2007

Spatial econometric analysis of agglomeration economies associated with the geographical distribution of the U.S. biotech industry

Pramod Reddy Sambidi
Louisiana State University and Agricultural and Mechanical College, psambi1@lsu.edu

SPATIAL ECONOMETRIC ANALYSIS OF AGGLOMERATION ECONOMIES ASSOCIATED WITH THE GEOGRAPHICAL DISTRIBUTION OF THE U.S. BIOTECH INDUSTRY

A Dissertation
Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in
The Department of Agricultural Economics & Agribusiness

by
Pramod R. Sambidi
B.Sc., ANGRAU, 1999
M.S., Louisiana State University, 2003
May 2007

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank everyone who contributed to the successful completion of this dissertation. First of all, I would like to thank my major professor, Dr. Wes Harrison, for his invaluable support, guidance and patience throughout my program of study. His role in my career development is vital, and I sincerely appreciate that. I have greatly enjoyed working with him and look forward to any opportunity that I may have to work with him in the future.

Sincere thanks are extended to my graduate committee members, Dr. Hector Zapata, Dr. Lynn Kennedy, Dr. Matthew Fannin, Dr. Douglas McMillin, and Dr. Yan Chen for their constructive criticism and recommendations for this study. Grateful acknowledgement is also expressed to Dr. Gail Cramer, head of the department of agricultural economics and agribusiness, the faculty, and staff, for helping me complete my research and studies here. I would also like to thank my fellow graduate students for their help and friendship throughout the course of my study. I especially like to thank Anil, James, Brian, Rebecca, Tina, Lee, Sachin, Pawan, Gus, Jose and Dwi for their invaluable friendship.

I would like to thank my friends for being there with their support during the times I needed the most. I am very grateful to my parents, Anitha and Indrasena Reddy, and to all my family members for their never ending love and support, and I could not have done it without their love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This dissertation analyzed spatial agglomeration economies associated with the geographical distribution of the U.S. biotech industry. Three location issues associated with the biotech industry were addressed in the study. The first study utilized a Bayesian spatial tobit model and examined the overall and regional differences in factors affecting the location of the U.S. biotech industry. The second study examined the inter- and intra-industry spatial association of biotech related research and development (R&D) and testing facilities across all contiguous U.S. counties employing a Spatial Two-Stage Least Squares model. Finally, the interdependence between different subsectors of the U.S. biotech industry was analyzed using a Seemingly Unrelated Regression model.

The first study confirmed the hypothesis of spatial agglomeration for the spatial structure of the biotech industry, indicating that biotech firms are positively correlated across counties, resulting in clustering of biotech production. Availability of venture capital firms, research institutions, and hospitals were found to have the most significant impact on the location of biotech firms. Results from regional models indicate that biotech firms willing to locate in the West prefer to establish in metro-counties with easy access to research institutes and skilled labor pool. Conversely, firms that are willing to locate in the Northeast prefer counties with easy access to funding sources and hospitals for research, testing and marketing of new biotech products.

Spatial clustering of biotech research and testing activities was confirmed in the second study. Proximity to manufacturing firms and research universities, and availability of venture capital firms were found to have the most significant impact on the location of R&D and testing facilities. Results indicated that public as well as private spillovers are at work in the R&D and testing industry, resulting in their spatial clustering.

Agricultural biotechnology firms' preference to locate in counties with large farmland, low median housing values and average hourly wage, and a high

unemployment rate was indicated in the third study. Conversely, results indicate that firms belonging to drug and pharmaceuticals, and medical devices and equipment subsectors prefer to locate in counties with high standards of living and in close proximity of research institutes and hospitals to access skilled-labor, and develop and test new drugs.

CHAPTER 1

INTRODUCTION

Over the past few decades, state and local planners as well as spatial and economic scientists have been interested in understanding the forces that contribute to the clustering of innovative activities and growth potential of firms and regions. A significant body of industry location literature focuses on economic externalities (technical and pecuniary externalities) and knowledge spillovers to analyze the reasons behind industry clusters (for recent review see Xia and Buccola, 2005; van Oort, 2004; Zuker, Darby, and Armstrong, 2003; Feldman, 2003; and Goetz and Rupasingha, 2002). Barkley and Henry (2001) defined industry cluster as "a loose, geographically bounded collection of similar and/or related firms that together create competitive advantages for member firms and the regional economy." In general, there are two types of industry clusters: intra- and inter-industry clusters. The concept of intra-industry clusters is associated with the collaboration of firms of the same industry that utilize similar technologies and face common problems. These firms tend to share information on a broad variety of issues from problem solving to the development of new production techniques (Lall, Koo, and Chakravorty, 2003). Conversely, inter-industry clusters are associated with collaboration of firms between different industries that are connected through buyer-supplier chains or shared factors (Lall, Koo, and Chakravorty, 2003). Scorsone (2002) classified industrial clusters into two types: one is Value Chain, which include groups of businesses that buy and sell from each other, and the second is Labor Pool, where clusters are based on occupational categories. One of the main reasons for the formation of industry clusters is to lower the overall transaction costs associated with the production, marketing, and distribution of outputs (Porter, 1990). According to

Barkley and Henry (2001), the advantages of industrial clusters are that they strengthen localization economies, facilitate industrial reorganization, encourage networking among firms, and permit greater focusing of public resources. However, the authors also point out some of the shortcomings with industrial clusters which include difficulty in identifying the targeted industry, competitive advantage of early clusters over latecomers, and difficulty in establishing supportive institutions. Some of the other disadvantages with industrial cluster development are that they tend to increase local land rents, wages, congestion, and utility costs, eventually diverting new firms away from the region.

While several studies have analyzed the role of agglomeration forces (economic externalities and knowledge spillovers) on industry location, not much is yet understood on the spatial extent of agglomeration. First of all, we want to know whether the agglomeration forces are significant in the location of an industry or not, and if significant, how far (distance) are they effective? It is critical to analyze the concept of agglomeration forces based on space, distance and spatial dependence (van Oort, 2004). Data obtained from points in space are generally associated with spatial dependence and heterogeneity. The concept of spatial dependence is a functional relationship between two locations, indicating that what happens at one point in space is determined by what happens elsewhere in the system (Kaliba, 2002). This spatial dependence is considered to decrease as the distance between the points increase. In terms of the agglomeration forces, economies of scale associated with a firm located in a neighboring (contiguous) county are supposed to be greater as compared to other firms located in non-neighboring counties. The second issue related to spatial data is spatial heterogeneity, which deals with variation in relationships over space (LeSage, 1999). Neglecting these spatial concepts in econometric modeling will result in biased estimates, leading to erroneous

interpretation and wrong conclusions (Anselin, 1988; LeSage 1999). This indicates that it is important to take into account spatial dependence and heterogeneity when analyzing the affects of agglomeration economies on industry location. Generally speaking, most of the previous studies on industry location have failed to incorporate spatial components in their econometric modeling (for recent exceptions see Roe, Irwin, and Sharp, 2002; Goetz and Rupasingha, 2002; Isik, 2004; van Oort, 2004; and Sambidi and Harrison, 2005).

The biotech industry, which is one of the fastest growing industries in the U.S.(Biotechnology Industry Organization (BIO), 2005), provides us with an opportunity to analyze the spatial concepts of industry location. It has all the necessary characteristics for a thorough understanding of the growth of a new industry. These characteristics include: i) the industry being in its early stages of development, ii) association of biotech firms in intra- and inter-industry clustering, and iii) the industry's uncertain future development in terms of locating new biotech firms (Feldman 2003).

