th>Pearson’s $r$</th>
<th>Lee’s $L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>States (49)</td>
<td>1969-1999</td>
<td>0.887</td>
<td>0.419</td>
</tr>
<tr>
<td></td>
<td>1969-1979</td>
<td>0.933</td>
<td>0.426</td>
</tr>
<tr>
<td></td>
<td>1979-1989</td>
<td>0.816</td>
<td>0.331</td>
</tr>
<tr>
<td></td>
<td>1989-1999</td>
<td>0.972</td>
<td>0.478</td>
</tr>
<tr>
<td>BEAs (170)</td>
<td>1969-1999</td>
<td>0.796</td>
<td>0.384</td>
</tr>
<tr>
<td></td>
<td>1969-1979</td>
<td>0.912</td>
<td>0.527</td>
</tr>
<tr>
<td></td>
<td>1979-1989</td>
<td>0.805</td>
<td>0.385</td>
</tr>
<tr>
<td></td>
<td>1989-1999</td>
<td>0.950</td>
<td>0.413</td>
</tr>
<tr>
<td>LMAs (391)</td>
<td>1969-1999</td>
<td>0.810</td>
<td>0.416</td>
</tr>
<tr>
<td></td>
<td>1969-1979</td>
<td>0.916</td>
<td>0.555</td>
</tr>
<tr>
<td></td>
<td>1979-1989</td>
<td>0.843</td>
<td>0.437</td>
</tr>
<tr>
<td></td>
<td>1989-1999</td>
<td>0.952</td>
<td>0.444</td>
</tr>
<tr>
<td>CZs (722)</td>
<td>1969-1999</td>
<td>0.735</td>
<td>0.396</td>
</tr>
<tr>
<td></td>
<td>1969-1979</td>
<td>0.875</td>
<td>0.536</td>
</tr>
<tr>
<td></td>
<td>1979-1989</td>
<td>0.810</td>
<td>0.430</td>
</tr>
<tr>
<td></td>
<td>1989-1999</td>
<td>0.934</td>
<td>0.477</td>
</tr>
</tbody>
</table>

(All the values are significant at the 95% confidence level)
lowest in CZs (735), implying the dominance of continuity. The fact that values for
CZs are always lower than ones for states is associated with the modifiable areal unit
problem that bivariate correlations usually rise as spatial aggregation proceeds. Among
the three sub-periods, lowest correlations are found in the period of 1979-1989 in
Pearson’s $r$ and Lee’s $L$. This means that restructuring in the regional income
distribution occurred more markedly during the 1980s in comparison with other periods,
and suggests that growing and/or declining regions are spatially dispersed. This
corresponds observations from Figure 8.2 and 8.3 where an obvious discontinuity in
spatial patterns is found between 1979 and 1989 maps.

Effects of the MAUP do not appear in Lee’s $L$ unlike in Pearson’s $r$. This is
partially due to the fact that spatial measures are more resistant to effects of the MAUP
than aspatial ones. Significant Lee’s $L$s indicate that bivariate spatial dependence is
highly pronounced for all periods. When bivariate spatial dependence is present,
statistical tests for Pearson’s correlation coefficients could be flawed, because the degree
of freedom of $n-2$ cannot be obtained. Practically, a high and significant Lee’s $L$
indicates that similar local correlations are spatially clustered. An interesting finding is
that Lee’s $L$ is much higher during 1969-1979 than during 1989-1999 for all
regionalization schemes except for states, even though the latter has higher Pearson’s $r$
than the former. This means that regional income distributions have become more
spatially dispersed in later years so that the level of spatial co-patterning was much higher
in the 1970s than in the 1990s. Hence, we may need to make a clear distinction between
two spatial processes: sporadic and contagious processes. Assuming an identical
Pearson’s $r$, sporadic processes yield a less degree of spatial co-patterning between two
forms of spatial interaction between adjacent LMAs are evident in the income
distribution through all periods, the continuation of high or low values in the 1990s
occurred in a rather sporadic way than in the 1970s. Obviously, the revival of large cities
in the 1990s has led to this trend. This will be more intensively discussed later on.

As can be seen from positive and high Pearson’s $r$ and Lee’s $L$ between 1969 and
1999 regional income distributions in Table 8.4, regional income disparity in the US over
last 30 years is characterized by continuity rather than change. However, it should not be
assumed that each local area equally follows the global trend. Bivariate ESDA
techniques presented in Chapter 7 are expected to reveal spatial heterogeneity in income
trajectory over the last 30 years that each local has experienced.

