EVOLUTION OF INVESTMENT FLOWS

IN U.S. MANUFACTURING:

A SPATIAL PANEL APPROACH

by

Jason P. Brown, Raymond J.G.M. Florax,
and Kevin T. McNamara

Working Paper # 08-06

(This paper was revised September 2009. Please see Working Paper #09-10
"Determinants of Investment Flows in U.S. Manufacturing")

August 2008

Dept. of Agricultural Economics

Purdue University

EVOLUTION OF INVESTMENT FLOWS IN U.S. MANUFACTURING: A SPATIAL PANEL APPROACH

by

Jason P. Brown,[1] Raymond J.G.M. Florax,[1,2] and Kevin T. McNamara[1]

[1] Dept. of Agricultural Economics, Purdue University
403 W. State Street, West Lafayette, IN 47907–2056, USA
Phone: +1 (765) 494–7960, Fax: +1 (765) 494–9176
E-mail: brown151@purdue.edu, rflorax@purdue.edu, mcnamara@purdue.edu

[2] Dept. of Spatial Economics, VU University
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

Working Paper # 08-06
August 2008

Abstract

The paper starts with a discussion of a conceptual model of location factors in U.S. manufacturing investment at the state level. The purpose of the paper is to test the relative importance of growth factors influencing investment and whether or not they have changed in importance over time. These factors include agglomeration, market structure, labor, infrastructure, and fiscal policy. A better understanding of investment flows in the manufacturing sector will help determine how growth factors have changed over time and which economic development policies may be most appropriate at targeting the sector. The analysis covers the time period 1994 to 2006 for the 48 contiguous states, with data taken from the Annual Survey of Manufactures, the Bureau of Economic Analysis, and the Bureau of Labor Statistics. Panel methods are used to test for fixed effects due to heterogeneity across states. Spatial panel methods with time effects are used for determination and specification of spatial and temporal effects. Empirical results are consistent across the empirical models put forth. Results suggest that market demand remains one of the most important location factors of manufacturing investment. Investment also goes to states with more productive labor and localized agglomeration of manufacturing activity.

Keywords: manufacturing, investment, location factors

JEL Codes: L60, R11, R30

1. Introduction

Manufacturing employment in the U.S. has been on a steady long-term decline but increasing for many multiyear periods since the 1960s until present. At the same time, real wages with the exception of late ‘70s and early ‘80s, real investment on total capital and manufacturing output have been increasing. During this 40 year-period manufacturing increased its capital intensity, especially with the integration of computer technology into manufacturing systems that began in the 1990s resulting in higher productivity of labor. This was not due merely to the increased availability of skilled labor, since the relative wages of skilled workers increased dramatically along with their employment. Rather, the patterns of wages and employment suggest technical change resulting in increased demand for skilled labor. The onset of the global market place led to shifts in production overseas in order to optimally satisfy market demand. The relationship between investment flows and its underlying determinants is important for the shaping of economic development policy. Specifically, the inverse relationship between manufacturing employment and output provides an opportunity to examine whether the shift of manufacturing to more capital intensive production influenced factors important in firms’ decisions about the geographic location of capital investment.

Location theory has a rich history of scientific research dating back to later part of the 19th and early part of the 20th centuries. Its beginnings are found in seminal works of von Thünen (1826), Weber (1929), and Lösch (1954). Renewed interest took hold in the middle 1900s in the work of Hoover (1948), McLaughlin and Robock (1949), Isard (1948) and Greenhut (1956). Common themes in this literature were aimed at the empirically testing the theoretical determinants of the distribution and choice of industrial location. Empirical research has continued throughout subsequent decades to better establish the determinants of industrial location. Recent work by Guimarães et al. (2004) has acknowledged a renewed interest in addressing these questions following the emergence of “new economic geography” where agglomeration forces are purported to be important in clustering of economic activity.

Moreover, the location determinants of manufacturing investment are likely to evolve as the composition of the manufacturing industry changes (Woodward and Glickman, 1991). Blair and Premus (1987) state that the determinants of locational choices change as conditions of production change. They cite the example of U.S. appliance companies who in the mid 1960s could gain a large cost savings if they relocated from the North to the South. However, their review of major factors in industrial location found that the traditional economic factors of location were becoming “quantitatively less significant”. In fact, the investigation of location determinants changing over time is the purpose of this research. What follows in subsequent sections is a discussion of a conceptual model of location factors, an exploratory spatial data analysis, an empirical model and estimation results, and finally conclusions and policy implications.

2. Conceptual Framework

Profit maximization is often cited as the objective behind industrial location. Greenhut (1956), in one of the first works to extend Lösch's framework, describes a firm's location decision based upon a profit maximization framework. According to Greenhut, firms choose a site from which there is sufficient demand (buyers) whose purchases of the firm's product are required for maximum sales that is served at the least possible cost. The location need not be lowest in total cost but rather a location from which monopolistic control over buyers makes it more profitable than a lower cost site as it relates to spatial monopoly power *á la* Hotelling (1929). More specifically, firms examine alternative location choices to find regional differences in both product demand and supply of necessary inputs, i.e. labor. Access to product markets may be dependent on whether an urban or rural location is chosen. The advantage of the profit maximizing approach is that it recognizes the interaction between demand and the cost of production in location choice. It also allows for the analysis of both demand and cost factors that may influence the location decision. While somewhat restrictive, firms are assumed to have perfect information in demand and supply markets. Wasylenko and McGuire (1985) relate profit to the local price and quantity sold of the firm's output as well as to the same for inputs, all of which vary across space. Product revenues and input expenditures may also vary to the extent that local areas tax or subsidize inputs and outputs (Gerking and Morgan, 1991).

An additional common framework assumes that location decisions of manufacturing investment occur as a two-stage process. McLaughlin and Robock (1949) describe firms first selecting a general area on the basis of the most important advantage of location for a given type of manufacturing. Secondly, firms choose a specific location within the general area. McLaughlin and Robock were among the first to put forth this framework although it was not empirically tested until Schmenner et al. (1987). Schmenner and his co-authors developed a conceptual model of location decisions which derived from the premise that a manufacturing plants choice is based upon considerations of long-run profitability. The location factors are assumed to affect the location decision in two stages. Three categories of state-specific characteristics are hypothesized to affect the expected profitability of a plant. The first category is an indicator of the cost and supply of inputs. The second category is fiscal impacts from the government or governmental influence in general. The third category is geographic or demographic features such as amenities, population density, etc. As part of their framework plant characteristics were also included as they are expected to change the influence of the state characteristics. This is similar to McLaughlin and Robock's (1949) categorization of plants into those that are market oriented, material oriented, and labor oriented. Location factors are expected to change in importance depending on the type of manufacturing plant. Others, such as Muller and Morgan (1962), have grouped manufacturing location decisions based whether it concerned a new firm, relocation, or expansion of existing facilities. Blair and Premus (1987) and Schmenner (1982) have suggested relocations are rare and when they do occur the majority are over short distances. Moreover, location patterns have been suggested to change depending on the life-cycle of the firm with interregional moves being uncommon (Wissen, 2000).

Subsequent work by Bartik (1989), Woodward (1992), Henderson and McNamara (1997), and Lambert, et al. (2006) have framed that location of manufacturing investment as a two-stage process as well. The locations factor approach was first suggested by Fredrich Hall in the 1900 Census of Manufactures (Jones and Woods, 2002). In the context of the location decision, the location factors can be formally expressed in a conceptual model as $X_i = g(A, S, L, I, F)$, where X is the location choice of the investment, i indexes the choice set, and A, S, L, I, F are vectors of state attributes corresponding to agglomeration forces (A), market structure (S), labor (L), infrastructure (I), and fiscal (F) factors that influence a firm's cost structure. No restriction is made on the functional form of *g*, the function assumed to minimize total costs. What follows next is a description of each of the location factors.

Agglomeration (A) economies is a well-known concept by now. Hoover (1948) summed up agglomeration into internal returns to scale which are firm-specific economies of agglomeration, localization economies which are industry-specific, and urbanization economies which are city-specific economies of agglomeration. Internal returns are firm-specific in that efficiency gains are a direct result of the size of the individual firm. Internal economies of scale may come about by a large level of investment that takes place at one particular location. Local economies occur when similar firms agglomerate across relatively close locations making it easier to find a large enough pool of skilled labor, or use of specialized services which are non-traded. There can also be information spillovers from close interaction among workers across firms or demand and supply-side spillovers if firms produce products used by each other or require similar inputs used in production. Glaeser et al. (1992) describes the Marshall-Arrow-Romer externalities that come about from knowledge spillovers between firms in an industry when internalized lead to innovation and increased growth. Porter's theory of agglomeration due to competition holds the same view (Porter, 1990). However, Jacobs (1969) states that knowledge spillovers come from outside the common industry via urban economies. Urbanization economies are those of agglomeration which are gained to firms across different sectors (McCann, 2001). This can be observed around areas where high concentrations of manufacturing and service sectors coinciding. Clustering occurs in response to large local market possibilities that exist. Agglomeration economies are hypothesized to positively impact manufacturing investment.

Market structure (S) and nearness of markets is important for industrial-products and was recognized rather early in the development of location theory and location factors. Neo-classical theory hypothesizes that production patterns are uniquely determined by relative price and supply factors. Demand concentration induces production concentration. Market influence is the tendency of supply to become geographically distributed according to demand (Wheat, 1986). Plant investment decisions are influenced by access to product markets, particularly when they are the source of final demand (Bartik, 1989; Woodward, 1992). Moreover, survey research has shown that market structure is consistently the most important factor in investment location decisions (Mueller and Morgan, 1962; Schmenner, et al.; Blair and Premus, 1987; Calzonetti and Walker, 1991; Crone, 2000). Breaking market structure into size and density, it is hypothesized that size will have a positive impact on investment, while density will detract due to sufficient product demand and excessive competition for resources respectively.