1.1 The U.S. Biotech Industry Facts

The biotech industry is defined as "the application of biological knowledge and techniques to develop products and services" (BIO 2005). Goetz and Morgan (1995) defined it as "any technique that uses living organisms to make/modify products, improve plants or animals, or develop microorganisms for a specific use." Biotech firms are mainly research and development (R&D) oriented and operate in collaboration with research-oriented universities, biomedical research centers, and other diversified companies that aid in the production and distribution of biotech products. Biotech products may be characterized as drugs and pharmaceuticals, agricultural, and environmental, which aid in improving the quality of health, increasing the production of

agricultural goods, improving food quality, minimizing environmental hazards and providing a cleaner environment.

The biotech industry, which emerged during 1973 (Feldman, 2003), is one of the fastest growing industries in the United States, with increasing sales from $7.7 billion in 1994 to $33.3 billion in 2004 (Ernst and Young, 2005). According to the BIO (2005), in 2003 there were 1,473 biotechnology companies in the U.S., employing 198,300 people and spending $17.9 billion on research and development. The top five states in terms of number of biotech companies are: California (420), Massachusetts (193), North Carolina (88), Maryland (84), and New Jersey (77) (Ernst and Young, 2004). The biotech industry is mainly concentrated in nine cities/regions (San Francisco Bay Area, Boston/Cambridge, San Diego, Los Angeles, New York, Philadelphia, Raleigh-Durham, Seattle and Washington, DC), which account for three-fourths of the nation's largest biotechnology firms and for three-fourths of the biotech firms formed in the past decade (Cortright and Mayer, 2002).

As a result of the increasing success of the biotech industry, several state and local economic development agencies are designing and implementing strategies to attract new biotech firms, resulting in stiff competition among and within states. For example, in 2004, 40 states have adopted strategies to stimulate the growth of biotechnology and 50 states have technology based economic development initiatives for biotech firms, compared to merely 14 states in 2001 (Battelle and State Science and Technology Institute (SSTI), 2004). A survey of 77 local and 36 state economic development departments indicated that 83% have listed biotechnology as one of the top two targets for industrial development (Grudkova, 2001; Cortright and Mayer, 2002).

1.2 Problem Statement

The rationale for concentration of the U.S. biotech industry in California and the Northeast has been attributed to proximity to highly research-oriented universities, research parks and laboratories, and well-developed infrastructure. However, our understanding of the spatial influence (spatial dependence and heterogeneity) on the intra- and inter-regional distribution of biotech establishments is anecdotal. Moreover, as a result of increased competition in attracting biotech firms and increasing commercial applications, the U.S. biotech industry is undergoing changes in its geographical distribution. For example, during 1991-2001, while the concentration of biotech firms in the top states increased slightly, the relative ordering of states changed with New Jersey dropping below Maryland and North Carolina in terms of the number of establishments (Feldman, 2003). Moreover, a study conducted by McCandless (2005) indicated that Dallas, Houston, San Antonio, Memphis, Richmond, and Miami-Dade are making significant progress in becoming important biotech regions of the future.

These changes in the spatial distribution of the biotech industry as a result of increasing commercial applications indicate that proximity to universities is no longer a sufficient condition to promote the biotech industry cluster (Feldman, 2003). Earlier, biotech firms were considered to locate close to universities with star scientists based on the observation that biotechnology discoveries are characterized by tacit knowledge, which is best communicated through face-to-face contact (Feldman, 2003). However, as a result of technology advancements, even tacit knowledge is now considered to spread quickly (Toole, 2003; Feldman, 2003). In addition, the previous literature on the biotech industry location has not taken into consideration the extent of inter-industry spatial

clustering of biotech firms and research institutes. Analyzing this will improve our understanding of the scale of spatial dependency of biotech firms on research institutes.

The geographical changes in the biotech industry may also be attributed to other location factors. Generally speaking, the location decision is always associated with trade-off between location factors to lower the costs associated with procuring raw materials and producing and distributing the final product. In order to analyze the importance of trade-offs between location factors, let us briefly discuss the case of the U.S. broiler industry (mature industry), which has undergone a drastic change in its geographic distribution (Sambidi and Harrison, 2005). During the 1940s, the broiler industry was mainly concentrated in the Midwest, which has the advantage of low cost feed ingredients and low product transportation cost to large Midwestern cities. However, as a result of increasing land costs and wages in the Midwest, the broiler industry started to move to the South, which is associated with low land and labor costs. This movement also benefited from low rail rates associated with transporting feed from the Midwest to South. Currently, 85% of the U.S. broiler production is concentrated in the South (Sambidi and Harrison, 2005), thus indicating the significance of trade-offs between factors in industry location-decision. This typical example may open some doubts on the future of the spatial distribution of the biotech industry, since it is still at an early stage of development, and we have already started to see some geographical changes in its location. In addition, the current biotech locations are associated with congestion and high land, labor, and utility costs, prompting biotech companies to look for other sites with a comparatively more space and less costs (Munroe, Craft, and Hutton, 2002). Therefore, it is critical to analyze other location factors and resource endowments that may pull biotech firms away from research institutes. Additionally, it is also important to

analyze the statistical implications of the spatial clustering of biotech firms, which was ignored by earlier studies (except Goetz and Rupasingha, 2002).

1.3 Justification

The present study focuses on spatial economies (geographically determined externalities) associated with county-level biotech establishments. Involving spatial component in the biotech industry location analysis will aid in capturing important intangible aspects concerning spatial dependence and heterogeneity, which was not acknowledged by most of the previous studies. This spatial analysis will also aid in identifying some of the intermediate industries (such as research institutes, hospitals, and venture capital) that may cluster with the biotech industry and aid in the production of biotech products. It also helps in measuring the extent of intra- and inter-industry clustering, indicating the range of correlation among biotech firms and between biotech firms and other intermediate industries. Moreover, the spatial analysis also involves estimating the regional differences in factors affecting the location of the U.S. biotech industry. Analyzing these spatial concepts and location factors will aid state and local economic development agencies in designing strategies to better retain and attract biotech firms and their clusters, which in turn will boost their economy and provide employment opportunities for local residents.

1.4 Research Question

The primary question this study seeks to address is: how do factors such as localization economies, poverty, unemployment rate, population, median household income, median housing value, property tax, crime index, education, agricultural production, and proximity to venture capital firms, colleges, and hospitals affect the location of the U.S. biotech industry? The specific question this study seeks to address is:

do spatial agglomeration economies (spatial externalities and knowledge spillovers) exist in the location of biotech industry? If they exist: does it spillover predominantly between firms within the biotech sector or between firms in different sectors? And, what is the scope of spatial autocorrelation and correlation between firms within the biotech sector and between firms in different sectors, respectively?

1.5 Objectives

The primary objective of this paper is to identify county level determinants of the spatial distribution of the U.S. biotech industry. Some of the specific objectives are:

1) Analyze the extent to which numerous firm-specific, location-specific, and inter- and intra-industry spatial agglomeration factors affect the location, movement and concentration of the U.S. biotech industry. This objective is achieved by utilizing a Bayesian spatial tobit model that captures the spatial organization of the biotech industry utilizing county-level data for the U.S.

2) Examine the regional differences in factors affecting the location of the U.S. biotech industry. This objective is achieved by utilizing a separate Bayesian spatial tobit model for each of the U.S. census regions.

3) Analyze the interdependence between different subsectors of the U.S. biotech industry. This objective is achieved by employing a Seemingly Unrelated Regression (SUR) and Spatial 2 Stage Least Square (Spatial 2SLS) model.

4) Evaluate the presence and extent of intra-industry and inter-industry spatial correlation. This objective is achieved by utilizing Global Moran's I Statistic and Local Indicators of Spatial Association (LISA) statistics.