Figure 8.4 utilizes local-r scatterplot and scatterplot map between 1969 and 1999
income levels. A local-r scatterplot (Figure 8.4-(a)) is created by putting z-scores of 1969
per capita income on the x-axis and ones of 1999 income on the y-axis. The scatterplot
categorizes local bivariate associations into four classes: low-low (lower-than-average
income in 1969 and lower-than-average income in 1999), low-high, high-low, and high-
high. As expected from a high correlation between 1969 incomes and 1999 incomes,
most observations fall within either the high-high quadrant or low-low quadrant. The
local-r scatterplot map (Figure 8.4-(b)) shows that a significant number of areas in the
Midwest and the Pacific regions have fallen from higher-than-average category to lower-
than-average category during the 30 years. It may also be interesting to notice that only
20 LMAs has experienced a shift from the lower-than-average class to the higher-than-
average class, and most of them are located in the South.
Figure 8.4: Local-r scatterplot and scatterplot map of logarithmic per capita personal income across the US LMAs, 1969-1999
spatially smoothed z-scores as defined in Chapter 3, the resulting scatterplot map (Figure 8.5-(a)) benefits pattern detection. Several distinctive patterns are detected. First, the traditional core areas are characterized by continuation of a higher income level with LMAs experiencing the high-low transition. Second, the Pacific counterpart shows a similar pattern, that is, most areas still enjoy higher-than-average income level and some occasional LMAs suffer from economic downturns. Third, areas from the Intermountain through the South to the southern Atlantic coast remain poor except for southern Florida. Within those areas, most of the low-high swing areas reside. They include areas in the Pediment, the Dallas-Houston corridor in Texas, and northern New Mexico centered on Santa Fe. Figure 8.5-(b) selects areas with a statistical significance from Figure 8.5-(a). Bivariate spatial hotspots or spatio-temporal hotspots include the Megalopolis, the central California, Chicago areas, and spatio-temporal coldspots include the lower Mississippi, the ORV region, the southern Texas, and northwestern New Mexico, part of the Four Corners.

8.3.4 c-convergence and spatial dependence of income distribution

Figure 8.6 displays the temporal trend of income dispersion measured by the coefficient of variation (CV) for different regionalization schemes. First, it should be noted that larger spatial units, states and BEAs, have larger CVs than smaller spatial units, LMAs and CZs. As mentioned, it happens because variances decrease as spatial aggregation proceeds. Strikingly, there is no evidence of income convergence during 1969-1999. Perhaps, the last 30 years is too short to display distinctive trends of
Figure 8.5: Local-L scatterplot map and significance map of logarithmic per capita personal income across the US LMAs, 1969-1999
Figure 8.6: Coefficients of variation of per capital personal income in the US, 1969-1999
convergence/divergence. A graph from Rey and Montouri (1999) shows a constant decrease of CV from 1930 and 1975 followed by a relatively flat line. However, some interesting patterns are detected. First, albeit a cyclical fluctuation, a general trend is a convergence until the mid-1970s and then a divergence. Especially during the late 1990s, the trend of divergence is remarkable.

Figure 8.7 provides a different insight into regional income convergence. A CV (coefficient of variation) trend for LMAs is compared to that of Moran's I. A complete correspondence between them indicates that income convergence/divergence is directly associated with spatial dispersion/clustering. Obviously, a contagious process of income distribution is more likely to result in spatial convergence or clustering than a sporadic process. First, spatial autocorrelation measured by Moran’s I gradually decreases during the years, which means that spatial clustering is less pronounced in recent years and can be evidenced from Figure 8.1 and Figure 8.2. Second, two peaks in income divergence in terms of CV, one in 1989 and the other in 1999, seem to be oppositely related to spatial clustering. The 1989 divergence exactly corresponds to spatial clustering in Moran’s I, while the 1999 divergence is oppositely associated with Moran’s I. It can be concluded that income growth or decline may have happened within particular spatial regimes during the late 1980s. Thus, the trend towards income divergence may have been driven by a contagious process. In contrast, another trend of the income divergence in the late 1990s may have been dictated by a sporadic spatial process such that income growth or decline occurred at particular classes of regions that may be represented by population size.
Figure 8.7: Coefficients of variation and Moran’s I of per capital personal income across the US LMAs, 1969-1999
This argument may be advocated by Figure 8.8 where spatial outliers, defined as areas significantly different from their neighbors, are displayed. Throughout the years, significant spatial outliers of high-low association (high values surrounded by low values) are found in the South. When 1979 map is compared to 1989 map in Figure 8.8, one may notice that the number of spatial outliers decreased, indicating spillover effects during the 1980s. Especially, disappearance of spatial outliers in the Pediment during the 1980s, e.g. Charlotte, Atlanta, and Birmingham, are clearly associated with spillover effects (see 1979 and 1989 maps in Figure 8.1). In contrast, 1999 map in Figure 8.8 shows that more areas have become significant spatial outliers during the 1990s. This implies that income growth/decline had been more selective in a spatial sense. For example, Charlotte and Birmingham in the Pediment resurrect as spatial outliers, and areas, including Columbus in Ohio, Traverse City in Michigan, Raleigh in South Carolina, San Antonio in Texas, and Phoenix in Arizona are newly identified as spatial outliers of high-low association. Re-orientation of economy towards selective large cities or economic aggravation in already-lagged areas, for example, may explain the trend. Apparently, this type of spatial process tends to reduce the level of spatial clustering, depending on a given spatial scale, LMAs.