After the cost of materials, labor (L) costs are the largest component of the average manufacturing plants operating expenses (Crone, 1997). Cost and availability of labor have been major factors in investment location decisions (Smith, et al., 1978; Schmenner, et al., 1987; Henderson and McNamara, 2000). However, skilled labor has increasingly become important (McGranahan, 2000). This trend is expected to continue as manufacturing production becomes relatively more capital intensive. A large proportion of the manufacturing industries that are labor intensive, such as the textile industry, have been outsourced to regions of the world where unskilled labor is ubiquitous and cheap relative to the U.S. Low wage levels, high labor availability, and human capital are expected to attract manufacturing investment holding other factors constant.

Infrastructure (I) aids economic development of a state by creating access to regional and national markets. Broadly speaking, infrastructure can include transportation systems and available land. Transportation has been a central focus in location theory dating back to von Thünen. New transportation technology and changing regional cost structures has tended to improve the advantages of certain areas. Throughout the literature, access to infrastructure has been shown to attract manufacturing investment (Smith et al., 1978; Bartik, 1989; Holl, 2000; Carlson, 2000; Cohen and Paul, 2004; Lambert et al., 2006a). Jones and Woods (2002) argue that the National Interstate Highway System has opened industrial possibilities to the vast majority of rural America, especially in the South. Available land has also been cited as an important location factor (Carlson, 2000). Manufacturing firms concerned with land availability tend to avoid cities and densely populated locations. Both measures of infrastructure are hypothesized to attract manufacturing investment.

Fiscal (F) factors are also a very common point of discussion in the industrial location literature. Direct and indirect policies affect the cost of conducting business via taxes and regulations imposed. Taxes are viewed as a deterrent to manufacturing investment in the literature (Carlton, 1983; Plaut and Pluta, 1983; Wheat, 1986; Bartik, 1989, 1992). Additionally, some surveys about plant location decisions find taxes as an important factor (Kieschnick, 1981; Hekman, 1982; Calzonetti and Walker, 1991). However, as pointed out by Papke (1991), it is effective or net taxes that are important to investment location. Oftentimes, manufacturing plants avoid or pay low property taxes due to economic development policies of state and local governments. State expenditures are also a part of fiscal factors. Higher spending is often seen as a benefit in the location literature (Plaut and Pluta, 1983; Goetz, 1997). Fiscal expenditures directed towards education, worker training, and infrastructure can impact a manufacturing plant's profitability. States with high tax rates are expected to deter investment while state expenditures are expected to attract investment.

3. Data and Exploratory Spatial Analysis

3.1 Data Description

The purpose of the conceptual model is to direct the development of the empirical model and data selection. Manufacturing investment in the literature is most commonly proxied by manufacturing activity via counts of new plant births, change in employment growth, or levels of capital expenditure. The least common proxy is capital expenditures as they are more difficult to obtain, only publicly available at the state level, and are only available for the two-digit NAICS code (31-33). Data on capital expenditures provide a different perspective on manufacturing investment as opposed to counts of new firms or employment growth both of which have been in decline. While all three measures are subject to economic cycles, capital investment has generally been on an upward trend over the last fifty years (Berman, et al., 1994). Other studies that have used a similar measure are Benson and Johnson (1986), Papke (1987), and Gupta and Hofmann (2003). Total capital expenditures for the aggregate manufacturing sector are taken from the Annual Survey of Manufactures (ASM) produced by the U.S. Census Bureau from 1994 to 2006. The 48 contiguous states are used excluding Washington D.C. as the investment data are not available in every year. The investment figures are converted to real dollars using the CPI with 2006 as the base year.

Manufacturer's share of employment serves as a measure of agglomeration that expresses localized economies to scale. Agglomeration due to urbanization economies is measured using the state density of business establishments per square kilometer (e.g. Bartik, 1989; Guimarães et al., 2004). Market demand structure is measured using total population and gross state product per capita (e.g. Henderson and McNamara, 2000; Davis and Schluter, 2005; Lambert et al., 2006b). Labor factors are measured using average hourly manufacturing production wage for labor cost, value of production per hour for productivity of labor, unemployment rates for availability, and worker union participation rates for labor climate. The productivity measure was calculated by taking the value of manufacturing shipments adding the net between beginning and ending inventories for the year. This value term was then divided by the total number of production hours. The current measure is much more appropriate than gross value added (GVA) constructed by the Bureau of Economic Analysis (BEA) as there measure includes purchased services. Infrastructure factors are measured using kilometers of Interstate highway in the state to proxy for transportation access and ratio of farm land to total area of a state to proxy for land availability (e.g. Henderson and McNamara, 2000). State personal income tax rates and state expenditure per capita are used as measures for fiscal factors (e.g. Plaut and Pluta, 1983; Bartik, 1989). Table 3.1 provides the descriptive statistics and sources of the data.

<< Insert Table 3.1 >>

What follows is a discussion of the data directly related to manufacturing activity. Investment ranges over the time period from $49 million (Wyoming, 2003) to $53.1 billion (Ohio, 1996) with an average $3.4 billion across all states and years. The states having the most investment in all years are California, Michigan, Ohio, and Texas. This is somewhat expected as

Michigan and Ohio have traditionally been strong manufacturing states, while California and Texas are very large states. However, size alone of a state does not guarantee proportional investment as will be shown later. Manufacturing share of employment provides insight as to where it is heavily concentrated. The average is 12% with the range being 3% to 26%. Indiana had the highest percentage in all years with the exception of North Carolina in 1995. States with the highest percentages have typically been Indiana, Michigan, and Ohio in the Midwest, and Arkansas, Mississippi, Alabama, North Carolina and South Carolina in the South. The average hourly production wage ranges from $13.01 to $23.41 with an average of $17.52. Wages are typically higher in Michigan, Indiana, and Ohio in the Midwest, Connecticut and Delaware in the Northeast, and Washington and Wyoming in the West. Moreover, wages are lower in the South with exception of Louisiana and in some years West Virginia. Value of manufacturing production measured in thousands of dollars per production hour ranges from $114 to $855 with an average of $220. The highest value is Louisiana in 2006. Prior to 2005 and Hurricane Katrina, Louisiana still had the highest value of production at $611 in 2004.

Looking at these data over time provides some insight on interesting trends in manufacturing. Figure 3.1 shows investment measured in millions of real dollars for each state. There is an expansion in investment during the mid to late '90s which coincides with the macroeconomic expansion of that period. Beginning in 2000, a contraction in investment occurred which worsened after 2001, coinciding with the economic recession over the same period. The level of investment appears to flatten out by 2006. This economic cycle provides an opportunity to investigate whether location factors change in importance during a time of economic expansion versus contraction. Figure 3.2 illustrates another interesting trend in value of manufacturing output measured in thousands of real dollars per production hour. An upward trend occurs across all states. Such a trend would be expected as capital investment in manufacturing continues to increase along side of technological advancement. Figure 3.3 shows that over the same period an upward trend in average hourly production wages occurred across most states. However, there are some states where wages have gone down slightly since about 2004. In light of these trends, the increasing need for skilled versus unskilled labor for manufacturing production seems very plausible and has been supported in the literature (McGranahan, 2000).

<<Insert Figure 3.1>>
<<Insert Figure 3.2>>
<<Insert Figure 3.3>>

3.2 Exploratory Spatial Data Analysis

Since states are generally smaller in the East and larger in the West, use of cartograms to geographically display the levels of manufacturing investment can be rather useful. Cartograms are useful when large areal units have small values of investment or when small areal units have large levels of investment. It is the proportional size of the attribute value that corresponds to the size and color of a shape used in constructing the cartogram versus the physical area (Tobler, 2004). Figure 3.4 shows cartograms constructed in ArcGIS 9.2[1] for manufacturing investment in 1994 (top) and

[1] ArcGIS uses the Gaster-Newman method (Newman and Gaster, 2004) in order to construct the cartograms.

2006 (bottom). We see that states out west generally have low levels of investment given the level of distortion that shrinks the states. However, in the Midwest and in parts of the South there are states with high levels of investment which are clustered together. States which have the highest levels of investment, shown in darker shades of red, are Michigan, Illinois, Indiana, and Ohio in the Midwest, Texas in the South, and California in the West. This same pattern persists across time except for 2001 when some of the Midwestern states such as Michigan and Ohio still have relatively high levels of investment, but are less distorted compared to other states. In 2006 California and Texas remained states with high investment compared to others. This pattern is likely explained by the composition of the manufacturing industry in those states. Texas has a high percentage of their manufacturing activity related to chemical manufacturing which is highly capital intensive. Similarly, California has a high proportion of computer and electronic product manufacturing compared to other states.

<<Insert Figure 3.4>>

Tests for spatial dependence are performed given the observed clusters of states that have low and high investment. In order to test for spatial dependence some structural assumptions must be made on neighboring criteria. Use of commonly shared borders as neighboring criteria is one possibility. States are non-uniform shapes and often share borders in all possible directions. Therefore, it seems most appropriate to use queen contiguity versus rook for this type of neighboring criteria. The contiguity matrix assumes that spatial interaction or dependence occurs only between states that share a common border. Bernat (1996) used a queen order one contiguity matrix when looking at manufacturing activity in the 48 contiguous states. Alternatively, distance can be used to define neighbors such as using the minimum distance for a state to have at least one neighbor or some other alternative cutoff distance. However, the distance must be large enough so that western states are not left as "islands" in the spatial system due to the smaller size of states in the Northeast. The downside to this weighting scheme is that it is less sparse as it results in regions having more neighbors. Figure 3.5 shows the connectivity of the spatial system under the queen order one weighting scheme. Construction of a spatial correlogram showed that the spatial range did not extend beyond the first order neighbors. As a result, queen order one was used as the weight matrix in the exploratory analysis and in the diagnostics of the empirical models.