1.6 Outline of the Dissertation

This study accomplishes the four objectives through a 'journal-article-style' dissertation, given in chapters 3, 4, and 5. Chapter 2 includes a summary of industry location theory and a review of previous work. In chapter 3, the spatial clustering of the U.S. biotech Industry is presented. Spatial clustering of innovative activities, a case of U.S. biotech related research and testing activities is presented in chapter 4. In chapter 5, the geographical distribution of biotech related manufacturing and Research & Development facilities in the U.S. is analyzed. Finally, an overall summary is included in chapter 6.

1.7 References

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

Barkley, David L and Mark S. Henry. Advantages and Disadvantages of Targeting Industry Clusters. REDRL Research Report, 2001.

Battelle Technology Partnership Practice and State Science and Technology Institute. 2004. " Laboratories of Innovation: State Bioscience Initiatives 2004." Biotechnology Industry Organization, Washington D.C.

Biotechnology Industry Organization. 2005-2006. "Guide to Biotechnology." Biotechnology Industry Facts. BIO, Washington D.C.

Cortright, J., & Mayer, H. (2002). "Signs of Life: The Growth of Biotechnology Centers in the U.S." Portland: The Brookings Institution Center on Urban and Metropolitan Policy.

Ernst & Young. 2004. "Resurgence: The Americas Perspective–Global Biotechnology Report 2004 Ernst & Young LLP.

Ernst & Young. 2005. "Beyond Borders: The Global Biotechnology Report." Ernst & Young LLP.

Feldman, M., 2003. "The locational dynamics of the US biotech industry: Knowledge externalities and the anchor hypothesis." Industry & Innovation, 10(3), 311-328.

Goetz, S. J. and A. Rupasingha. 2002. "High-Tech Industry Clustering: Implications for Rural Areas." American Journal of Agricultural Economics 84(5).

Goetz, S. and S. Morgan. 1995. "State-Level Locational Determinants of Biotechnology Firms". Economic Development Quarterly, (9):174-184.

Grudkova, V. (2001). The Technology Economy: Why do Tech Companies Go Where They Go? EDA National Forum, Washington, DC (May 30, 2001).

Isik, M.2004. "Environmental Regulation and the Spatial Structure of the U.S. Dairy Sector." American Journal of Agricultural Economics 86: 949-962.

Kaliba R.M.A. 2002. Participatory Evaluation of Community-Based Water and Sanitation Programs: The Case of Central Tanzania. Dissertation, Department of Agricultural Economics, Kansas State University.

Lall, Somik, Jun Koo and Sanjoy Chakravorty, (2003) Diversity Matters: The Economic Geography of Industry Location in India, World Bank Policy Research Working Paper Series #3072.

LeSage, J.P. 1999. "The Theory and Practice of Spatial Econometrics." Unpublished, Dept. of Econ., University of Toledo.

McCandless M.E. 2005. "Biotechnology Locations: Hitting the Mark." Business Facilities. http://www.businessfacilities.com/bf_05_05_cover.asp (Date last visited 14rth May 2006).

Munroe T., G. Craft, and D. Hutton. 2002. "A Critical Analysis of the Local biotechnology Industry Cluster — Counties of Alameda, Contra Costa, & Solano in California." Volume III, Appendix. Munroe Consulting Inc., Craft Consulting and Hutton Associates. Research Monograph prepared for a Consortium of Bay Area Organizations, May 2002.

van Oort G. F. 2004. "Urban Growth and Innovation: Spatially Bounded Externalities in the Netherlands." Published by Ashgate Publishing Company, USA.

Porter, M. (1990). The Competitive Advantage of Nations. London: Macmillan.

Roe, B., E. Irwin, and J. S. Sharp. 2002. "Pigs in Space: Modeling the Spatial Structure of Hog Production in Traditional and Nontraditional Production Regions." American Journal of Agricultural Economics 84: 259-278.

Sambidi P.R. and W. Harrison. 2005. "Spatial Dependency of the Geographically Concentrated U.S. Broiler Industry." Paper Presented at the American Agricultural Economics Association Annual Meeting, Providence, Rhode Island.

Scorsone E. 2002. "Industrial Clusters: Enhancing Rural Economies Through Business Linkages." Southern Rural Development Center. http://srdc.msstate.edu/publications/scorsone.pdf

Toole, A. 2003: Understanding entrepreneurship in U.S. biotechnology: what are its characteristics, facilitating factors, and policy challenges?, Chapter 9, in D. Hart (ed.), The Emergence of Entrepreneurship Policy: Government, Start-ups, and Growth in the Knowledge Economy. Cambridge, UK: Cambridge University Press.

Xia, Y., and Buccola, S. 2005. University life science programs and agricultural biotechnology. American Journal of Agricultural Economics, 87(1), 229-243.

Zucker L. G. & M.R. Darby & J.S. Armstrong, 2003. "Commercializing knowledge: university science, knowledge capture and firm performance in biotechnology," Proceedings, Federal Reserve Bank of Dallas, issue Sep, pages 149-170.

CHAPTER 2

REVIEW OF INDUSTRY LOCATION THEORY

2.1 Introduction

The purpose of location theory is to determine the reasons for and explain why a particular factor is important to one industry and not to another (Greenhut, 1982). It also involves the principle of substitution, where an industry selects a site from alternative locations, which in terms of the economic theory, is similar to the problem of substituting labor for capital or land and vice versa (Greenhut, 1982). Numerous factors are considered when determining a suitable site location for a firm. These factors vary depending on the particularities of an industry, but many factors associated with site selection are tied to finding the least-cost location of procuring raw materials and producing and distributing the final product.

Generally speaking location theory is broadly divided into three theoretical schools (Brülhart, 1998), which include: 1) Neo-classical theory, 2) New trade theory, and 3) New Economic Geography (NEG). Since we are interested in industry location within the U.S., we will only discuss the first and the third theoretical schools. Neo-Classical theory deals with the early stages of economic development before the industrial revolution, and is characterized by perfect competition, homogenous products and non-increasing returns to scale (Brülhart, 1998). This theory indicates that the location of an industry is determined exogenously, by the so called *first nature* (term used by Cronon (1991) and Krugman (1993a)), which indicates a given spatial distribution of natural endowments, climate, technology, and factors of production. Conversely, in the NEG theory, location is determined endogenously, by *second nature* (term used by Cronon (1991) and Krugman (1993a)), where spatial distribution of economic activities is

independent from natural advantage (Ottaviano and Thisse, forthcoming). The concept of *second nature* is characterized by mobility of production factors and/or firms, pecuniary and technical externalities, and input-output linkages (Brülhart, 1998). In the geographical economics, *second nature* is analyzed as the component of spatial conditional variation of economic activities that cannot be explained in terms of *first nature*, which is considered to be a control variable associated with natural advantage (Ottaviano and Thisse, forthcoming). It is this concept of *second nature* that leads to several theories on industrial location, which are grouped under the so-called New Economic Geography (NEG). NEG has been pioneered by Fujita (1988), Krugman (1991a, b, 1993b), and Venables (1996). As mentioned earlier in the NEG, location is determined by factor mobility, pecuniary and technical externalities, and input-out linkages, which in turn produce self reinforcing agglomeration processes (Brülhart, 1998). These agglomeration forces will result in clustering of firms/industries, which further result in intra-inter industry specialization, as well as regional specialization. These concepts of the NEG theory have been discussed in the work of early economic geographers and location theorists such as von Thünen (1826), Weber (1909), Marshall (1920), and Hoover (1948). However, the important contribution of the recent economic geographers is that they all use general equilibrium models with monopolistic competition to explain the spatial variation of economic activities, making the concepts more acquiescent to empirical scrutiny and policy analysis (Ottaviano and Thisse, forthcoming).