8.3.5 \( \beta \)-convergence and spatially autocorrelated errors

A negative relationship between an initial income level and an income growth rate during a given period of time constitutes the rationale of \( \beta \)-convergence. Table 8.5 lists correlations between the two variables computed with different regionalization schemes. Pearson's correlation coefficient column indicates that there is a significant negative
Figure 8.8: Moran significance maps: detection of spatial outliers in per capita personal income across the US LMAs, 1969-1999
<table>
<thead>
<tr>
<th>Regionalization Scheme</th>
<th>Years</th>
<th>Pearson’s r</th>
<th>Lee’s L</th>
</tr>
</thead>
<tbody>
<tr>
<td>States (49)</td>
<td>1969-1999</td>
<td>-0.370*</td>
<td>-0.252*</td>
</tr>
<tr>
<td></td>
<td>1969-1979</td>
<td>-0.684*</td>
<td>-0.447*</td>
</tr>
<tr>
<td></td>
<td>1979-1989</td>
<td>0.045</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>1989-1999</td>
<td>-0.373*</td>
<td>-0.335*</td>
</tr>
<tr>
<td>BEAs (170)</td>
<td>1969-1999</td>
<td>-0.490*</td>
<td>-0.413*</td>
</tr>
<tr>
<td></td>
<td>1969-1979</td>
<td>-0.529*</td>
<td>-0.336*</td>
</tr>
<tr>
<td></td>
<td>1979-1989</td>
<td>-0.244*</td>
<td>-0.198*</td>
</tr>
<tr>
<td></td>
<td>1989-1999</td>
<td>-0.217*</td>
<td>-0.310*</td>
</tr>
<tr>
<td>LMAs (391)</td>
<td>1969-1999</td>
<td>-0.481*</td>
<td>-0.440*</td>
</tr>
<tr>
<td></td>
<td>1969-1979</td>
<td>-0.548*</td>
<td>-0.385*</td>
</tr>
<tr>
<td></td>
<td>1979-1989</td>
<td>-0.201*</td>
<td>-0.177*</td>
</tr>
<tr>
<td></td>
<td>1989-1999</td>
<td>-0.175*</td>
<td>-0.276*</td>
</tr>
<tr>
<td>CZs (722)</td>
<td>1969-1999</td>
<td>-0.464*</td>
<td>-0.353*</td>
</tr>
<tr>
<td></td>
<td>1969-1979</td>
<td>-0.491*</td>
<td>-0.298*</td>
</tr>
<tr>
<td></td>
<td>1979-1989</td>
<td>-0.242*</td>
<td>-0.167*</td>
</tr>
<tr>
<td></td>
<td>1989-1999</td>
<td>-0.164*</td>
<td>-0.181*</td>
</tr>
</tbody>
</table>

( * significant at the 95% confidence level)

Table 8.5: Correlations between 1969 income levels and income growth rates between 1969 and 1999
relationship between initial income levels and income growth rates during 1969-1999, which may confirm the income convergence hypothesis. However, the latest period, 1989-1999, show a very poor correlation between the two variables especially in LMAs and CZ. This is mainly due to the fact that income divergence is most pronounced in the period as seen in Figure 8.6.

The Lee’s $L$ column, however, reports contrasting information that needs to be explained. Albeit the lowest Pearson’s $r$ (-0.175) in the latest decades across LMAs, Lee’s $L$ is larger than one for 1979-1989 for all regionalization schemes. Since Lee’s $L$ is larger when spatial autocorrelation of variables involved is larger in a technical sense, income growth rates are more spatially clustered in the 1980s than in the 1980s, given information that levels of spatial autocorrelation in 1979 and 1989 are almost identical, seen from Figure 8.6. Another finding is that the highest negative correlation in both columns is found in the period of 1969-1979. Especially Pearson’s $r$ between 1969-1979 is higher in magnitude than that for the entire period, 1969-1999. It can be concluded the catch-up forces were vivid in the 1970s and then finally faded away in the 1990s.

Figure 8.9 is an OLS regression between logarithmic 1969 income levels and annual income growth rates between 1969 and 1999, following equation 8.1. The slope is $-0.010$ and $R$-squared is 0.231 (Table 8.1). Therefore, $\beta$-coefficient for the US LMAs over 30 years is .010 that means that US regional income has converged at a speed of 1% annually. This is too low to conform to the myth of 2% convergence and the exploratory power is rather low. Further, the trend towards income convergence varies sub-period to sub-period. Table 8.6 lists different $\beta$-coefficients for different sub-periods: 0.022 for 1969-1979, 0.012 for 1979-1989, and 0.005 for 1989-1999. Obviously, the thesis of
Figure 8.9: OLS regression between 1969 logarithmic per capital personal income and annual income growth rate across the US LMAs, 1969-1999
Income convergence works best for the 1970s and worst for the 1990s. 1989 income levels only explain 3% of total variance in income growth rates. It thus should be concluded that there has been no $\beta$-convergence since the early 1980s.

As discussed, presence of spatial autocorrelation in OLS residuals may invalidate significance of regression coefficients. Table 8.6 shows that Moran's $I$ tests find a significant spatial autocorrelation in OLS residuals for all sub-periods let alone the entire period. This necessitates use of spatial autoregressive models. Here, I utilize a SAR (simultaneous autoregressive) model. Following Tiefelsdorf's notation (2000:43-44), a SAR model is written:

$$y = X\beta + \varepsilon = \rho V\varepsilon + \eta,$$

where $\varepsilon$ is the correlated error term and $\eta$ is a random white noise. From (8.1), variation of a dependent variable is decomposed into three parts, respectively what Haining (1990:258-259) calls trend, signal, and noise. If there is no spatial autocorrelation, $\rho$, spatial autocorrelation coefficient, will be zero, thus, variance of a dependent variable is decomposed into vectors of predicted values and non-correlated errors. Table 8.6 reports that $\rho$-coefficient for all sub-periods is not negligible, and is highest in 1979-1989. By applying $\varepsilon = y - X\beta$ to the equation, we have:

$$y = X\beta + \rho Vy - \rho VX\beta + \eta$$
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable:</strong> Income growth rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(intercept)</td>
<td>0.143</td>
<td>0.272</td>
<td>0.170</td>
</tr>
<tr>
<td></td>
<td>(20.31)</td>
<td>(19.58)</td>
<td>(6.60)</td>
</tr>
<tr>
<td>Initial income level</td>
<td>-0.010</td>
<td>-0.022</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>(-10.81)</td>
<td>(-12.93)</td>
<td>(-4.05)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.231</td>
<td>0.301</td>
<td>0.040</td>
</tr>
<tr>
<td>Moran's I</td>
<td>0.283*</td>
<td>0.373*</td>
<td>0.573*</td>
</tr>
<tr>
<td></td>
<td>0.123</td>
<td>0.118</td>
<td>0.161</td>
</tr>
</tbody>
</table>