<< Insert Figure 3.5 >>

Moran's I was calculated for investment using one thousand permutations to provide a statistical test. The log of investment was used as this is the dependent variable described in the next section. The resulting value was 0.33 statistically significant at the one percent level. Moran's I value suggests states with high levels of investment are expected to neighbor states that have high levels of investment. This same relationship holds also with low levels of investment. Figure 3.6 depicts this relationship in a standardized Moran's I scatter plot. The top-right quadrant represents states with high levels of investment that have neighbors with high levels of investment. The figure shows that there are numerous states which have low levels of investment that have neighbors with low levels of investment. Moreover, these states are more scattered in the low-low quadrant of Figure 3.6 versus the high-high quadrant.

<< Insert Figure 3.6>>

Local measures of spatial dependence can also be used to explore data. The most common statistical tool used is Anselin's local indicators of spatial autocorrelation (LISA) found in Anselin (1995). LISA allows for the identification of clusters of "hot spots" or "cold spots" as well as potential outliers when states with low levels of investment are located next to states with high levels and vice versa. In doing so, it actually reveals more about heterogeneity even though it is measuring local autocorrelation. For each state, LISA values allow for the computation of its similarity with neighboring states as well as statistical significance. LISA statistics were calculated on the pooled investment data. The same weight specification was used as in the global test of spatial dependence. States that had a significant LISA remained so across time with a few exceptions particularly in the Northeast and South. Significant clusters included Maine, Pennsylvania, and New Hampshire in the Northeast; Indiana, Illinois, Michigan, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin in the Midwest; Georgia and North Carolina in the South; Idaho, Montana, Utah and Wyoming in the West. Figure 3.7 shows LISA maps generated for 1994 and 2006 under 999 permutations using GeoDa™ (Anselin, 2004). Red colors indicate states with high levels of investment whose neighbors have high levels of investment. Blue states represent the same relationship but with low levels of investment. Each LISA value represented in the maps was statistically significant at the 95% level. If the critical value were expanded out to the 10% level, more of the Southern states would have been included particularly Georgia and North Carolina.

4. Empirical Model and Estimation

4.1 A-Spatial Cross-Sectional Models

According to Blair and Premus (1987), econometric studies examine where firms locate or investigate other factors that indicate change in or levels of manufacturing activity. Econometric models used in these studies are typically expressed in "change/level" or "change/change" (Bartik, 1985). The first classification is known as the disequilibrium-adjustment model in which the change in the dependent variable over the period is related to the levels of independent variables at the beginning of the period. This model is based on the assumption that differences in manufacturing profitability at the beginning of the period are large enough to cause differences in the rate of manufacturing growth. In contrast, the "change/change" model is where growth in a region is a result of the changes in a region's characteristics. Bartik (1985) summarizes the advantages and disadvantages of both models. Generally, the "change/level" is preferred as it does not eliminate potential fixed effects due to differencing and requires less data to operationalize. It is also common in the empirical literature for industrial location models to be expressed in log-linear form in the case of counts of new establishments or in double-log form when using levels of capital expenditure (Carlton, 1983; Guimarães, et al., 2004; Gupta and Hofmann, 2003). The double-log specification has the advantage of interpreting the regression coefficients as an elasticity measure, i.e. how the dependent variable will change given a 1% change in an independent variable *ceterus paribus*. Therefore, the base empirical model is formally given as:

$$\begin{aligned}\ln INV &= \beta_0 + \ln INV_{t-1}\beta_1 + \ln MEMP\beta_2 + \ln UE\beta_3 + \ln GDPC\beta_4 + \ln POP\beta_5 \\ &+ \ln PWAGE\beta_6 + \ln PROD\beta_7 + \ln UNEMPR\beta_8 + \ln UNION\beta_9 + \ln INTERST\beta_{10} \quad (4.1) \\ &+ \ln AVLAND\beta_{11} + \ln PIT\beta_{12} + \ln STEC\beta_{13} + \varepsilon\end{aligned}$$

where ε is assumed to be an i.i.d. error term. None of the previous literature using capital expenditure as a measure of manufacturing investment included lagged investment into the econometric model. However, a portion of the investment made in each period could be linked to the same overall investment as well to replace depreciated capital during the same time of an expansion. Therefore, having tested the appropriate lag structure, investment from the previous period is also included into the model. Investment from more than one previous period was not statistically significant.

Equation (4.1) can be modified to allow for time effects and or spatial heterogeneity. Time effects are necessary due to upward trends in real wages, value of output, and real investment. They also help capture economic cycles of expansion and contraction. Discussion of the data in the previous section pointed out how some of the proposed location factors vary widely across the U.S. Assuming a global coefficient for each location factor imposes the restriction that the impact on investment will be the same everywhere. Spatial heterogeneity is discussed further in section 4.4.

4.2 Spatial Process Models

Traditional industrial location models have tended to ignore the possibility of spillovers in investment from one region to the next. When spatial dependence is present and appropriately modeled more accurate estimates can be obtained of location factors. Tests for global spatial dependence using Moran's I can be conducted on the error terms from the a-spatial model in equation (4.1). Lagrange multiplier tests are used to determine which spatial process is most appropriate following guidelines in Anselin et al. (1996). There is no theoretical base for one over the other in the industrial location literature. Lambert et al. (2006a) tested for and incorporated spatial processes using geographically weighted regression (GWR) and Poisson regressions when looking at location factors of counts of new manufacturing firms. A distance decay function was used which applies geostatistical concepts, the direct approach for modeling spatial dependence. Klier and McMillen (2005) use a GMM estimated spatial logit model to explain clustering of auto supplier plants. Lambert et al. (2006b) in looking at food manufacturing investment, use a spatial probit model to estimate a spatial lag model to account for spillover effects.

Spatial dependence is routinely modeled as an additional covariate in the form of a spatially lagged dependent variable Wy, or in the error structure where $E[\varepsilon_i \varepsilon_j] \neq 0$. The first is referred to as a spatial lag model and is utilized when importance is granted to the presence of spatial interaction. Spatial dependence in the error term can take many forms via the spatial error model and is commonly referred to as nuisance dependence (Anselin, 2003). This specification is appropriate when trying to correct for spatial autocorrelation. A spatial error model can be expressed as:

$$y = X\beta + \mu$$
$$\mu = \lambda W\mu + \varepsilon \tag{4.2}$$

where y is a vector $(N \times 1)$ of observations on the dependent variable, X an $N \times K$ matrix of observations on the explanatory variables given in equation (3.1), and μ an error term that follows a spatial autoregressive (SAR) specification with autoregressive coefficient λ. In the spatial autoregression, the vector of errors is expressed as a sum of a vector of random terms (ε) and a so-called spatially lagged error, $W\mu$. Formally, a spatial lag model is given by

$$y = \rho Wy + X\beta + \varepsilon \tag{4.3}$$

where ρ is a spatial autoregressive coefficient often referred to as a spatial correlation coefficient, ε is a vector of error terms, Wy is a spatial lag operator which is a weighted average of y variables at the neighboring locations (Anselin, 2003). Neighboring criteria determines the structure of W which is routinely based off of contiguity (queen or rook) or distance criteria. The weights in W are usually row-standardized so that elements sum to one. Equations 4.2 and 4.3 are most commonly estimated with maximum likelihood. However, the spatial lag model expressed in reduced form shows that Wy is correlated with the error term,

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon \tag{4.4}$$

Under this specification an investment decision in a given state is connected to all other investment decisions by the spatial multiplier $(I - \rho W)^{-1}$ and the error term (Anselin, 2003). Equations 4.2 and 4.3 are often estimated via maximum likelihood. Spatial econometric literature has identified the advantages of maximum likelihood estimation (MLE) in the presence of spatial dependence (e.g. Anselin, 1988; Anselin and Hudak, 1992; Elhorst, 2003). Kelejian and Prucha (1999) provide an alternative estimation procedure of spatial process models via generalized method of moments (GMM). They point to at least two potential advantages of this approach: 1) relaxing the assumption of normality of error terms, and 2) less computationally burdensome compared to MLE.

4.3 Spatial Heterogeneity

A further consideration in space is that of spatial heterogeneity. Spatial dependence can come about if un-modeled relationships between location factors and investment vary across space. Spatial regime models allow for modeling this relationship in a discrete way by allowing coefficients to vary across space. For example, it is possible that union participation rates will have a different impact on investment location in different regions in the U.S. Unions can cause wages to be sticky and lead to higher labor costs. Southern states generally have lower union participation as compared to states in the Northeast and Midwest and therefore investment may not be as sensitive to labor unions in the South. This of course can be tested using the spatial regime model, which in general form where an intercept is included can be expressed as:

$$y_G = \sum^{G-1} \beta_G + X_G \beta_G + \varepsilon_G \tag{4.5}$$

where G is the number of groups or regions. All regional dummies can be included in the equation as long as the intercept term is suppressed. Chow tests and t-tests can be used to determine if there is a significant difference across models or in coefficients. One potential downside to this specification is loss of degrees of freedom. Therefore, it is best to limit the number of regions to only a few. The exploratory spatial data analysis in section three showed that local measures of spatial clustering in investment occurred heavily in the Midwest, and a few states in the West and Northeast. A similar pattern of U.S. manufacturing activity was found by Bernat (1996) which define regions based on local Moran Statistics. This classification is also similar to the four regions used by the U.S. Census Bureau. As a result, states are broken into four categories to define the spatial regimes with regions one to four being Northeast, Midwest, South, and West.