To better understand location theory, this study presents the location theories of leading writers as explained by Greenhut (1982). Examination of these writings will build a framework to better understand the rationale behind the biotech industry location

decisions. Since the U.S. biotech industry is mainly concentrated in the urban areas, we begin with von Thünen's location theory, which emphasizes the concept of spatial heterogeneity in industry location. von Thünen's theory of location is based on evaluating tradeoffs between product-specific transportation costs and location-specific land rents. He developed a model to predict the type of agricultural product to be grown on geographically dispersed plots of land. Key assumptions of the model include homogeneity of all land attributes except for its distance to a central market, i.e., the urban center. He assumes land is more valuable in the city relative to the country; implying land rents decrease the further away from the city a farm is located. This implies agricultural products grown on plots of land closer to the city are charged higher land rent relative to product grown further away from the city. On the other hand, products grown closer to the city are associated with lower transportation costs relative to product produced at a more distant location. The type of product grown on a particular site is determined by selecting the type of production that yields the lowest cost, given tradeoffs between product-specific transportation costs and land rents.

In contrast to von Thünen's model, which assumes resources are given and the type of industry is chosen, Weber's model assumes the type of industry is given and the optimal site is chosen. Weber assumed that input supply and output demand are known, and there is an unlimited supply of labor at fixed locations at a given wage rate. He considered three general factors of location: transportation cost, labor cost, and agglomeration forces. When transportation cost is the only factor affecting the location of an industry, the site with lowest transportation cost will be selected. This site may be close to the output market, to the input market, or between input and output markets, depending upon the product. Weber argued the orientation of industries is determined by

substitution between transportation and non-transportation cost factors, which include labor costs and agglomeration forces. This substitution involves non-transportation costs exerting a "locational pull", where in some cases they attract an industry from the point of minimum transportation cost to a point of higher transportation cost. This change occurs as long as the savings in non-transportation cost factors exceed the additional transportation costs incurred. Though Weber's location model is a general theory of location for all industries, his assumption of constant demand and omission of institutional factors such as interest rate, insurance, taxes and others leaves gaps in the theory (Greenhut, 1982).

Hoover's theory of industry location bridges this gap by focusing on demand determinants as well as transportation and production factors. Hoover's inclusion of institutional factors provides a more comprehensive theory of firm location than either von Thünen or Weber. Hoover argued that local property taxes were an important element of land cost, thereby influencing the location decision. A distinguishing feature of Hoover's theory is the introduction of fuel and raw material costs, agglomeration forces, and the costs generated by factors such as taxes and climate on the location decision.

2.2 Industrial Clustering and the Economics of Agglomeration

Over the past decade, industrial cluster development is gaining importance as the most vital strategy for economic development to enhance overall regional growth. Moreover, most of the state and local economic development agencies and research economists believe industrial cluster analysis as a policy solution for all regional problems (vom Hofe and Chen 2006). Akundi (2003) conducted a survey of state cluster initiatives and found that as many as 40 states in the U.S. consider industrial cluster analysis as their

critical strategy to promote economic development. Perroux (1950) proposed growth pole / development pole theories to explain how industrial cluster aided in economic development. He argued that well-established firms in a region serve as catalyst (growth poles) to smaller firms in geographic proximity by spreading positive economic effects. Perroux believes economic space as conceptual and homogenous environment where firms buy from and sell to one another following centrifugal and centripetal forces. Perroux (1988) added time to his growth pole theory and indicated that economic development by industrial clusters involve two stages: a first stage involve clustering of business and firms, and in second stage growth spreads to other regions through the goods, investment and information (vom Hofe and Chen 2006).

Agglomeration economies are considered to be one of the driving forces behind clustering of industries. Over the past few decades economic geographers and location theorists have analyzed different forms of agglomeration economies to better understand the theory behind industry location. The concept of agglomeration economies indicate that the performance of one firm is influenced by other firms located nearby. If a firm benefits by locating near an existing firm, it indicates positive economies of scale. Conversely, if a firm is deterred by locating near an existing firm, it indicates negative economies of scale. Agglomeration economies are further divided into Localization economies and Urbanization economies. Localization economies involve technical externalities and knowledge spillovers (Marshall-Arrow-Romer (MAR) externalities) that are specific to an industry whereby the productivity or growth of a firm in a given industry in a given region is assumed to increase the performance of others firms in that industry (van Oort, 2004). The externalities and spillovers include: the formation of a skilled labor pool and the production of new ideas (based on accumulation of human

capital and face to face communication) and the availability of specialized input services (Marshall, 1920; Ottaviano and Thisse, forthcoming). Urbanization economies reflect economic externalities that are transferred to firms of different industries as a result of savings from large scale operation of a city as a whole (van Oort, 2004). It indicates economies of scale associated with generalized location factors such as good infrastructure, favorable community attitude, tax credits and subsidies, and favorable socioeconomic factors. These factors are not specific to a particular industry, but, favor any kind of industry. This in turn results in sectoral diversity, and that is why we see a wide variety of industries in major metropolitan areas. Chinitz (1961) argued that urbanization economies have a higher prospect of successful economic development than localization economies. According to him, established older cities act as an incubator that creates new firms, business and economic opportunities. The level of success for urban economic regeneration or continued development depends on the level of diversification of industrial cluster (Chinitz 1961; vom Hofe and Chen 2006).

van Oort (2004) analyzed several theories that were proposed to determine the performance of agglomeration economies on industry growth and innovation. The MAR theory which focuses on localization economies argues that knowledge spillovers are more important when there is little local competition, so that rents associated with sector-specific knowledge can be internalized. Moreover, in terms of better growth, the theory favors local monopoly over local competition, as the former restricts the flow of ideas to other firms, and therefore allows innovator-internalization (van Oort, 2004). Conversely, Porter (1990), who also focuses on localization economies, argues that it is the local competition that favors rapid adoption of new technology, as different firms in an industry wants to capture as much market share as possible. Jacobs (1969) agrees with

Porter's thought that local competition favors growth, but argues that regional diversity in economic activity (Urbanization economies) will result in higher growth and innovation as knowledge and technical externalities resulting from one sector can be successfully adopted by other sectors (Van Oort, 2004). Scott (1988) argues that vertical disintegration is positively associated with geographical agglomeration. According to the author, firms tend to cluster in territorial space in order to reduce the time and costs of transacting.

2.3 Analytic Framework

The analytic framework to examine the location of the U.S. biotech industry stems from the New Economic Geography (NEG) literature, which stresses the importance of agglomeration forces in the formation of industry clusters. Fujita and Mori (2005) identified four key elements of NEG, which include: 1. general equilibrium modelling of spatial economy, which distinguishes NEG compared to other traditional location theories and economic geography; 2. increasing returns at the level of individual producers leading to imperfect competition; 3. transport costs, which is a key factor in firm location-decision, and 4. locational movement of producers and consumers leading to agglomeration economy (Fujita and Mori 2005). The core-periphery model (CPM), introduced by Krugman (1991a) provides a central framework of the NEG and illustrates the role of above mentioned key elements in changing the spatial structure of economic geography (Fujita and Mori 2005). This study summarizes the CPM model as demonstrated by Fujita and Mori (2005). The basic structure of the CPM assumes that there are two initially symmetric regions (north and south), two production sectors (agriculture and manufacturing), and fixed endowments of two sector-specific factors (industrial workers and agricultural labors). Agricultural labors are not geographically mobile whereas industrial workers do migrate with response to the regional wage

differences. The manufacturing sector produces horizontally differentiated product with scale economies where each firm producing a separate product utilizing workers as the only input. Conversely, agriculture sector produces a homogenous good under perfect competition and constant returns using only labor as input. Both manufacture and agriculture goods are traded, but trade of manufactures involves positive transport costs.