*: significant at the 99% confidence level

Table 8.6: OLS and SAR models for logarithmic initial per capita personal income levels and annual income growth rates
This equation allows a further decomposition (Tiefelsdorf 2000:44): (i) the spatially independent influences of the exogenous component $X\delta$; (ii) the spatially dependent endogenous observations $\rho V\gamma$; (iii) the spatial trend values $\rho VX\delta$; (iv) independent disturbances $\eta$. Further, the variance-covariance matrix $\Omega(\rho)$ among the error terms can be written (Tiefelsdorf 2000:44):

Figure 8.10 shows spatial patterns of decomposition of income growth rates between 1969 and 1999 based on equation 8.1. What a SAR model does is to decompose residuals into spatial autocorrelated errors (signal) and non-autocorrelated ones (noise). The signal map in Figure 8.10 shows that positive residuals are spatially clustered in the South. By eliminating spatially autocorrelated parts from residuals, the noise map rarely displays spatial autocorrelation. Moran's $I$ test for noise resulting from SAR models in Table 8.6 does not reject the null hypothesis that there is no spatial dependence in residuals. A crucial finding is that SAR models significantly lower $t$-values of $\beta$-coefficients. Even though all the $\beta$-coefficients in OLS models are significant at the 99% confidence level, the SAR counterparts are not except for one in the 1969-1979 model. This implies, as Bailey and Gatrell (1995:285) indicates, that OLS models tend to inflate the significance of regression coefficients. Thus, an ultimate conclusion is that there is no statistical evidence of regional income convergence in the US over the last three decades.

8.3.6 Spatial heterogeneity in $\beta$-convergence

It is noteworthy that the negative relationship between initial income levels and income growth rates should not be assumed to apply to an entire study region. The
Figure 8.10: Simultaneous autoregressive (SAR) model decomposition: 1969 logarithmic per capita personal income and income growth rate, 1969-1999

Figure 8.11 utilizes local-\( r \) and local-\( L \) scatterplot maps. The latter is simply a spatially smoothed version of the former, which may benefit pattern detection. From Figure 8.11-(a), one can notice that urban effects are dominant for high-high association. They started at higher-than-average income levels in 1969 and have enjoyed a higher-than-average income growth rates during the last 30 years. Figure 8.11-(b) may help generalize spatial patterns. Areas with the higher-than-average income level in 1969 are associated with lower-than-average growth rates. In contrast, areas with lower-than-average income level in 1969 are associated with higher-than-average growth rate. When Figure 8.11-(b) is compared to Figure 8.5-(a), it is revealed that there is a structural distinction among areas characterized by the continuation of lower income levels; areas from the Mountain region down to western Texas are discernable from ones in the South and the northwestern part of the Midwest; the former has never been involved in the catch-up process; the latter has positively contributed to the catch-up scenario.

Figure 8.12 shows local-\( L \) significance maps for the entire period and three sub-periods. The map for 1969-1999 selects LMAs from Figure 8.11-(b) that are significant. Some interesting patterns are detected. First, the northern part of Megalopolis, southern Florida, and Denver areas have significantly built on their initial higher-than-average income level. Second, economic slowdowns have mostly occurred in the Midwest and the Pacific. Third, most areas in the South except for several urban centers in the
Figure 8.11: Local-r and local-L scatterplot maps of 1969 logarithmic per capita personal income and income growth rate, 1969-1999
Figure 8.12: Local-L significance maps: initial logarithmic per capita personal income levels and annual income growth rates, 1969-1999
Pediment turn out to be significant spatial clusters for the $\beta$-convergence; that is, they conform to the 'catch-up' scenario.

Each sub-period, however, displays a particular level and form of spatial heterogeneity in $\beta$-convergence. In the 1970s, deindustrialization in the traditional industrial belt or rust belt and industrialization in the South centered on the lower Mississippi constituted a dominant pattern. The 1980s experienced re-orientation towards the Megalopolis and southern Florida, marked deindustrialization in the Pacific, and economic slowdowns in the Great Plain, and economic upswings in the South centered on the Pediment. Finally, the 1990s is characterized by economic slowdowns in the Megalopolis, the California region, and southern Florida, and economic revitalization in the South centered on the lower Mississippi.

8.3.7 A geographically weighted regression (GWR): spatially drifting $\beta$-coefficients

Another way to investigate spatial heterogeneity in statistical parameters is to fit a geographically weighted regression (GWR) model (Brunsdon et al. 1996; 1998a; 1998b; 1999; Fotheringham et al., 1997a; 1997b; 1998; 2000). This approach is simply a combination of weighted least squares (WLS) regression and kernel regression (Schimek 2000). However, it is different from the former in the sense that weights matrix in WLS is constant across observations, and is different from the latter in the sense that the weights matrix in GWR is based on spatial proximity, rather than numerical similarity (kernel regression). GWR is also different from any spatial autoregressive models because it produces a set of localized estimates. Fotheringham et al. (1998:1908)
 contend that "looking at a GWR model estimation gives some insight into how localized effects affect coefficients attached to specific variables".