4.5 A-Spatial Panel Models

An additional way to deal with heterogeneity is in panel data methods. Fixed and random effects models become useful when dealing with cross-sectional heterogeneity due to a vast number of unmeasured variables that determine y, or manufacturing investment in this case. This phenomenon suggests that OLS is biased unless the influence of these omitted variables is uncorrelated with the included explanatory variables (Kennedy, 2003). There are two ways routinely suggested to improve estimation by using different intercepts for each cross-sectional unit, i.e. states (Wooldridge, 2002). The fixed effects model in standard notation is given by:

$$y_{it} = \alpha_i + \beta x_{it} + e_{it} \tag{4.6a}$$

and after transformation $(y_{it} - \bar{y}_i) = \beta(x_{it} - \bar{x}_i) + (\varepsilon_{it} - \bar{\varepsilon}_i), t = 1,...,T, i = 1,...,N$ (4.6b)

where each cross-sectional unit has its own intercept, but after subtracting the mean from each cross sectional element, the intercept term disappears. There are two downsides to the fixed effects model: 1) loss of degrees of freedom due to the dummy variables, and 2) the transformation used in the estimation process potentially removes explanatory variables that do not vary over time. Only the first will have an impact in this analysis.

Random effects panel models are designed to overcome the two drawbacks of the fixed effects model. The model is comparable to the fixed effects model in that it hypothesizes a unique intercept for each individual, but it interprets the differing intercepts in a unique way (Kennedy, 2003). The procedure views the different intercept for each cross sectional unit as have being randomly drawn from a distribution of possible intercepts. Consequently, they are interpreted as random and considered to be part of the error term. The usual specification is:

$$y_{it} = \mu + \beta x_{it} + (u_i - \varepsilon_{it}) \tag{4.7}$$

where μ is the mean of the random intercepts $\alpha_i = \mu + u_i$, and the errors u_i and ε_{it} make up the composite error term. A Hausman test (Hausman, 1978) can then be used to compare the fixed versus random effects under the null hypothesis that the individual effects are uncorrelated with the other regressors. If correlated (H_0 is rejected), a random effect model produces biased estimators, violating one of the Gauss-Markov assumptions. Consequently, this would suggest that a fixed effect model is preferred (Greene, 2003). However, there are differences in terms of generalization between the fixed and random effects models with fixed effects not being generalizable beyond the sample at hand. If a random effects model is chosen, a Breusch-Godfrey test for autocorrelation in the idiosyncratic component of the error term can be performed. If autocorrelation is present the random effects model will generate biased estimates.

Estimation of the fixed and random effects models can be further refined by using feasible generalized least squares (FGLS). A treatment of this model is provided in Wooldridge (2003) and in Cameron and Trivedi (2005). Estimation is based on a two-step procedure. First a model is estimated using OLS (random or fixed effects). The residuals are then used to estimate an error covariance matrix for use in the FGLS estimation. This setup is useful as it allows for the error covariance structure inside every group (state) of observations to be unrestricted. As a direct result FGLS is robust against any type of intra-group heteroskedasticity as well as autocorrelation (serial). However, the estimation is inefficient in the presence of group-wise heteroskedasticity. One drawback of this approach is that it requires the estimation of $T(T+1)/2$ variance parameters, where T is the number of time periods.

4.6 Spatial Panel Models

Recent developments in the econometric literature have brought to forefront the specification and estimation of spatial panel data models (Anselin, 1988, 2001; Baltagi, et al. 2003a, 2003b; Elhorst, 2001, 2003, Kapoor et al., 2007). Typical treatments are fixed and or random effects models with either the spatial error or spatial lag process. To date the so-called SARAR process has not been applied to the spatial panel models where both the error and lag are estimated jointly as in the cross-sectional case (e.g. Kelejian and Prucha, 2004). For now we apply the spatial process to either the error or lagged dependent variable as described in section 4.2. In order to incorporate this framework in a panel data model, the weight matrix W and its elements are often assumed to be constant over time (Anselin et al., 2006). The weights matrix then with W_N in the cross-sectional case becomes:

$$W_{NT} = I_T \otimes W_N \tag{4.8}$$

with I_T being an identity matrix of dimension T. Incorporating the panel weights matrix into the spatial error model (4.2) and spatial lag model (4.4) results in the following:

$$y = X\beta + \left[I_T \otimes (I_N - \lambda W_N)^{-1}\right]\mu \tag{4.9}$$

$$y = \left[I_T \otimes (I_N - \rho W_N)^{-1}\right]X\beta + \left[I_T \otimes (I_N - \rho W_N)^{-1}\right]\varepsilon \tag{4.10}$$

where dimensions of the matrices are noted by the subscripts. If unobserved heterogeneity occurs in panel data models then it is possible that spatial heterogeneity is a portion of that problem. As is well known, consistent estimation of the individual fixed effects is not possible when $N \to \infty$, because of the incidental parameter problem (Elhorst, 2003). Since spatial models rely on asymptotics in the cross-sectional setting to obtain consistency and asymptotic normality of estimators, concern has been raised about the appropriateness of spatial models in the fixed effects context (e.g. Anselin, 2001, Anselin et al 2006). Elhorst (2003) showed that the spatial fixed effects model could in fact be estimated by first *demeaning* the data followed by standard maximum likelihood techniques provided by Anselin (1988) and Anselin and Hudak (1992). However, the procedure still requires that the number of cross-sectional units does not get too large. A fixed effects spatial error and lag model as in equation 4.10 and 4.11 may be expressed as:

$$y = (\gamma_T \otimes \alpha) + X\beta + \left[I_T \otimes (I_N - \lambda W_N)^{-1}\right]\mu \tag{4.11}$$

$$y = \rho(I_T \otimes W_N)y + (\gamma_T \otimes \alpha) + X\beta + \varepsilon \tag{4.12}$$

where α is a $N \times 1$ vector of individual effects in both models. The standard assumptions of well-behaved error terms applies to these specifications.

Spatial random effects models have also been provided in the literature. Like the a-spatial case there must be no correlation between u_i as in equation 4.8. Elhorst (2003) finds this particular assumption restrictive suggesting a compelling argument for the fixed effects model irrespective of the size of *N* and *T*. Spatial panel data econometrics remains an on-going area of research. Our purpose here is not to extend the field, but to apply these recent methods in a practical and useful way. As such, a thorough discussion and survey of the relevant literature is provided in Anselin et al. (2006).

5. Empirical Results

The empirical models where estimated using the software packages R and MATLAB®. The results are reported in Tables 5.1 and 5.2 for a-spatial and spatial models respectively. Model 1 is a pooled cross-sectional model without time or spatial effects estimated using OLS. A Breusch-Pagan test rejected homoskedasticity, which is likely to do the vast differences in size of U.S. states. To correct for this a heteroskedasticity-consistent (HC) covariance matrix was estimated as described in Long and Irvin (2000) and Zeilies (2004). As expected, lagged investment, manufacturing share of employment, population, and productivity have positive coefficients and are statistically significant. Available land is statistically significant and has a negative coefficient indicating that having too much proportional farm land is a deterrent of investment. Unemployment rates were expected to positively contribute to investment location but in fact have a negative and significant coefficient. This may be as a result of the spatial level of aggregation. Since states represent a fairly large area, a high unemployment rate may signal those with a sluggish economy. Manufacturers may choose to invest where economic expectations are high. Similarly, state personal income tax had the unexpected positive sign. It was thought that high tax rates relative to other states would deter investment. This result could be due to the fact in the present sample states which do not use this tax have very small levels of manufacturing investment such as New Hampshire, Nevada, South Dakota, Tennessee, Washington, and Wyoming. However, these states may have other taxes that deter investment which are unaccounted for in the proposed models. Productivity had the largest coefficient with 1% increase in the value of output per hour leading to a 0.5% increase in investment *ceterus paribus*. Overall, the model explains approximately 95% of the variation in manufacturing investment from the variation in the covariates.

Model 2 extends the base model by including fixed time effects using dummy variables for each year. The year 1995 is used as the reference group. As before, homoskedasticity is rejected. The same correction procedure was applied as in Model 1. Most of the coefficients are the same compared to Model 1 with a few important exceptions to note. Production hourly wage was positive and statistically significant. It was expected *a priori* that higher wages would deter investment. However, the model does not have a measure for labor quality such as educational attainment. It is very likely that high wages are picking up the effect of skilled labor instead of being a cost measure. Also to note is that union participation rates have the expected negative coefficient and is statistically significant. The time effects are not reported in Table 5.1. However, they indicate a significant negative trend in investment beginning in 2002 and continuing so through 2006. This confirms the downward trend observed in the raw investment data following the recession of 2001. Population had the highest coefficient in this model indicating that a 1% increase in population would lead to a 0.6% increase in manufacturing investment.