There are two forces driving the spatial structure of the economy, the centripetal and centrifugal forces. The immobility of agricultural labor is considered as a centrifugal force as they consume both goods. The cause of centripetal force, however, is complicated. First, a region with more number of firms is considered to produce a large variety of goods. Workers in that region will have access to wide variety of goods, thus increasing their real income. As a result of this, workers from other region migrate toward the region with access to large variety of goods. Secondly, as number of workers in a region increase, the market size of that region increases, inducing producers to concentrate in that region to benefit from scale economies and at the same time lower transport costs. This movement of workers to be close to producers of consumer goods (forward linkages) and the preference of producers to concentrate in a region with large market size (backward linkage) results in a centripetal force (Krugman 1991a; Fujita and Mori 2005). As a result of these centrifugal and centripetal forces, the economy exhibits a core-periphery pattern, turning one region into industrial core and the other into non-industrial periphery (Krugman 1991a). This theory explains how locational movement of consumers and producers lead to an agglomerated economy.

A firm/industry tries to locate close to other firms/industries in order to obtain positive economies of scale, thus resulting in intra- and inter-industry specialization (clusters). Spatial clustering is eventually limited by offsetting diseconomies such as high

land rent, wages, traffic congestion, and density-related pollution (van Oort, 2004). However, new firms still try to locate in close spatial proximity of the existing clusters in order to utilize economic externalities (Phelps, Fallon, and Williams, 2001). Therefore, it is important to incorporate agglomeration economies to measure the performance of an industry.

To analyze the location of the U.S. biotech industry, this study develops a location model based on drawing from Lall, Koo, and Chakravorty (2003), Fujita and Thisse (1996), and Fujita (1988). Let b(x, y, z) be the benefit (localization economies) obtained by a firm at x from a firm (same industry) located at y, which is at a distance z, from x. This benefit is assumed to be inversely proportional to the distance z, indicating that as the distance between two firms' increases, economies of scale associated with the externalities decreases. Let $f(y)$ be the density (number) of all firms at each location y ϵ $\boldsymbol{X}$ then,

$$B(x) = \int_X b(x, y, z) f(y) dy \tag{2.1}$$

where $B(x)$ represents the aggregate benefit (in dollars) obtained by a firm at x from the externalities created in region $\boldsymbol{X}$. As an aggregate term, the density of firms at each location, $f(y)$, represent urbanization economies, which are external to industries and depend on overall scale as well as scope of economic activity in a given region, indicating inter-industry specialization (Lall, Koo, and Chakravorty, 2003). Thus, the aggregate benefit $B(x)$ is a function of localization and urbanization economies, representing the overall agglomeration economies.

Assume that production utilizes land (K) and labor (L) with rents of $R(x)$ and $W(x)$ respectively at x, where x ϵ $\boldsymbol{X}$. Let $TR(x)$ represent total revenue for a firm located at x

and let $TC(x)$ represent the transportation costs incurred by that particular firm, which maximizes profits subject to:

$$\Pi(x) = TR(x) - R(x)K - W(x)L - TC(x) \tag{2.2}$$

$$TR(x) = pf(B(x), n, o, l)$$

Where p represents the vector of output prices, $f(.)$ represents a function of quantity of output produced, n is the vector of inputs, o is the vector of outputs, and l is the vector of location factors other than agglomeration economies.

2.4 Review of Spatial Econometric Modeling in Industry Location

Most of the previous empirical studies on industry location have employed non-spatial econometric models, such as Ordinary Least Squares (OLS), Poisson, Negative Binomial, and Tobit (for recent exceptions see Roe, Irwin, and Sharp,2002; Goetz and Rupasingha, 2002; Isik, 2004; and Sambidi and Harrison, 2005). Since the data are collected with reference to points in space, employing OLS (and models mentioned above) as an econometric tool will produce spatially autocorrelated residuals, resulting in biased estimates and all inferences based on the model may be incorrect (Anselin, 1988; LeSage, 1999). This section summarizes some of the spatial concepts associated with data collected from points in space and discusses issues related to its modeling. Spatial issues presented in this section are mainly derived from Anselin (1988) and LeSage (1999).

Data obtained from points in space are generally associated with spatial dependence and heterogeneity (Anselin 1988; LeSage 1999). Spatial dependence indicates that observations at a particular location depend on observations at other locations. This spatial dependence in some instances is also expressed as spatial autocorrelation (Anselin 1988). Conversely, spatial heterogeneity refers to variations in

spatial structure in the form of non-constant error variance or model estimates (Anselin 1988). It indicates that relationship between sample data observations varies as we move across the space, thus violating Gauss-Markov assumption that a single linear relationship with constant variance exists across the sample (LeSage 1999). Alternative estimation procedures are required to deal with this variation and draw inferences. Spatial econometrics is considered as the collection of techniques that deal with spatial issues (spatial autocorrelation and heterogeneity) in regression models that utilize data collected from points in space (Anselin 1988). Spatial econometric models include variables that account for spatial interaction within the data by employing a spatial weights matrix.

2.4.1. Spatial Weights

In the spatial literature, the spatial weight matrix provides the composition of assumed spatial relationships among different points in space. This matrix is utilized as a variable in econometric modeling to capture the spatial effects (spatial dependence and heterogeneity) within the data. It is generally built based on the distance between spatial units or the contiguity between spatial unites. Generally speaking, spatial weights matrix that are based on contiguity include two values 0 and 1. If two spatial units share a common border then they are assumed to be connected and are given a value of 1 and 0, otherwise. The weight matrix for three spatial units is represented as follows:

(2.3)
$$\mathbf{W} = \begin{bmatrix} 0 & w_{12} & w_{13} \\ w_{21} & 0 & w_{23} \\ w_{31} & w_{32} & 0 \end{bmatrix}$$

where w_{ij} may be inverse distance between two spatial units (i and j) or 0, 1 if they share a border and/ or vertex. The spatial weight matrix **W** is symmetric, and by principle it

always has zeros on the main diagonal (LeSage 1999). There are different types of contiguity based weights matrices as discussed by LeSage (1999):

1. Linear contiguity: w_{ij}=1 for counties that share a common edge to the immediate right or left of the county of interest.

2. Rook contiguity: w_{ij}=1 for counties that share a common side with the county of interest.

3. Bishop contiguity: w_{ij}=1 for counties that share a common vertex with the county of interest.

4. Double linear contiguity: w_{ij}=1 for two counties to the immediate right or left of the county of interest.

5. Double rook contiguity: w_{ij}=1 for two counties to the right, left, north and south of the county of interest.

6. Queen contiguity: w_{ij}=1 for counties that share a common side or vertex with the county of interest.

The spatial weights matrix is often standardized (row-sums to unity) and multiplied with a vector of observation y (dependent variable) in empirical work, where it is represented as a new variable ($\mathbf{W}y$) known as spatial lag variable. This variable is equal to the mean of observations of the dependent variable from contiguous counties (LeSage 1999). This indicates that what happens at one point in space is function of what happens elsewhere, i.e. $y_i = f(y_j), j \# i$ (LeSage 1999). The linear relationship between the dependent variable and the spatial lag variable is expressed as follows:

$$y = \rho \mathbf{W} y + \varepsilon \tag{2.4}$$

where ρ is a spatial autoregressive parameter, which represents the spatial dependence in the sample data. It measures the average influence of neighboring or contiguous counties

on counties in the vector y (LeSage 1999). A significant parameter ρ will indicate that the data under consideration is spatially dependent and employing OLS models will result in inconsistent and biased estimates.