In the regular OLS regression, regression parameters at \( i \)th location are estimated by:

\[
b = \left( X^T X \right)^{-1} X^T y
\]  

(8.3)

In the GWR, they are given:

\[
b_i = \left( X^T W_i X \right)^{-1} X^T W_i y
\]  

(8.4)

where \( W_i \) is an \( n \)-by-\( n \) local spatial weights matrix, which is a diagonal matrix composed of entries in an \( i \)th row in the corresponding global spatial weights matrix. A global spatial weights matrix for GWR is based on inter-distances, and various kernel functions apply to postulate a distance-decay relation. A quartic kernel function is given:

\[
w_{ij} = \begin{cases} 
1 - \left( \frac{d_{ij}}{h} \right)^2 & \text{if } d_{ij} < h \\
0 & \text{otherwise}
\end{cases}
\]  

(8.5)

where \( d_{ij} \) is a distance between spatial objects, and \( h \) is a bandwidth or range beyond which spatial autocorrelation does not exist. In order to determine a \( h \), a cross-validation algorithm can be utilized. However, I fit a variogram for the dependent variable, annual
income growth rates between 1969 and 1999 (Figure 8.13-(a)). A exponential function yields a range value around 215 miles.

Figure 8.13-(b) shows spatial distribution of $\beta$-coefficients. Even though ideally the map is expected to be compatible to Figure 8.11-(b), a considerable degree of discrepancy for some areas is observed. This is mainly because they are based on different perspectives in specifying spatial weights matrices; connectivity-based and distance-based. As suggested, a set of spatially adaptive kernel functions, rather than a global kernel function, may perform better to depict spatial dependence in regression parameters (Brunsden 1995; Fotheringham et al. 2000).

The first two classes in Figure 8.13-(b) belong to negative $\beta$-coefficients which conform to the notion of $\beta$-convergence, while the third and fourth classes indicate a positive relationship between 1969 income levels and income growth rates between 1969 and 1999. Note that the global $\beta$-coefficient is $-0.01$ in Table 8.6. Most areas belong to the first two classes, which is a major observation in Figure 8.11. Focus here is placed on positive values which are against the global trend. High values from California to New Mexico is associated with a combination of high-high association in San Francisco areas and low-low association in the rest of the region (see Figure 8.11(a)). High values in area centered on Seattle, the Megalopolis, and southern Florida are related to high-high association in Figure 8.11.

In comparison with Local-r and Local-L scatterplot maps, GWR seems to provide less information. This suggests that GWR may perform better in a multivariate situation, rather than a bivariate situation. Thus, GWR is more suitable for examining conditional convergence that deals with additional shock variables besides an initial income level.
(c) Variogram

(d) Spatially drifting $\beta$-coefficients

Figure 8.13: A geographically weighted regression
CHAPTER 9

CONCLUSIONS

9.1 Summary of findings

This study aimed to develop a new set of spatial association measures (SAMs), to provide generalized significance testing methods, to propose a set of ESDA techniques using the developed SAMs, and finally to illustrate rationales and usefulness of the proposed ESDA techniques by applying the methods to spatio-temporal income dynamics of the US labor market areas from 1969 to 1999.

In Chapter 2, I attempted to postulate an ESDA-GIS framework based on SAMs. This requires a formulation of the nature of spatial data, which revolves around four interrelated concepts; spatial scale, spatial structures and processes, spatial dependence, and spatial heterogeneity. Statistical results are often dependent upon spatial scale of a study. Various effects of the modifiable areal unit problem (MAUP) have crucial impacts on the understanding of the reality and dictate research results to a large extent. A spatial weight matrix as a mathematical abstraction of spatial structure is necessary for a spatial
spatial processes, tend to induce different spatial structures over relative space, spatial clustering and dispersion. The presence of spatial dependence often erodes the validity of uniform statistical inferences, and the overall level of spatial dependence should be tackled by global SAMs. Since a global trend in a spatial pattern is not expected to apply to all locales, spatial heterogeneity should be investigated to identify spatial outliers and spatial regimes by utilizing local SAMs.

I emphasized the importance of a generalized significance testing framework for SAMs on which various SAMs, whether univariate or bivariate, global or local, or a spatial weights matrix with zero-diagonal or not, are commonly predicated. A well-founded significance testing is necessary for ESDA, because a pattern detection using SAMs will be theoretically more consistent and practically more efficient. I proposed a SAM-based ESDA-GIS framework that is defined as a GIS-based research platform equipped with various ESDA techniques. Developments of the ESDA-GIS framework are strongly connected to the emergence of GIS as a general purpose platform for spatial data analysis. A SAM-based ESDA-GIS framework is characterized by a continuous interaction between GIS and ESDA. A GIS takes advantage of ESDA's statistical integrity and computational efficiency, and ESDA takes advantage of GIS's spatial data management systems and visualization capabilities.

In Chapter 3, rationales for a univariate SAM were first discussed. A global univariate SAM parameterizes the univariate spatial dependence or captures the level of spatial clustering. A local univariate SAM gauges an observation's relative contribution
to the corresponding global trend. A new set of univariate SAMs, $S$ and $S'$, were developed. Lee's $S$, spatial smoothing scalar, captures the degree of spatial smoothing when a variable is transformed to its spatial moving average vector. If a spatial pattern is more spatially clustered, a higher value of $S$ results. It was argued that local Lee's $S_i$ has some advantages over other local univariate SAMs such as local Moran's $I_i$ and local Geary's $c_i$, because it less depends upon an reference area and still works effectively in detecting spatial clusters.