Model 3 incorporates spatial heterogeneity as described in section 4.3 as well as time fixed effects in the pooled cross-sectional data. The regions are defined based on the results of the ESDA as discussed in section 3.2. Region 3, the South, is dropped as the reference group. To conserve space, the regional coefficients are not reported due to lack of statistical significance. The coefficient estimates did not change in sign or significance from Model 2 with the exception of union participation rates. It was no longer statistically significant. It was at this point that an unrestricted model was estimated as described by equation 4.6 in section 4.4 by allowing the intercepts and coefficients to vary by region with the exception of the time effects. A Chow test was conducted on the joint significance between the restricted and unrestricted model. The models were not statistically significantly different indicating that the coefficients and intercepts need not vary across space based upon the regional structure chosen. [2]

Beginning with Model 4, linear panel methods are used to estimate the impact of the location factors on manufacturing investment. Use of the fixed (spatial and time) effects model has a stronger theoretical argument as one would expect a high degree of heterogeneity across states. Spatial refers to individual state effects in the present study. A Hausman test was conducted to test between the fixed and random effects model. The null hypothesis of a consistent model between the two was not rejected indicating that the random effects model may be appropriate. However, a Breusch-Godfrey test was conducted on the residuals of the random effects model. The hypothesis was rejected indicating serial-correlation in the idiosyncratic portion of the error term.[3] Such a

[2] F-tests also failed to reject the hypotheses that individual coefficients were significantly different across space.
[3] This result was discovered for panel model specifications. As a result, only fixed effects models are reported.

condition makes the estimates of the beta coefficients biased. As a result a fixed effects models were estimated using space and time effects for Models 4-6. In the presence of heteroskedasticity a robust covariance matrix estimator *á la White* was used for the a-spatial panel models. Model 4 with spatial fixed effects gave almost the same results as Model 3, the expected signs for lagged investment, manufactures share of employment, population, and worker productivity. Model 5 was estimated with time fixed effects. Notable differences are significance of unemployment rates, and available land both deterring investment. Production wages were not significant in this specification. Model 6 incorporated both space and time fixed effects. The results were comparable to Models 3 and 4. One exception was the lack of significance for personal income tax rates. However, the coefficient still had the unexpected sign.

Starting with Model 7-10 in Table 5.2 the spatial process models were estimated for the pooled cross-section and panel data. Prior to the estimation of Model 7, spatial diagnostics were conducted on the first three pooled cross-sectional models. Moran's I calculated on the residuals of the models was not statistically significant at conventional levels. However, the Lagrange multiplier test was statistically significant for the spatial lag model in Models 1 and 2. As a result, the spatial lag model was estimated using MLE. A Breusch-Pagan test for homoskedastic error terms was rejected. Heteroskedasticity was corrected using the HC estimator of the variance co-variance matrix with corrected standard errors as in the a-spatial models. Generally speaking there are minimal differences between it and the previous specifications. Production wages were not statistically significant as in Model 1. However, worker productivity was with the coefficient suggesting that a 10% increase in productivity leads to a 5.5% increase in investment. The spatial lag parameter, ρ, was small (0.055) but statistically significant at the one percent level. The coefficient suggests that spillovers from states with high levels of investment to their neighbors may occur, but in a very small fashion. It is also evident that the spillover dies out rather quickly. A states investment raises 0.5% when its neighboring states raise their investment 10% *ceterus paribus*.

The spatial panel models are operationalized in MATLAB® using routines developed by Elhorst (2004) which are publicly available at his website and also at the spatial econometrics library in "Econometrics Toolbox" maintained by LeSage (2005). Following the earlier analysis of the pooled cross-sectional model, the SAL was identified as the appropriate spatial process. Using LM tests for panel models would be more appropriate. At present those tests are not readily available and are difficult to implement. For now we rely on the cross-sectional diagnostics. Empirical findings in the a-spatial panel models and the theoretical arguments discussed in sections 4.5 and 4.6 suggest that a fixed effect model is more appropriate than random effects. Model 8 is a SAL panel model including spatial fixed effects. Time lagged investment, manufacturing's share of employment, and productivity are statistically significant and have the expected signs. Similar to some of the first six models, unemployment rates and state personal income tax rates are significant but with the unexpected signs. A point of departure is the coefficient on available land which is positive statistically significant in this specification. The spatial auto-regressive coefficient ρ is not significant. Model 9 is a SAL panel with time fixed effects only. The results follow closely that of Model 2. Time lagged investment, manufacturing's share of employment, population, and productivity are significant with the expected sign. State expenditure per capita has a negative and significant coefficient suggesting increased government expenditure deters investment. Unemployment rates, available land, and state personal income tax rate are significant but with the unexpected signs. In this model, ρ is statistically significant at the 95^{th} percentile with a value of 0.048. As in Model 7, ρ provides some empirical evidence of spillovers of investment from state to state albeit rather small and likely die out quickly. Model 10 includes spatial and temporal fixed effects. The results are almost identical to the previous SAL fixed effects model with the exception of state per capita expenditures being statistically significant.

Overall, the empirical models provide robust results. The goodness of fit measures were comparably the same explaining about 94% of the variation in investment. Investment in the previous period, manufacturer's share of employment, population, and productivity all had significant and positive coefficients across most of the models. These are indications that local agglomeration, market size, and productive labor are attractions to manufacturing investment. Urbanization economies were not significant in every model, but the coefficients were negative in every case except the SAL panel models that included spatial fixed or spatial and time fixed effects. At first glance this may seem at odds with the available land hypothesis. However, it is likely that manufacturers do not necessarily want to locate in areas of high competition or in those which have a large proportion of open space. Gross state product per capita had the unexpected sign in six out of nine models but was only significant in Model 4. This seems to go against a large amount of literature which says that higher income regions should attract investment. This unexpected result could also be tied in the production wages being endogenously determined by education. Production wage was statistically significant in three of the nine models but had the unexpected sign. Unemployment was significant in seven of nine models but also with the unexpected sign as it is believed to be a proxy for available labor. This may indicate that investment goes to states with a healthy economic climate as opposed to one with abundant and cheap labor. It is more likely to have a different effect at a lower level of spatial aggregation. Transportation infrastructure was not significant across any of the models. This does not mean that transportation infrastructure is unimportant. A more logical explanation is that the measure, kilometers of Interstate Highway, is a poor proxy. State personal income tax rates were significant across all models but with the unexpected sign. It is difficult to come up with an explanation of this result. States in the sample which do not use state income taxes generally have low levels of investment, but they may in fact have high levels of other taxes that discourage investment, which are unaccounted for the proposed empirical models put forth.

<<Insert Table 5.2>>

6. Conclusions

This paper has revisited an on-going question of what location factors influence where manufacturing firms choose to invest. Location factors are likely to change in importance as methods of production change with technology and as markets mature. A conceptual model is put forth after an extensive review of the literature. Common location factors include agglomeration, market structure, labor, infrastructure, and fiscal policies. These factors are then proxied by measures used previously in the literature or in some cases with the best data at hand. Spatial diagnostics indicate a low and almost insignificant level of spatial dependence. Spatial spillovers are found to be positive and statistically significant, but rather small. A-spatial panel models are used to estimate fixed and random effects models. The results are similar to each other but both indicate remaining autocorrelation despite estimation via FGLS. Spatial panel models including fixed effects and a spatially lagged dependent variable provide comparable results to the cross-sectional SAL and a-spatial fixed effects panel models. Model results are fairly robust across the nine models estimated. Market structure was found to be the most important factor in investment location. On average, a 1% increase in population lead to a 0.5% increase in manufacturing investment which suggests that the manufacturing sector as a whole still prefers to locate near demand centers. One potential policy implication is that policy makers should focus economic development policies that attract people if they wish to attract manufacturing investment.

One obvious limitation in this study is the level of spatial aggregation. Analysis at the state level is usually not desirable. Heterogeneity within the each state is likely to be of the same magnitude as across states. However, manufacturing investment data (i.e. capital expenditures) at a lower level of spatial aggregation is not readily available. Low levels and insignificance of spatial dependence is also likely due to the high level of spatial aggregation. Additionally, the models are likely to suffer from omitted variable biased from excluding some measure of educational attainment. The data are also difficult to obtain on an annual basis. Future research should attempt to address these issues. Careful thought should also be given on how to best correct for simultaneous serial correlation and idiosyncratic autocorrelation due to unobserved factors that point to the heterogeneity across states. A further consideration is to test for heteroskedasticity in the spatial panel models and adjust the standard errors appropriately.

Future research should consider how location factors also effect manufacturing investment by sector and size of the firm over time. These are likely to be different if the firm is demand-oriented, supply-oriented, or footloose. The size of the firm is also likely to have an impact with smaller firms potentially being more inelastic to location factors.

References

Anselin, L. 1988. *Spatial Econometrics: Methods and Models.* Kluwer Academic Publishers.

______ (1995) "Local Indicators of Spatial Association – LISA." *Geographical Analysis* 27: 93 115.

______ (2003) "Spatial Externalities, Spatial Multipliers, and Spatial Econometrics." *International Regional Science Review* 26(2): 153-66.

______(2001) "Spatial Econometrics." In Baltagi, Badi (Eds.) *A Companion to Theoretical Econometrics* 310-330. Blackwell Oxford.

______ (2004) "GeoDa 0.9.5-i Release Notes." Spatial Analysis Laboratory. University of Illinois Urbana Champaign.

Anselin, L., A. Bera, R.JG.M. Florax, and M. Yoon (1996) "Simple Diagnostic Tests for Spatial Dependence." *Regional Science and Urban Economics* 26(1): 77-104.

Anselin, L., J. Le Gallo and H. Jayet (2006) "Spatial Panel Econometrics." in *The Econometrics of Panel Data, Fundamentals and Recent Developments in Theory and Practice* (3rd Edition), edited by L. Matyas and P. Sevestre.

Anselin, L., and S. Hudak (1992) "Spatial Econometrics in Practice: A Review of Software Options." *Regional Science and Urban Economics* 22: 509-36.

Baltagi, B.H., S.H. Song, B.C. Jung, and W. Koh (2003a) "Testing for Serial Correlation, Spatial Autocorrelation and Random Effects Using Panel Data." Working paper, Texas A&M University, College Station, TX.

Baltagi, B.H., S.H. Song, and W. Koh (2003b) "Testing Panel Data Regression Models with Spatial Error Correlation." *Journal of Econometrics* 117: 123–50.

Bartik, T.J. (1985) "Business Location Decisions in the United States: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States." *Journal of Business and Economic Statistics* 3(1): 14-22.

______ (1989) "Small Business Start-Ups in the United States: Estimates of the Effects of Characteristics of States." *Southern Economic Journal* 55(4): 1004-18.