2.4.2 Spatial Econometric Models

To overcome the problem of spatial dependence in standard linear regressions, two specifications are available: adding an explanatory variable in the form of a spatially lagged dependent variable, or in the error structure (Anselin 1988). Spatial econometric model associated with the former specification is referred to as a spatial autoregressive model (SAR), which is appropriate when spatial dependence operates in the form of a spatially lagged dependent variable. The model with latter specification is referred as spatial error model (SEM), which is relevant when the spatial dependence operates through the disturbance term (Anselin 1988). The SAR model is also referred as spatial lag model and is expressed as:

$$y = \rho \mathbf{W} y + \mathbf{X}\beta + \mu \qquad (2.5)$$
$$\mu \sim N(0, \sigma^2 I_n)$$

where y is a n x 1 vector of the dependent variable, $\mathbf{W}y$ is a spatially lagged dependent variable for spatial weights matrix $\mathbf{W}$, ρ is the scalar for spatial lag coefficient, β is the k x 1 parameter vector, $\mathbf{X}$ is the n x k matrix of exogenous explanatory variables, $\boldsymbol{\mu}$ is an n x 1 vector of normally distributed error terms with zero mean and variance σ^2. The spatial lag $\mathbf{W}y$ can be considered as a spatially weighted average of the dependent variable at neighboring counties. This spatial lag term is assumed to be correlated with error terms, even though the later are independent and identically distributed (Anselin 1999). This relation is expressed as follows:

(2.6) $$y = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\beta + (I - \rho\mathbf{W})^{-1}\mu$$

Accordingly, the spatial lag should be considered and estimated as an endogenous variable; failing to do so will generate inconsistent and biased estimates for the parameter coefficients in the model (Anselin et al. 2000). The spatial lag coefficient ρ implies positive spatial correlation if $\rho > 0$, negative correlation if $\rho < 0$, and no correlation if $\rho = 0$.

The second form of spatial dependence, which is often expressed as a spatial autoregressive process caused by the misspecification of error term (eg. omitted variables) is expressed as:

$$y = \mathbf{X}\beta + \varepsilon$$

(2.7) $$\varepsilon = \lambda\mathbf{W}\varepsilon + \mu$$

$$\mu \sim N(0, \sigma^2 I_n)$$

where ε represents non-spherical error terms expressing spatial dependence in the off-diagonal elements of its covariance matrix, λ is a spatial autoregressive coefficient and μ is a standard spherical error term. Ignoring spatial dependence in the error term does not lead to biased estimates, but the estimate of their variance will be biased, leading to erroneous interpretations and wrong conclusions (Anselin et al. 2000). The spatial autoregressive coefficient indicates that error terms are positively correlated if $\lambda > 0$, negatively correlated if $\lambda < 0$, and not correlated if $\lambda = 0$.

If there was evidence indicating spatial dependence in both forms (spatial lag and error terms), a more general specification can be employed. This general spatial specification is called general spatial model (SAC), which accounts for both the spatially lagged term as well as a spatial error structure. The SAC model is represented as follows:

$$y = \rho \mathbf{W} y + \mathbf{X}\beta + \varepsilon$$

(2.8) $$\varepsilon = \lambda \mathbf{W} \varepsilon + \mu$$

$$\mu \sim N(0, \sigma^2 I_n)$$

This specification is employed if there is evidence that spatial dependence existed in the error structure from a SAR estimation, which is tested by employing a Lagrange Multiplier test. Before estimating a spatial econometric model, the presence of spatial dependence in data can be analyzed utilizing spatial exploratory models, which are discussed in the following section.

2.4.3 Spatial Exploratory Analysis

2.4.3.1 Global Moran's I

The spatial association of data collected from points in space is tested using a Global Moran's I, which measures similarities and dissimilarities in observations across space (Anselin, 1995). For the number of biotech establishments *y*, Moran's I is:

(2.9) $$I = \left(\frac{n}{\sum_i \sum_j w_{ij}} \right) \times \left(\frac{\sum_i \sum_j (y_i - \mu)(y_j - \mu)}{\sum_i (y_i - \mu)^2} \right)$$

where w_{ij} indicates elements of the spatial weight matrix $\boldsymbol{W}$ (distance/contiguity weight matrix) between two points (*i* & *j*), μ the mean of all *y* observations, and i, j=1,…,n. A positive and significant value for Moran's I indicate positive spatial correlation, showing that counties have a high or low number of establishments similar to their neighboring counties. Conversely, a negative and significant value for Moran's I indicates negative spatial correlation, showing that counties have high or low number of establishments unlike neighboring counties (Pacheco and Tyrrell, 2002). The study calculate's Moran's I

for the number of biotech establishments across all contiguous U.S. counties, employing GeoDa, a spatial data analysis software.

2.4.3.2 Local Indicators of Spatial Association (LISA)

In the case of uneven spatial clustering, global spatial indicators such as Moran's I are found to be less useful. This resulted in a new general class of local spatial indicators such as Local Indicators of Spatial Association (LISA, also known as Local Moran), which measures the contribution of individual counties to the global Moran's I statistic (Anselin, 1995). The LISA statistic is calculated for the i^{th} county as:

(2.10) $$I_i = z_i \sum_j w_{ij} z_j$$

where w_{ij} indicates elements of the spatial weight matrix $\boldsymbol{W}$ (distance/contiguity weight matrix) between two points (i & j), z_i and z_j indicates the standardized number of establishments for county i and j, respectively. The sum of LISAs ($\sum_i I_i$) for all observations is proportional to global Moran's I, implying that LISA statistic can be interpreted as indicators of local spatial clusters and as diagnostics for local instability (spatial outliers) (Anselin, 1995).

The LISA cluster map indicates the locations with a significant Local Moran statistic classified by type of spatial correlation: (a) high-high association (HH), a county with many biotech firms has neighboring counties with many biotech firms; (b) low-low association (LL), a county with few biotech firms has neighboring counties with few biotech firms; (c) low-high association (LH), a county with few biotech firms has neighboring counties with many biotech firms; and (d) high-low association (HL) , a county with many biotech firms has neighboring counties with few biotech firms. The HH and LL locations suggests clustering of similar values (positive spatial correlation),

whereas the HL and LH locations indicate spatial outliers (negative spatial correlation) (Anselin, 1995).

2.4.3.3 Multivariate Global Moran's I

In addition to the spatial autocorrelation, multivariate spatial correlation is also analyzed employing a multivariate Moran's I statistic. The multivariate spatial correlation "centers on the extent to which values of one variable (z_k) observed at a given location show a systematic (more than likely under spatial randomness) association with another variable (z_l) observed at the neighboring locations." (Anselin, Syabri, and Smirnov, 2002). The multivariate Moran's I is as follows:

(2.11) $$I_{kl} = \frac{z_k' W z_l}{n}$$

where n indicates the number of observations, $\boldsymbol{W}$ indicates rook contiguity weight matrix, and z_k and z_l indicate standardized variables with mean zero and standard deviation equal to one (Anselin, Syabri, and Smirnov, 2002).

2.4.3.4 Multivariate LISA

Using a similar rationale as in the development of LISA, a Multivariate Local Moran Statistic (MLMS) was developed by Anselin, Syabri, and Smirnov (2002). This is defined as follows:

(2.12) $$I_{kl}^i = z_k^i \sum_j w_{ij} z_l^j$$

where w_{ij} indicates elements of the spatial weight matrix, $\boldsymbol{W}$ (Rook contiguity weight matrix), and z_k^i and z_l^j indicates the standardized variables for county *i* and *j*, respectively. The MLMS "gives an indication of the degree of linear association (positive or negative) between the values for one variable at a given location *i* and the

average of another variable at neighboring locations." (Anselin, Syabri, and Smirnov, 2002). Similar to LISA, MLMS suggests two classes of positive spatial correlation, or spatial clusters (HH and LL), and two classes of negative spatial correlation, or spatial outliers (HL and LH) (Anselin, Syabri, and Smirnov, 2002).