In Chapter 4, need for a bivariate SAM was discussed by elaborating on a conceptual decomposition of association into pairwise point-to-point association and univariate spatial association. A global bivariate SAM captures spatial co-patterning by simultaneously gauging the two types of association. Lee's $L$ is defined as an adjusted Pearson's correlation coefficient between spatial moving averages drawn from the original variables scaled by the square root of the bivariate spatial smoothing scalar that is the product of univariate spatial smoothing scalars. Local Lee's $L_i$ is conceptualized as a spatially varying correlation or localized correlation. Under the presence of bivariate spatial dependence, neighboring locations tend to retain similar bivariate associations across two variables, and thus a set of local Pearson's $r_i$ should display a spatial clustering when it is mapped. Thus, local Lee's $L_i$ can be seen as a spatially smoothed version of local Pearson's $r_i$.

In Chapter 5, a generalized significance testing method based on normality assumption was presented. When a SAM is defined as ratio of quadratic forms, the preexisting algorithm can be utilized to compute first four moments of the measure.
show how various univariate SAMs, whether global or local, can be transformed to ratio of quadratic forms and how the generalized procedure is customized for a particular SAM.

In Chapter 6, a generalized significance testing method based on the randomization assumption was presented. I showed that two general procedures, the Extended Mantel Test and the generalized vector randomization, can yield first two moments for all the SAMs, whether global or local, or whether total or conditional randomization is assumed. I first provided a general procedure and then illustrated how a particular SAM is tailorized to fit to the procedure.

In Chapter 7, I proposed a new set of ESDA techniques utilizing various SAMs and attempted to demonstrate their usefulness with a hypothetical data set. For a univariate situation, local-S significance map and Geary significance map were proposed in comparison with the preexisting Moran scatterplot and Moran significance map. For bivariate situations, local-r and local-L maps, local-r and local-L scatterplot, and local-L significance maps were proposed. A local-L scatterplot is obtained by utilizing two vectors of spatially smoothed z-scores. Four quadrants in the scatterplot stand for different bivariate spatial associations. The lower left and upper right quadrants indicate positive bivariate associations, while the upper left and lower right quadrants indicate negative associations. When areas are referenced by their quadrant locations, categorical maps can be created, local-L scatterplot map. A local-L significance map is created by combining a local-L scatterplot map and a vector of p-values of local $L_s$ derived from a
As bivariate hot spots or cold spots.

In Chapter 8, I analyzed the US annual regional income data from 1969 to 1999 in order to examine the regional income convergence hypothesis by utilizing various ESDA techniques developed in Chapter 7. LMAs (Labor Market Areas) along with other regionalization schemes are used as spatial units. A series of local-S significance maps evidenced the presence of spatial dependence in regional income distribution resulting in distinctive spatial clusters, and show a spatial disintegration within traditional industrial cores in the U.S. over time. Geary significance maps report that local homogeneity is more obvious within hot spots (significantly high income areas) than within cold spots (significantly low income areas).

Extremely high Pearson's $r$ and Lee's $L$ between 1969 and 1999 income distributions indicate the dominance of continuity over change during the 30 years. Correlations within sub-periods show that restructuring in income distribution occurred rather markedly during the 1980s, implying that growing and/or declining areas are spatially dispersed during the decade. A local-L scatterplot map between 1969 and 1999 regional income distributions reveal a significance level of heterogeneity across the US areas. While most areas in the tradition industrial cores including the Midwest industrial belt and the Megalopolis and areas in the Pacific are still enjoying a higher-than-average income level, most areas in the South suffer from the continuation of lower-than-average income.
The notion of \( \alpha \)-convergence was not empirically evidenced. Rather, a general trend towards income divergence was detected since the late 1970s. It was observed that temporal trends in coefficients of variation and Moran's \( I \) do not necessarily correspond to each other. Especially, two peaks of income divergence in terms of coefficients of variation, one in the late 1980s and the other in the late 1990s, seem to be associated with different spatial processes. Contagious spatial processes leading to spatial clustering were dominant in the former, while sporadic spatial processes inducing spatial dispersion somewhat prevailed in the latter. This was supported by Moran significance maps identifying spatial outliers.

The hypothesis of \( \beta \)-convergence was partially evidenced: coefficient was 0.01 and significant. However, the trend varies among sub-periods: 2% convergence rate was found in the 1970s, but the coefficient for the 1990s was minimal (0.5%). A SAR model was fitted to deal with spatial autocorrelation in regression residuals. The results indicate that \( \beta \)-coefficient for the entire period is not significant at the 99% confidence level, which may lead to a conclusion that there is no statistical evidence of regional income convergence in the US over the last three decades.

A local-L scatterplot map and a local-L significance map show that there was a substantive level of spatial heterogeneity in the catch-up process, and suggested possible spatial regimes. While some areas with higher-than-average 1969 incomes, including the northern part of the Megalopolis, southern Florida, and Denver areas, have enjoyed a higher-than-average income growth rate, areas with lower-than-average 1969 income level, LMAs from central Texas to the Four Corners, have experienced a lower-than-
average income growth rate. A series of local-L significance maps for sub-periods show spatio-temporal heterogeneity in \( \beta \)-convergence: different sub-periods display different facets of spatial restructuring. A geographically weighted regression (GWR) model also showed significant level of spatial heterogeneity in \( \beta \)-coefficients.

9.2 Future research agenda

First, \( L \) measure can be extended to calibrate a spatial principal components analysis, which is expected to yield a set of spatially smoothed principal components scores as a multivariate SAM.