Benson, B.L. and R. N. Johnson (1986) "The Lagged Impact of State and Local Taxes on Economic Activity and Political Behavior." *Economic Inquiry* 24(3): 389-401.

Berman, E., J. Bound, and Z. Griliches (1994) "Changes in Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufactures." *Quarterly Journal of Economics* 109(2): 367-97.

Bernat, G.A. (1996) "Does Manufacturing Matter? A Spatial Econometric View of Kaldor's Laws." *Journal of Regional Science* 36(3): 463-77.

Blair, J.P. and R. Premus (1987) "Major Factors in Industrial Location: A Review." *Economic Development Quarterly* 1(1): 72-85.

Calzonetti, F.J. and R.T. Walker (1991) "Factors Affecting Industrial Location Decisions: A Survey Approach." in *Industry Location and Public Policy*, edited by Herzong, H.W., and A.M. Schlottmann, University of Tennessee Press, Knoxville, pp. 221-40.

Cameron, A.C. and P.K. Trivedi (2005) *Microeconometrics: Methods and Applications*. Cambridge University Press, New York, NY.

Carlson, V. (2000) "Studying Firm Locations: Survey Responses vs. Econometric Models." *Journal of Regional Analysis & Policy* 30(1): 1-22.

Carlton, D.W. (1983) “The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables.” *Review of Economics and Statistics* 14:209-20.

Cohen, J.P. and C.M. Paul (2004) “Agglomeration Economies and Industry Location Decisions: The Impacts of Spatial and Industrial Spillovers.” *Regional Science and Urban Economics* 35: 215-37.

Crone, T.M. (1997) “Where Have All the Factory Jobs Gone – And Why?” *Business Review, Federal Reserve Bank of Philadelphia*. May/June.

Crone, T.M. (2000) “A New Look at Economic Indexes for the States in the Third District,” *Business Review, Federal Reserve Bank of Philadelphia*. November/December.

Davis, E.D., and G.E. Schluter (2005) “Labor-Force Heterogeneity as a Source of Agglomeration Economies in an Empirical Analysis of County-Level Determinants of Food Plant Entry.” *Journal of Agricultural and Resource Economics* 30(3): 480-501.

Elhorst, J.P. (2001) “Dynamic Models in Space and Time.” *Geographical Analysis* 33:119-40.

______(2003) “Specification and Estimation of Spatial Panel Data Models.” *International Regional Science Review* 26(3): 244-68.

______(2004) “MATLAB Software to Estimate Spatial Panels.” available at http://www.regroningen.nl/index_en.html.

Gerking, S. and W. Morgan (1991) “Measuring Effects of Industrial Location and State Economic Development Policy: A Survey,” in *Industry Location and Public Policy*, edited by Herzong, H.W., and A.M. Schlottmann, University of Tennessee Press, Knoxville, pp. 31-56.

Glaeser, E.L., H.D. Kallal, J.A. Scheinkman, and A. Shleifer (1992) “Grow in Cities.” *Journal of Political Economy* 100(6): 1126-51.

Goetz, S. (1997) “State- and County-Level Determinants of Food Manufacturing Establishment Growth: 1987-1993.” *American Journal of Agricultural Economics* 79(3): 838-50.

Greene, W.H. (2003) *Econometric Analysis*, 5th ed. NJ: Prentice Hall.

Greenhut, M.L. (1956) *Plant Location in Theory and in Practise*. University of North Carolina Press,
Chapel Hill.

Guimarães, P., O. Figueiredo, and D. Woodward (2004) “Industrial Location Modeling: Extending the Random Utility Framework.” *Journal of Regional Science* 44(1): 1-20.

Gupta, S. and M.A. Hofmann (2003) “The Effect of State Income Tax Apportionment and Tax Incentives on New Capital Expenditures.” *Journal of the American Taxation Association* 25: 1-25.

Hausman, J.A. (1978) “Specification Tests in Econometrics.” *Econometrica* 46(6): 1251-71.

Hekman, J.S. (1982) “Survey of Location Decisions in the South.” *Economic Review* (June):6-19.

Henderson, J.R. and K.T. McNamara (1997) “Community Attributes Influencing Local Food Processing Growth in the U.S. Corn Belt.” *Canadian Journal of Agricultural Economics* 45: 235-50.

______ (2000) “The Location of Food Manufacturing Plant Investments in Corn Belt Counties.” *Journal of Agricultural and Resource Economics* 25(2): 680-97.

Holl, A. (2000) “Manufacturing Location and Impacts of Road Transport Infrastructure: Empirical Evidence from Spain.” *Regional Science & Urban Economics* 34: 341-63.

Hoover, E.M. (1948) *The Location of Economic Activity*. McGraw-Hill, New York.
Hotelling, H. (1929) "Stability in Competition." *The Economic Journal* 39(153): 47-57.
Isard, W. (1948) "Some Locational Factors in the Iron and Steel Industry Since the Early Nineteenth Century." *Journal of Political Economy* 56(3): 203-17.
Jacobs, J. (1969) *The Economy of Cities*. New York: Vintage.
Jones, L.L. and M.D. Woods (2002) "Economic Location Theory and Practice." Texas A&M University.
Kapoor, M., H.H. Kelejian, and I.R. Prucha (2007) "Panel Data Models with Spatially Correlated Error Components." *Journal of Econometrics* 140: 97-130.
Kelejian, H. and I. Prucha (2004) "Estimation of Simultaneous Systems of Spatially Interrelated Cross Sectional Equations." *Journal of Econometrics* 118:27-50.
Kennedy, P. (2003) *A Guide to Econometrics*. 5th ed. MIT Press.
Kieschnick, M. (1981) "Taxes and Growth: Business Incentives and Economic Development." Council of State Planning Agencies, Washington D.C.
Klier, T. and D.P. McMillen (2005) "Clustering of Auto Supplier Plants in the U.S.: GMM Spatial Logit for Large Samples." Federal Reserve Bank of Chicago. WP2005-18.
Lambert, D.M., M.I. Garret, and K.T. McNamara (2006a) "An Application of Spatial Poisson Models to Manufacturing Investment Location Analysis." *Journal of Agricultural and Applied Economics* 38(1): 105-121.
______ (2006b) "Food Industry Investment Flows: Implications for Rural Development." *The Review of Regional Studies* 36(2): 140-62.
LeSage, J.P. (2005) Econometrics Toolbox for MATLAB. Available at http://www.spatial-econometrics.com/.
Long, J.S., and L.H. Ervin (2000) "Using Heteroskedasticity Consistent Standard Errors in the Linear Regression Model." *American Statistical Association* 54(3): 217-24.
Lösch, A. (1954) *The Economics of Location*. Yale University Press, New Haven.
McCann, P. (2001) *Urban and Regional Economics*. Oxford University Press.
McGranahan, D.A. (2000) "The Geography of New Manufacturing Technology: Implications for the Nonmetropolitan South." *Southern Rural Sociology* 15: 84-103.
McLaughlin, G.E. and S. Robock (1949) *Why Industry Moves South*. NPA Committee of the South. Report No.3. Kingsport Press, Inc.
Muller, E. and J.N. Morgan (1962) "Location Decisions of Manufacturers." *American Economic Review* 52(2): 204-17.
Newman, M. and M. Gaster (2004) Proceedings of the National Academy of Science, Vol 101, Number 20, pp 7499-7504.
Papke, L.E. (1987) "Subnational Taxation and Capital Mobility: Estimates of Tax-Price Elasticities." *National Tax Journal* 40(2): 191-203.
Papke, L.E. (1991) "Interstate Business Tax Differentials and New Firm Location: Evidence from Panel Data." *Journal of Public Economics* 45: 47-68.
Plaut, T.R. and J.E. Pluta (1983) "Business Climate, Taxes and Expenditures, and State Industrial Growth in the United States." *Southern Economic Journal* 50: 99-119.
Porter, M.E. (1990) *The Competitive Advantage of Nations*. New York: Free Press.
Schmenner, R.W., J.C. Huber, and R.L. Cook (1987) "Geographic Differences and the Location of New Manufacturing Facilities." *Journal of Urban Economics* 21: 83-104.

Smith, E.D., B.J. Deaton, and D.R. Kelch (1978) "Location Determinants of Manufacturing Industry in Rural Areas." *Southern Journal of Agricultural Economics* 10: 23-32.

Tobler, W. (2004) Thirty-Five Years of Computer Cartograms. *Annals of the Association of American Geographers*, Vol. 94 (1), March, pp. 58-73.

Thünen, J.H. von, (1826) *Der Isolierte Staat in Bezeihung auf Landwirtschaft und Nationalökonomie*. Hamburg.

Wasylenko, M. and T. McGuire (1985) "Jobs and Taxes: The Effects of Business Climate on States' Employment Growth Rates." *National Tax Journal* 38:497-511.

Weber, A. (1929) *Theory of the Location of Industries*. translated by C. Friedrich. University of Chicago Press, Chicago.

Wheat, L. (1986) "The Determinants of 1963-77 Regional Manufacturing Growth: Why the South and West Grow." *Journal of Regional Science* 26(4): 635-60.

Wissen, L. van (2000) "A Micro-Simulation Model of Firms: Applications of Concepts of the Demography of the Firm." *Papers in Regional Science* 79: 11-34.

Woodward, D.P. (1992) "Location Determinants of Japanese Manufacturing Start-Ups in the United States." *Southern Economics Journal* 58: 690-708.

Woodward, D.P. and N. Glickman (1991) "Regional and Local Determinants of Foreign Firm Location in the United States." in *Industry Location and Public Policy*, edited by Herzong, H.W., and A.M. Schlottmann, University of Tennessee Press, Knoxville, pp. 190-220.