2.5 Conclusion

This chapter briefly discussed different theories related to industry location. It developed an analytical framework that emphasizes the significance of agglomeration economies on the biotech industrial cluster development. It also argued some of the spatial issues associated with data collected from points in space and discussed suitable econometric models that successfully incorporate those spatial issues in modeling the location biotech industry. The following chapter presents the spatial clustering of the U.S. biotech Industry. Spatial clustering of innovative activities, a case of U.S. biotech related research and testing activities is presented in chapter 4. In chapter 5, the geographical distribution of biotech related manufacturing and Research & Development facilities in the U.S. is analyzed.

2.6 References

Akundi, K. 2003 Cluster-based Economic Development, Part 1: A Survey of State Initiatives. Texas Economic Development, Business and Industry Data Center, www.bidc.state.tx.us.

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

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

Anselin L. 1999. Interactive techniques and exploratory spatial data analysis, in Longley P.A., Goodchild M.F., Maguire D.J. and Rhind D.W. (eds.) Geographical Information Systems: Principles, Techniques, Management and Applications, JohnWiley, New York, NY, 251–264.

Anselin, L., I. Syabri, and O. Smirnov. 2002. “Visualizing Multivariate Spatial Correlation with Dynamically Linked Windows.” In: Proceedings of the SCISS Specialist Meeting "New Tools for Spatial Data Analysis". Santa Barbara, California, USA. May 10-11, 2002.

Brülhart, M. 1998. “Economic Geography, Industry Location, and Trade: The Evidence.” World Economy 21(August):775-802.

Chinitz, B. 1961 Contrast in Agglomeration: New York and Pittsburgh. American Economic Review 51: 279-89.

Cronon, William. Nature’s Metropolis: Chicago and the Great West. New York: Norton, 1991.

Fujita M (1988) A monopolistic competition model of spatial agglomeration: differentiated product approach. Regional Science and Urban Economics 18: 87-124.

Fujita, Masahisa and Jean F. Thisse, “Economics of Agglomeration,” CEPR discussion paper 1344 (1996).

Fujita, M. and Mori, T. (2005a): “Frontiers of the New Economic Geography”, *Papers in Regional Science*, 84(2), 377-405.

Goetz, S. J. and A. Rupasingha. 2002. “High-Tech Industry Clustering: Implications for Rural Areas.” American Journal of Agricultural Economics 84(5).

Greenhut Melvin L. “Plant Location in Theory and in Practise, the Economics of Space”. The Theory of Least-Cost Location, Chapter one. Greenwood Press, Publishers. 1982.

vom Hofe, R. and Chen, K. 2006 Whither or not Industrial Cluster: Conclusions or Confusions. The Industrial Geographer 4: 2-28.

Hoover, Edgar M. Location of Economic Activity, 1st ed. New York; McGraw-Hill, 1948.

Isik, M.2004. “Environmental Regulation and the Spatial Structure of the U.S. Dairy Sector.” American Journal of Agricultural Economics 86: 949-962.

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

Krugman P (1991a) Increasing returns and economic geography. Journal of Political Economy 99: 483-499.

Krugman P (1991b) Geography and Trade. MIT Press, Cambridge, MA.

Krugman P (1993a) First nature, second nature, and metropolitan location. Journal of Regional Science 33: 129-14

Krugman P (1993b) On the number and location of cities. European Economic Review 37: 293-298.

Lall, Somik, Jun Koo and Sanjoy Chakravorty, (2003) Diversity Matters: The Economic Geography of Industry Location in India, World Bank Policy Research Working Paper Series #3072.

LeSage, J.P. 1999. "The Theory and Practice of Spatial Econometrics." Unpublished, Dept. of Econ., University of Toledo.

Marshall, A. 1920. Principles of Economics. Eighth edition. London: Macmillan.

National Science Foundation. 2006. Industrial Funding of Academic R&D Continues to Decline in FY 2004. NSF 06-315.

van Oort G. F. 2004. "Urban Growth and Innovation: Spatially Bounded Externalities in the Netherlands." Published by Ashgate Publishing Company, USA.

Ottaviano, G. and J.-F. Thisse. New Economic Geography: what about the N? Environment and Planning A, forthcoming.

Pacheco, Andrada I., and Timothy J. Tyrrell. 2002. "Testing Spatial Patterns and Growth Spillover Effects in Clusters of Cities." Journal of Geographical Systems 4:275-285.

Perroux, F. 1950 Economic Space: Theory and Applications. Quarterly Journal of Economics 64: 89-104.

Perroux, F. 1988 Growth Pole Theory and Strategy Reconsidered: Domination, Linkages and Distribution. Economic Development Essays in Honour of Francois Perroux.

Phelps, N.A., Fallon, R.J. and Williams, C.L. (2001): "Small firms, borrowed size and the urban–rural shift", Regional Studies 35: 613–624.

Porter, M. (1990). The Competitive Advantage of Nations. London: Macmillan.

Roe, B., E. Irwin, and J. S. Sharp. 2002. "Pigs in Space: Modeling the Spatial Structure of Hog Production in Traditional and Nontraditional Production Regions." American Journal of Agricultural Economics 84: 259-278.

Sambidi P.R. and W. Harrison. 2005. "Spatial Dependency of the Geographically Concentrated U.S. Broiler Industry." Paper Presented at the American Agricultural Economics Association Annual Meeting, Providence, Rhode Island.

Scott, A. (1988), New Industrial Spaces: Flexible Production, Organization and Regional Development in North America and Western Europe, Pion, London.

von Thünen JH (1826) Der Isolierte Staat in Beziehung auf Landtschaft und Nationalökonomie. Hamburg. English translation: Wartenberg CM (1966) von Thünen's Isolated State. Pergamon Press, Oxford.

Venables AJ (1996) Equilibrium locations of vertically linked industries. International Economic Review 37: 341-359

Weber,A (1909): Ueber den Standort der Industrien, Tuebingen. "Theory of the Location of Industries". Simplifying Assumptions (Chapter 2), Transport Orientation (Chapter 3), and Labor Orientation (Chapter 4). The University of Chicago Press. Seventh impression 1969. Translated by C. J. Friedrich.

Zellner, A. (1962), "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," *Journal of the American Statistical Association*, 57, 348-368.