Second, a way of computing higher moments for SAMs based on the randomization assumption should be obtained. Since the sampling distributions of Lee's \( L \) and \( L_t \) show a significant level of skewness and kurtosis, the inferential test based on the normal approximation with first two moments does not yield an accurate \( p \)-value. In relation to this, an extension of the generalized significance testing method based on normality assumption or the exact distribution approach should be undertaken for bivariate SAMs as can be exemplified by Tiefelsdorf (2001).

Third, more empirical studies utilizing the ESDA techniques presented in this dissertation should be conducted. An instant application could be an attempt to compute local spatial segregation indices. Since residential segregation intrinsically involves a certain level of spatial clustering, a comparison between spatial distributions of two different racial/ethnic groups should embrace bivariate spatial dependence. While a
global Lee’s $L$ provides a reliable measure for overall spatial exclusion between two
groups with a statistical test, local Lee’s $L_i$ can capture a local degree of residential
segregation. Another field that may effectively utilize the approach could be a
comparison between two raster-based layers such as remotely sensed imageries. The
utilization of local Lee’s $L_i$ could provide a feasible way of conducting a bivariate image
generalization by generating a spatially smoothed layer of correspondence/discrepancy
between two layers.

Fourth, development of a full-fledged platform implementing SAM-based ESDA-
GIS framework proposed in this dissertation should be undertaken. As mentioned
(Goodchild 2000; Marble 2000), current GIS programs and operating systems are
expected to provide users with a better environment for customization of a program and
integration between programs.

This study orients itself to a broader field of geographical information sciences
(GISc) where various disciplines interact with each other and a new academic division of
labor occurs. I would argue that quantitative geography or spatial data analysis in
geography need to move into the new terrain with being equipped with various ESDA
techniques. This is based on a clear distinction between GISc and GIS. Goodchild
(1992:43-44) contends that “the handling of spatial information with GIS technology
presents a range of intellectual and scientific challenges of much greater breadth than the
phrase ‘spatial data handling’ implies—in fact, a geographical information science” and
that “geographical information systems are a tool for geographical information science”.
Marble (2000:32) similarly argues that “the recent rise of GISc as an integrative concept
covering both GIS and spatial analysis certainly works in favor of a broadly based view of spatial analysis and places us in a better position to move rapidly and effectively towards a closer integration of GIS technology and spatial analysis.” This conceptualization implies that we need to retreat much of the discipline’s intellectual resource from technical aspects of GIS and to bring it back to the implementation and sophistication of geographical inquiries with substantive research objectives in the GIS environment. In this sense, SAM-based ESDA-GIS framework presented here could be a solution.
BIBLIOGRAPHY


Amrhein, C. G. and Reynolds, H., 1997, Using the Getis statistic to explore aggregation effects in Metropolitan Toronto Census data, Canadian Geographer, 41:2, 137-149


Anselin, L., 1992, Spatial Data Analysis with GIS: An Introduction to Application in the Social Science, NCGIA Technical Report 92-10, University of California, Santa Barbara


Anselin, L., 1995, Local indicators of spatial association: LISA, Geographical Analysis, 27:2, 93-115

205


Anselin, L. and Griffith, D. A., 1988, Do spatial effects really matter in regression analysis?, Papers of the Regional Science Association, 65, 11-34

Anselin, L. and Smirnov, O., 1996, Efficient algorithms for constructing proper higher order spatial lag operators, Journal of Regional Science, 36, 67-89


Armstrong, H. W., 1995a, Convergence among regions of the European Union, Papers in Regional Science, 74, 143-152

206
Armstrong, H. W., 1995b, Trends and disparities in regional GDP per capita in the European Union, United States and Australia, European Commission Report 94/00/74/017, Brussels


Bailey, T. C. and Gatrell, A. C., 1995, Interactive Spatial Data Analysis, New York: Longman

Bailey, T. C., 1990, GIS and simple systems for visual, interactive, spatial analysis, Cartographic Journal, 27:2, 79-84


207


Bivand, R. S., 1980, A Monte Carlo study of correlation coefficient estimation with spatially autocorrelated observations, *Quaestiones Geographicae*, 6, 5-10


Bivand, R. S., 2000, Using the R statistical data analysis language on GRASS 5.0 GIS database files, *Computers & Geosciences*, 26:9-10, 1043-1052


Casetti, E., 1972, Generating models by the expansion method: applications to geographic research, *Geographical Analysis*, 4, 81-91


Clifford, P., Richardson, S., and Hémon, D., 1989, Assessing the significance of the correlation between two spatial processes, *Biometrics*, 45, 123-134


Cromartie, J. B. and Swanson, L., 1996, Census tracts more precisely define rural populations and areas, Rural Development Perspectives, 11:3, 31-39


Dickey, H., 2001, Regional earnings inequality in Great Britain: a decomposition analysis, Regional Studies, 35, 605-612


Dunford, M., 1996, Disparities in employment, productivity and output in the EU: The roles of labour market governance and welfare regimes, Regional Studies, 30:4, 339-357


Dunford, M. and Perrons, D., 1994, Regional inequality, regimes of accumulation and economic development in contemporary Europe, Transactions of the Institute of British Geographers, 19:2, 163-182


Durbin, J. and Watson, G. S., 1950, Testing for serial correlation in least square regression – I, Biometrika, 37, 409-428

212
Durbin, J. and Watson, G. S., 1951, Testing for serial correlation in least square regression – II, Biometrika, 38, 159-178