Wooldridge, J.M. (2002) *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Zeiles, A. (2004) "Econometric Computing with HC and HAC Covariance Matrix Estimators." *Journal of Statistical Software* 11(10): 1-16.

Table 3.1

Variable	Definition	Mean	Std Dev.	Min	Max
inv[1]	Manuf. investment ($ Millions)[*]	3,409.26	4,108.99	49.67	53,129.12
memp[2]	Manuf. share of total employment	0.12	0.05	0.03	0.26
ue[2]	Establishment density (no. per square km)	1.86	2.63	0.07	12.58
pwage[1]	Prod. employ.. average hourly wage ($ per hour)	17.52	1.85	13.01	23.41
product[1]	Value of production per hour ($ Thous)	220.49	71.16	114.51	855.76
pop[3]	Population (in Thous)	5,846.21	6,236.83	485.16	36,249.87
gspc[4]	Gross state product per capita ($ Thous)	38.38	7.07	26.09	70.78
unempr[5]	Unemployment rates	4.70	1.10	2.30	8.10
union[5]	Unionization rates	11.81	5.34	2.30	27.70
interst[6]	Kilometers of interstate highway	1,514.75	936.87	64.37	5,204.48
avland[7]	Farm area / total area	0.42	0.25	0.06	0.94
pit[4]	State personal income tax rate	2.14	1.10	0.00	4.48

* All currency values are in real dollars, [1] Annual Survey of Manufactures 1994-2006, [2] County B.P., U.S. Census Bureau

[3] Population Division, U.S. Census Bureau, [4] Bureau of Economic Analysis, Regional Economic Information System

[5] Bureau of Labor Statistics, [6] U.S. Department of Transportation, [7] National Agricultural Statistics Services

Table 5.1 – A-Spatial Pooled Cross-Section and Panel Models

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	FGLS	FGLS	FGLS	FGLS	FGLS	FGLS
Constant	-1.463***	-2.644***	-2.230***	-----	-----	-----
	(0.498)[a,b]	(0.655)	(0.663)			
$lninv_{t-1}$	0.486***	0.431***	0.424***	0.428***	0.431***	0.383***
	(0.076)	(0.086)	(0.087)	(0.120)	(0.105)	(0.114)
lnmemp	0.527***	0.542***	0.542***	0.543***	0.606***	0.592***
	(0.093)	(0.099)	(0.106)	(0.128)	(0.114)	(0.119)
lnue	-0.024	-0.034	-0.012	-0.037	-0.016	-0.017
	(0.024)	(0.025)	(0.034)	(0.024)	(0.028)	(0.026)
lngdpc	-0.158*	-0.010	0.019	-0.063	-0.232	-0.135
	(0.095)	(0.097)	(0.116)	(0.089)	(0.149)	(0.173)
lnpop	0.503***	0.558***	0.526***	0.565***	0.541***	0.570***
	(0.082)	(0.090)	(0.095)	(0.121)	(0.114)	(0.119)
lnpwage	0.107	0.488**	0.420**	0.498***	0.068	0.537**
	(0.177)	(0.191)	(0.186)	(0.183)	(0.214)	(0.223)
lnprod	0.542***	0.436***	0.447***	0.464***	0.541***	0.440***
	(0.106)	(0.113)	(0.112)	(0.131)	(0.118)	(0.116)
lnunempr	-0.174***	-0.029	-0.063	-0.052	-0.235***	-0.111
	(0.059)	(0.080)	(0.075)	(0.085)	(0.068)	(0.083)
lnunion	0.040	-0.062*	-0.026	-0.046	0.062**	-0.029
	(0.029)	(0.033)	(0.036)	(0.027)	(0.031)	(0.056)
lninfr	0.022	0.019	0.059	0.018	0.028	0.060
	(0.055)	(0.055)	(0.071)	(0.046)	(0.048)	(0.050)
lnavland	-0.074**	-0.049	-0.057	-0.053	-0.058*	-0.043
	(0.034)	(0.035)	(0.036)	(0.036)	(0.034)	(0.035)
lnpit	0.010*	0.012**	0.012**	0.013**	0.006	0.008
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)
lnstec	-0.158	-0.002	-0.015	0.012	-0.130	0.044
	(0.111)	(0.112)	(0.112)	(0.122)	(0.125)	(0.134)
Regional dummies	no	no	yes	no	no	no
Time fixed effects	no	yes	yes	no	yes	yes
Spatial fixed effects	no	no	no	yes	no	yes
Adjusted R^2	0.949	0.952	0.952	0.954	0.938	0.940
F-statistic	828***	472***	419***	813***	639***	604***
Moran's I	0.01	-0.04	-0.04			
LM-error	0.07	1.44	2.12			
LM-lag	8.99***	2.85*	2.47			
Robust LM-error	0.574	3.44*	4.61**			
Robust LM-lag	9.49***	4.84**	4.95**			
LM-SARMA	9.57***	6.29**	7.08**			

[a] Standard errors are in parentheses. Significance is at the 1, 5, and 10% level as noted by, ***, **, and *, respectively.

[b] Standard errors adjusted for heteroskedasticity using HC estimator of variance co-variance matrix.

Table 5.2 – Spatial Pooled Cross-Section and Panel Models

	Model 7 MLE	Model 8 MLE	Model 9 MLE	Model 10 MLE
Constant	-2.214***	----	----	----
	(-4.790)[a]			
$lninv_{t-1}$	0.468***	0.235***	0.465***	0.245***
	(6.862)	(5.996)	(13.163)	(6.438)
lnmemp	0.517***	0.699***	0.528***	0.665***
	(6.146)	(4.710)	(10.892)	(4.617)
lnue	-0.030	0.573	-0.024	0.680
	(-1.337)	(0.871)	(-1.115)	(1.059)
lngdpc	-0.086	0.202	-0.171	0.183
	(-0.946)	(0.562)	(-1.604)	(0.514)
lnpop	0.492***	0.381	0.498***	0.238
	(6.617)	(0.518)	(10.173)	(0.334)
lnpwage	0.079	0.231	0.075	0.298
	(0.170)	(0.764)	(0.467)	(0.287)
lnprod	0.537***	0.320***	0.567***	0.326***
	(0.100)	(2.730)	(8.162)	(3.028)
lnunempr	-0.181***	-0.285***	-0.224***	-0.311***
	(0.056)	(-3.709)	(-4.411)	(-4.442)
lnunion	0.046*	0.078	0.071**	0.075
	(0.027)	(0.676)	(2.523)	(0.658)
lninfr	0.050	0.049	0.049	0.053
	(0.051)	(0.089)	1.163)	(0.097)
lnavland	-0.084**	2.106***	-0.080***	2.061***
	(0.032)	(3.514)	(-3.502)	(3.463)
lnpit	0.012**	0.434***	0.012**	0.469***
	(0.005)	(3.107)	(2.275)	(3.453)
lnstec	-0.165	-0.329	-0.213**	-0.399**
	(-2.00)	(-1.585)	(-2.541)	(-2.051)
ρ	0.055***	0.002	0.048**	0.005
	(2.93)	(0.037)	(2.527)	(0.111)
Spatial fixed effects	no	yes	no	yes
Time fixed effects	no	no	yes	yes
Adjusted R^2	0.950	0.957	0.947	0.955
Log-Likelihood	-28.010	38.200	-40.194	36.300

[a] Asymptotic t-statistics are in parentheses. Significance is at the 1, 5, and 10% level as noted by, ***, **, and *, respectively.

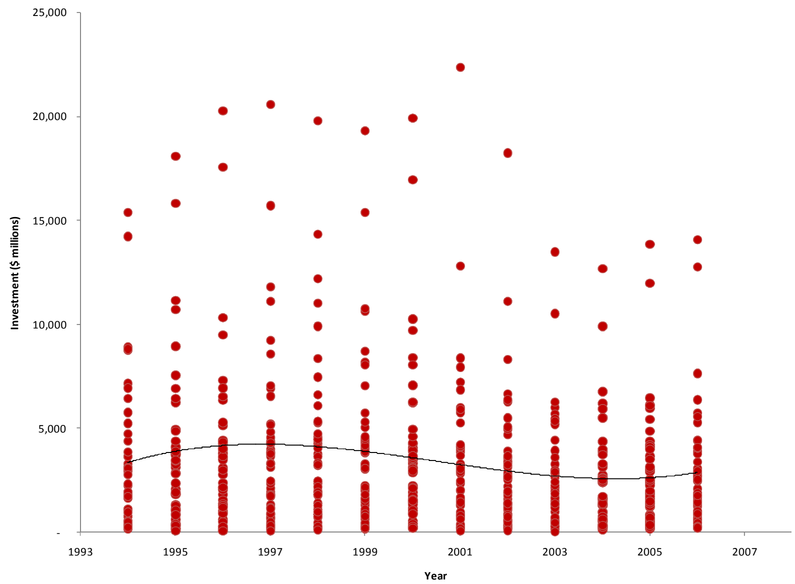

Figure 3.1 – Real Manufacturing Investment

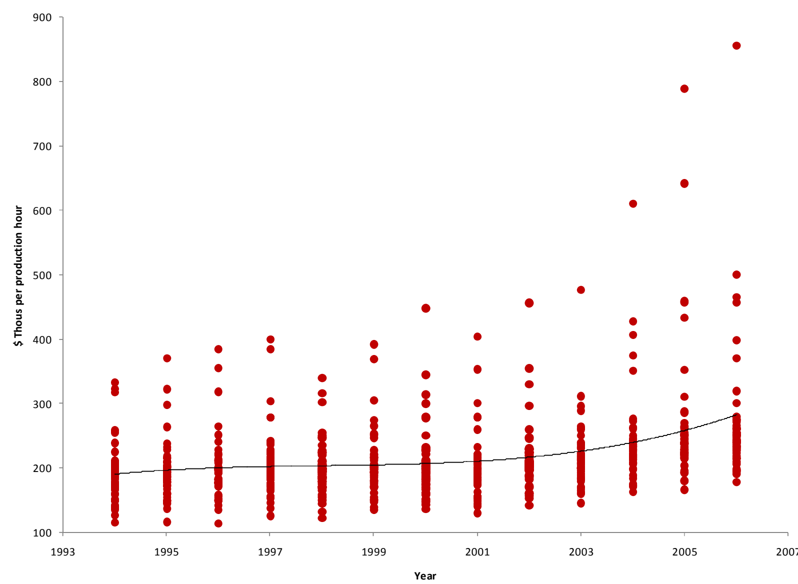

Figure 3.2 – Value of Output per Production Hour

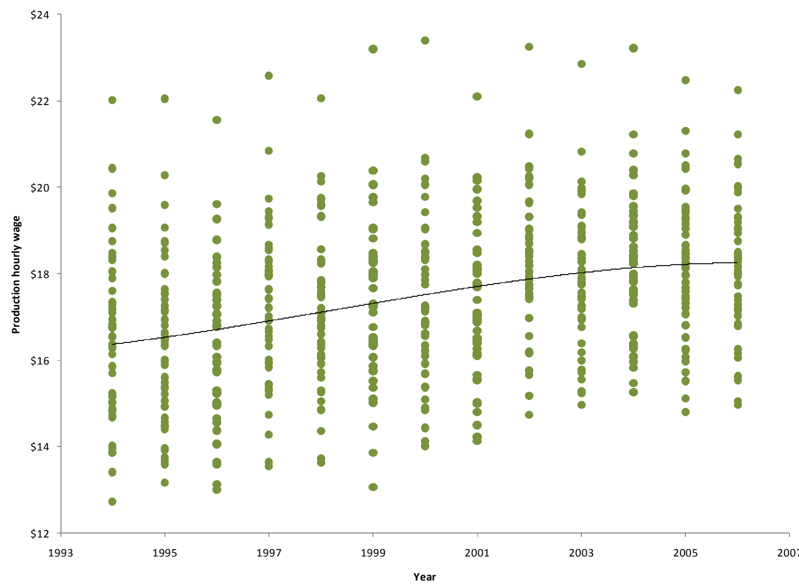

Figure 3.3 – Average Hourly Production Wage

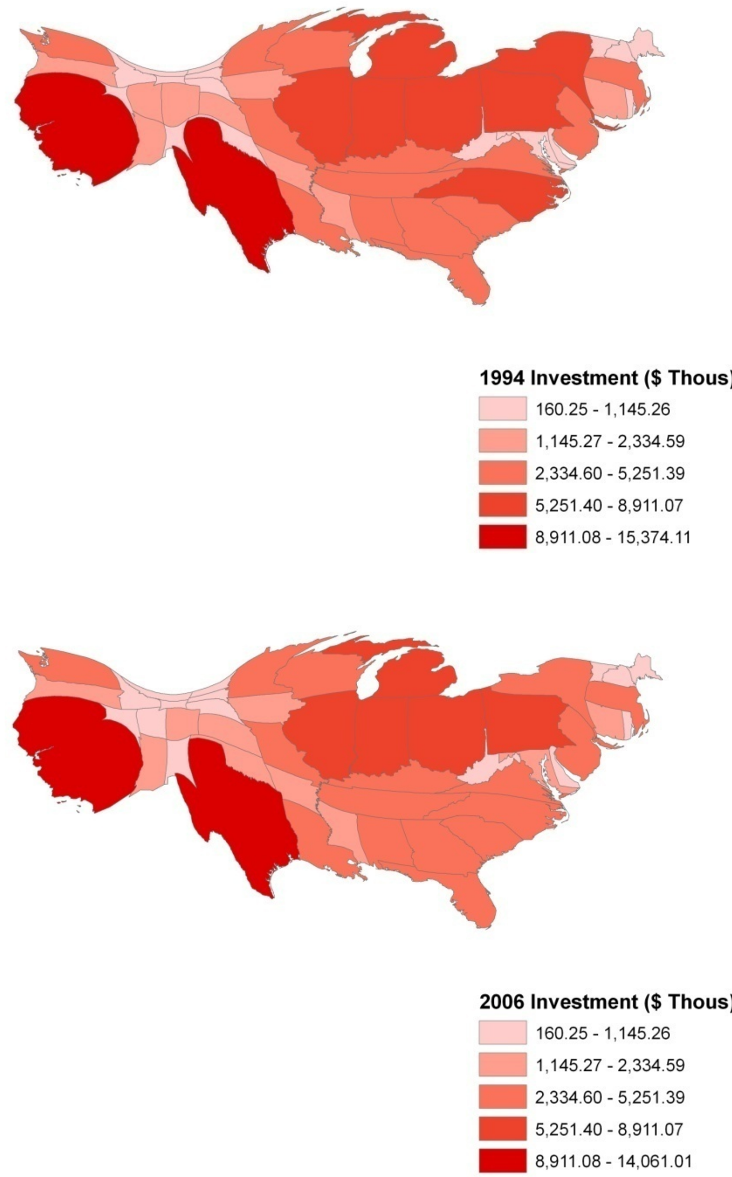

Figure 3.4 – Cartograms of U.S. Manufacturing Investment

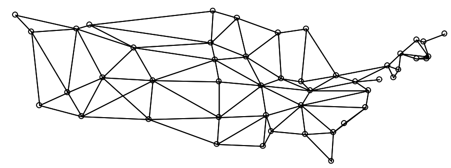

Figure 3.5 – Connectivity of Spatial System

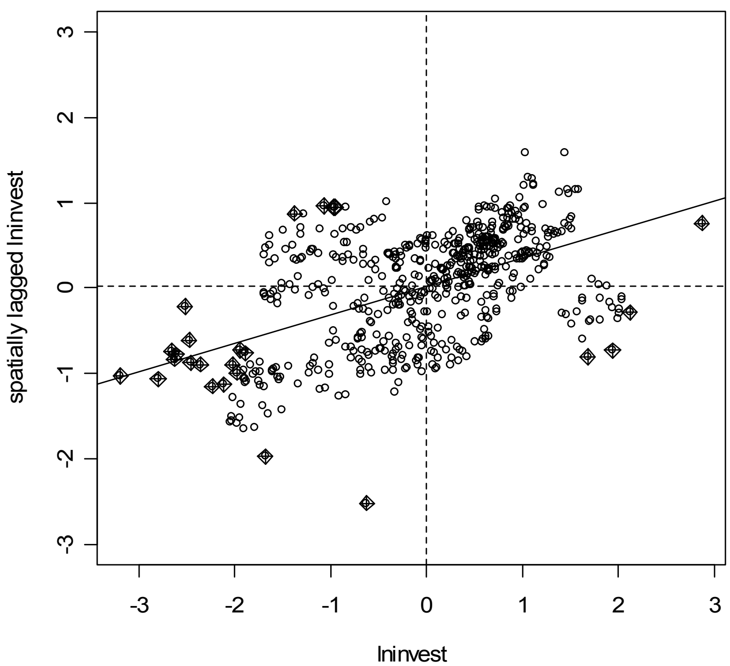

Figure 3.6 – Moran's I Scatterplot

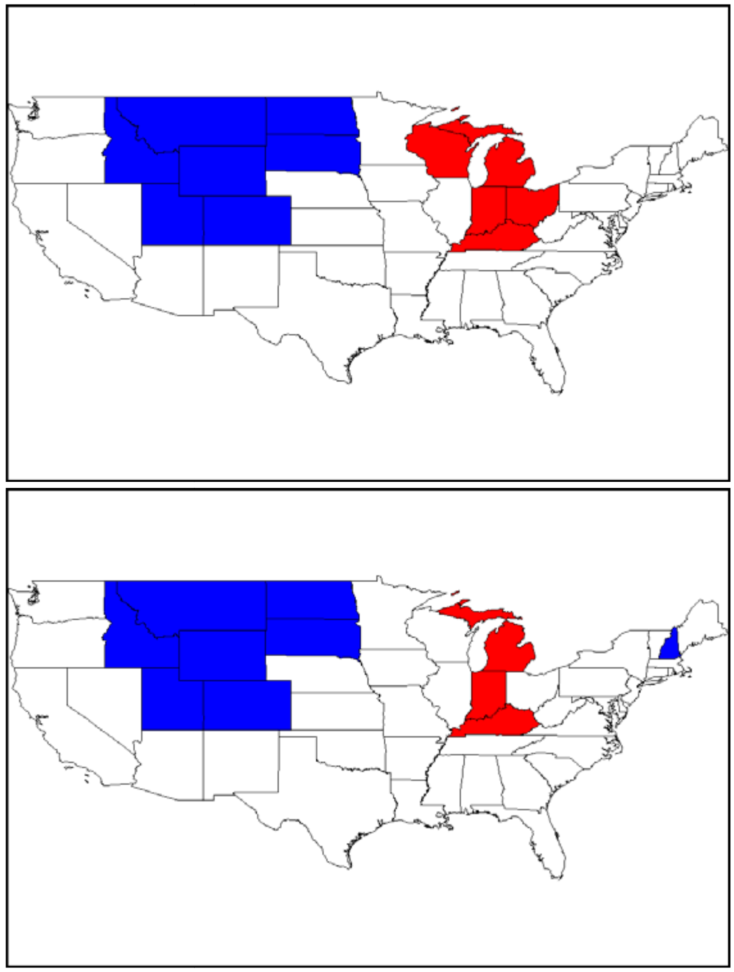

Figure 3.7 – Local Indicators of Spatial Autocorrelation (top 1994, bottom 2006)
High values (red/grey), Low values (blue/black)