Dutilleul, P., 1993, Modifying the t test for assessing the correlation between two spatial processes, Biometrics, 49, 305-314


213


Fotheringham, A. S. and Charlton, M., 1994, GIS and exploratory spatial data analysis, an overview of some research issues, *Geographical Systems*, 1, 315-327


Glick, B. J., 1979, Tests for space-time clustering used in cancer research, Geographical Analysis, 11, 202-208


216


217


218


Hung, T. and Weber, K. E., 1999, Roles of growth centers in regional development: a case study in Northern Thailand, Geographic Information Sciences, 5:2, 131-142


Johnson, K., 1995, Redefinition of the BEA Economic Areas, Survey of Current Business, February, 75-81


Jones, K., 1991b, Specifying and estimating multilevel models for geographical research, Transactions of the Institute of British Geographers, 16:2, 148-159


220


Krishna Iyer, P. V., 1949, The first and second moments of some probability distributions arising from points on a lattice and their application, Biometrika, 36:1-2, 135-141


Lee, S.-I., 2001a, A spatial statistical approach to migration studies: exploring the spatial heterogeneity in place-specific distance parameters, Journal of the Korean Association of Regional Geographers, 7, 107-120

221


Leung, Y., Meo, C.-L., and Zhang, W.-X., 2000, Statistical tests for spatial nonstationarity based on the geographically weighted regression, Environment and Planning A, 32:1, 9-32


Lucas, R. E., 1988, On the mechanics of economic development, Journal of Monetary Economics, 22, 3-42


Majure, J. J. and Cressie, N., 1997, Dynamic graphics for exploring spatial dependence in multivariate spatial data, Geographical Systems, 4:2, 131-158


McDonald, J. A., 1982, Interactive graphics for data analysis, Technical Report, Orion II, Statistics Department, Stanford, University, Stanford, CA


Moellering, H. and Tobler, W., 1972, Geographical variances, *Geographical Analysis*, 4, 34-50


Monmonier, M., 1989, Geographic brushing: enhancing exploratory analysis of the scatterplot matrix, *Geographical Analysis*, 21, 81-84


224


Openshaw, S. and Taylor, P., 1979, A million or so correlation coefficients: three experiments on the modifiable areal unit problem, in *Statistical Applications in the Spatial Science*, edited by N. Wrigley, Pion: London, 127-144


Páez, A., Uchida, T., and Miyamoto, K., 2001, Spatial association and heterogeneity issues in land price models, *Urban Studies*, 38:9, 1493-1508


225


226
Rogerson, P. A., 1999, The detection of clusters using a spatial version of the chi-square goodness-of-fit statistic, Geographical Analysis, 31:1, 130-147


Rowlingson, B. S. and Diggle, P. J., 1993, SPLANCS: Spatial point pattern analysis code in s-plus, Computers & Geosciences, 19:5, 627-655


Scott, A. J., 1988b, Metropolis: From the Division of Labor to Urban Form, Berkely: University of California Press


Sibley, D., 1987, Spatial Applications of Exploratory Data Analysis, Concepts and Techniques in Modern Geography 49, Norwich: Geo Books


Talen, E., 1997, The social equity of urban service distribution: an exploration of park access in Pueblo, Colorado, and Macon, Georgia, Urban Geography, 18:6, 521-541


Tiefelsdorf, M., 2001, Specification and distributional properties of the spatial cross-correlation coefficient $C_{x,y}$, Paper presented at the Western Regional Science Conference, Palm Springs, February 26


Tsionas, E. G., 2000, Regional growth and convergence: evidence from the United States, *Regional Studies, 34*, 231-238


229


Sociology*, 88, 1008-1018

Wilhelm, A. and Steck, R., 1998, Exploring spatial data by using interactive graphics and
local statistics, *Journal of the Royal Statistical Society D: The Statistician*, 47:3,
423-430

Williamson, J. G., 1965, Regional inequality and the process of national development: a
description of the patterns, *Economic Development and Cultural Change*, 13, 3-
45

exploratory analysis of area data, *Environment and Planning A*, 31, 1825-1838


Wong, D. W. S., 1996, Aggregation effects in geo-referenced data, in *Practical
Handbook of Spatial Statistics*, Arlinghaus, S. L., Griffith, D. A., Arlinghaus, W.
C., Drake, W. D., and Nystuen, J. D., eds., Boca Raton: CRC Press, 83-106

Wrigley, N., 1995, Revisiting the modifiable areal unit problem and the ecological
fallacy, in *Diffusing Geography: Essay for Peter Haggett*, edited by Cliff, A. D.,
Gould, P. R., Thrift, N. J., Oxford: Blackwell, 49-71

Wulder, M. and Boots, B., 1998, Local spatial autocorrelation characteristics of remotely
sensed imagery assessed with the Getis statistic, *International Journal of Remote
Sensing*, 19:11, 2223-2231

Ying, L. G., 2000, Measuring the spillover effects: some Chinese evidence, *Papers in
Regional Science*, 79:1, 75-89

Zhang, Z. and Griffith, D. A., 1997, Developing user-friendly spatial statistical analysis
modules for GIS: An example using ArcView, *Computers, Environment and
Urban Systems*, 21, 5-29

Zhang, Z. and Griffith, D. A., 2000, Integrating GIS components and spatial statistical
analysis in DBMSs, *International Journal of Geographical Information Science*,
14:6, 543-566

Zhao, X. B. and Tong, S. P., 2000, Unequal economic development in China: spatial
disparities and regional policy reconsideration, 1985-1995, *Regional Studies*, 34,
549-561

231