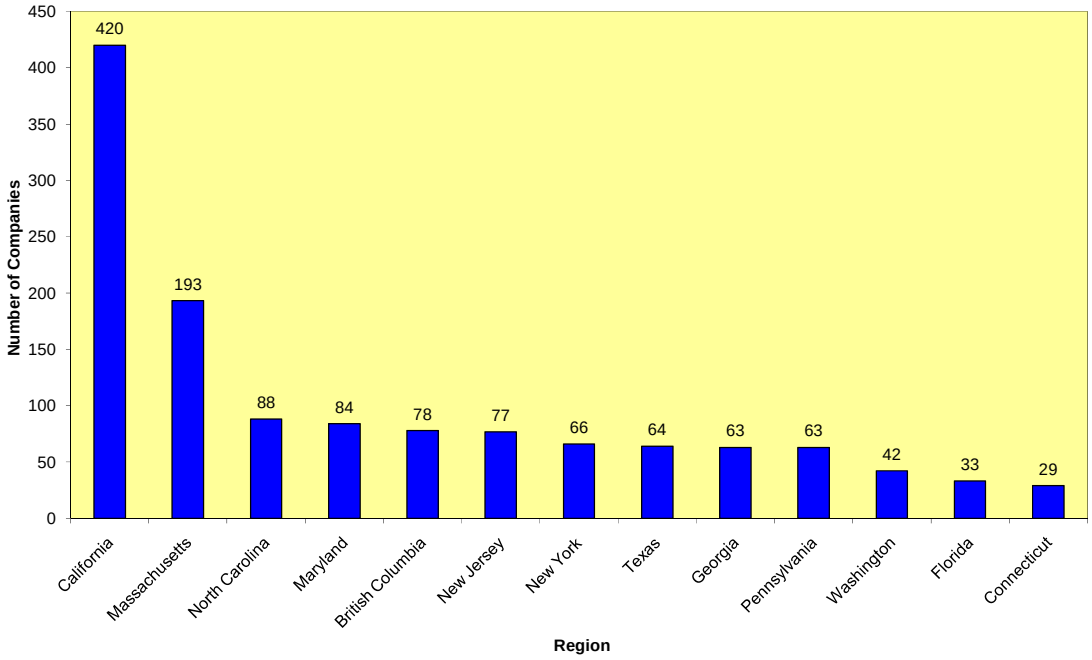
CHAPTER 3

SPATIAL CLUSTERING OF THE U.S. BIOTECH INDUSTRY

3.1 Background

The biotech industry is defined as "the application of biological knowledge and techniques to develop products and services" (Cortright and Mayer, 2002; BIO, 2005). Goetz and Morgan (1995) defined it as "any technique that uses living organisms or parts of organisms to make/modify products, improve plants or animals, or develop microorganisms for a specific use." Biotech firms are mainly research and development (R&D) oriented and operate in collaboration with research-oriented universities, biomedical research centers, and other diversified companies that aid in the production and distribution of biotech products. Biotech products may be characterized as drugs and pharmaceuticals, agricultural, and environmental, which aid in improving the quality of health, increasing the production of agricultural goods, improving food quality, minimizing environmental hazards and providing a cleaner environment (Ernst and Young, 2005).

The biotech industry is one of the fastest growing industries in the U.S., increasing sales from $7.7 billion in 1994 to $33.3 billion in 2004 (Ernst and Young, 2005). According to the Biotechnology Industry Organization (BIO, 2005), in 2003 there were 1,473 biotechnology companies in the U.S., employing 198,300 people and spending $17.9 billion on research and development. The top five states in terms of number of biotech companies are: California (420), Massachusetts (193), North Carolina (88), Maryland (84), and New Jersey (77) (figure 3.1; Ernst and Young, 2004). The biotech industry is mainly concentrated in nine cities/regions (San Francisco Bay Area, Boston/Cambridge, San Diego, Los Angeles, New York, Philadelphia, Raleigh-Durham,



Source: Ernst & Young LLP, America's Biotechnology Report: Resurgence, 2004.

Figure 3.1. U.S. Biotech Companies by State and Province.

Seattle and Washington, DC), accounting for three-fourths of the nation's largest biotechnology firms and for three-fourths of the biotech firms formed in the past decade (Cortright and Mayer, 2002).

As a result of the increasing success of the biotech industry, several state and local economic development agencies are pursuing strategies to attract new biotech firms, resulting in stiff competition among and within states. For example, in 2004, 40 states have adopted strategies to stimulate the growth of biotechnology and 50 states have technology based economic development initiatives for biotech firms, compared to merely 14 states in 2001 (Battelle and State Science and Technology Institute (SSTI), 2004). Moreover, a survey of 77 local and 36 state economic development departments

indicated that 83% have listed biotechnology as one of the top two targets for industrial development (Grudkova, 2001; Cortright and Mayer, 2002).

As a result of increased competition in attracting biotech firms, the U.S. biotech industry is undergoing changes in its geographical distribution. For example, between1991 and 2001, the concentration of biotech firms in the top states increased slightly, but the relative ordering of states changed with New Jersey dropping below Maryland and North Carolina in terms of the number of establishments (Feldman, 2003). Moreover, a study conducted by McCandless (2005) indicated that Dallas, Houston, San Antonio, Memphis, Richmond, and Miami-Dade are making significant progress in becoming important biotech regions of the future. Therefore, it is critical to analyze the economic factors and resource endowments that may affect the location of biotech firms. Additionally, it is also important to analyze statistical implications of the spatial clustering of biotech firms, which was excluded in earlier studies (except Goetz and Rupasingha, 2002).

The primary objective of this paper is to identify county level determinants of the spatial distribution of the U.S. biotech industry. Specifically, this study analyzes the extent to which numerous firm-specific, location-specific, and inter- and intra-industry spatial agglomeration factors affect the location, movement and concentration of the U.S. biotech industry. The study utilizes a Bayesian spatial tobit model that captures the spatial organization of the biotech industry utilizing county-level data for the United States. Analyzing these factors will aid state and local economic development agencies in designing strategies to better retain and attract biotech firms, which in turn will boost state and local economies and provide employment opportunities for their residents.

3.2 Literature Review

Several studies have empirically examined the location aspects of the U.S. biotech industry (Goetz and Morgan, 1995; Darby and Zucker, 1996; Gray and Parker, 1998; Prevezer, 1998; Lerner and Merges, 1998;; Zucker, Darby, and Brewer, 1998; Brennan, Pray and Courtmanche, 1999; Zucker, Darby, and Armstrong, 2003; Xia and Buccola, 2005). The rationale for concentration of the U.S. biotech industry in California and the Northeast has been attributed to proximity to highly research-oriented universities, research parks and laboratories, and well developed infrastructure. Gray and Parker (1998) examined the theoretical arguments surrounding the location and organization of biotech firms and analyzed the prospects for industrial renewal and regional transformation. The authors segregate the U.S. biotech industry into three different categories/regions based on the functions performed by biotech firms in those regions. The first category includes mature drug producing regions, such as New York, New Jersey, Pennsylvania, Delaware, Illinois, and Indiana. These regions include mature pharmaceutical firms that were established prior to the commercialization of biotechnology (before 1970s), and are now primarily involved in the manufacturing (53 %) and marketing (72%) of new drugs. Another category includes emerging drug-producing regions, such as San Francisco, San Diego, Los Angeles, Seattle, and Boston. Firms in these regions were established mainly during and after the commercialization of biotechnology, and are primarily involved in R&D activities (82%) for new drugs. A third category includes low-cost periphery regions, such as Puerto Rico, the Southern states of the U.S., and other scattered isolated rural areas. Biotech firms in these regions undertake the production of drug products that have achieved commercial scale and other intermediate products (Gray and Parker, 1998).

Munroe, Craft, and Hutton (2002) conducted a survey of biotech companies in three California counties (Alameda, Contra Costa, and Solano). The results indicated that proximity to leading research centers (i.e. ready supply of skilled labor, access to ongoing research activities, new technology, etc.) as primary determinants for their current locations. Access to venture capital, a well-trained workforce, space for expansion and access to new technology were considered to be the most critical requirements for the growth and prosperity of respondents' business. The results also indicated that state and local economic development agencies that provide financial assistance packages (i.e. subsidies, tax advantages, loan guarantees, etc.), biotechnology incubators/research parks (with appropriate zoning, infrastructure and public transportation), promote public awareness and training programs for the workforce are more likely to attract biotech firms. Other results indicated that some of the respondents were willing to locate in regions associated with lower costs (housing, space, wages, etc.), less congestion/commuting, and good incentives such as subsidies and tax credits (Munroe, Craft, and Hutton, 2002).

Goetz and Rupasingha (2002) analyzed the site-specific determinants of the U.S. high-tech industry, which includes firms that are involved in biotech activities, such as drug and pharmaceutical manufacturing firms and R&D services. Their results indicated that the availability of an existing high-tech firm, number of college graduates, local property taxes, population (urbanization economies), total county income, highway access, and county amenity scale have a positive and significant impact on the location of high-tech firms. Conversely, a county's unemployment and unionization rate, per capita pollution, and the percentage of blacks were found to have a negative and significant impact on the high-tech firm's location (Goetz and Rupasingha 2002). The present article differs from previous literature in that we examine the economies of scale associated with

county-level spatial agglomeration factors, exclusively for firms involved in the biotech activities. Moreover, our study includes several biotech related sub-industries (Agricultural Feedstock and Chemicals) that are of potential interest to agricultural research institutions.

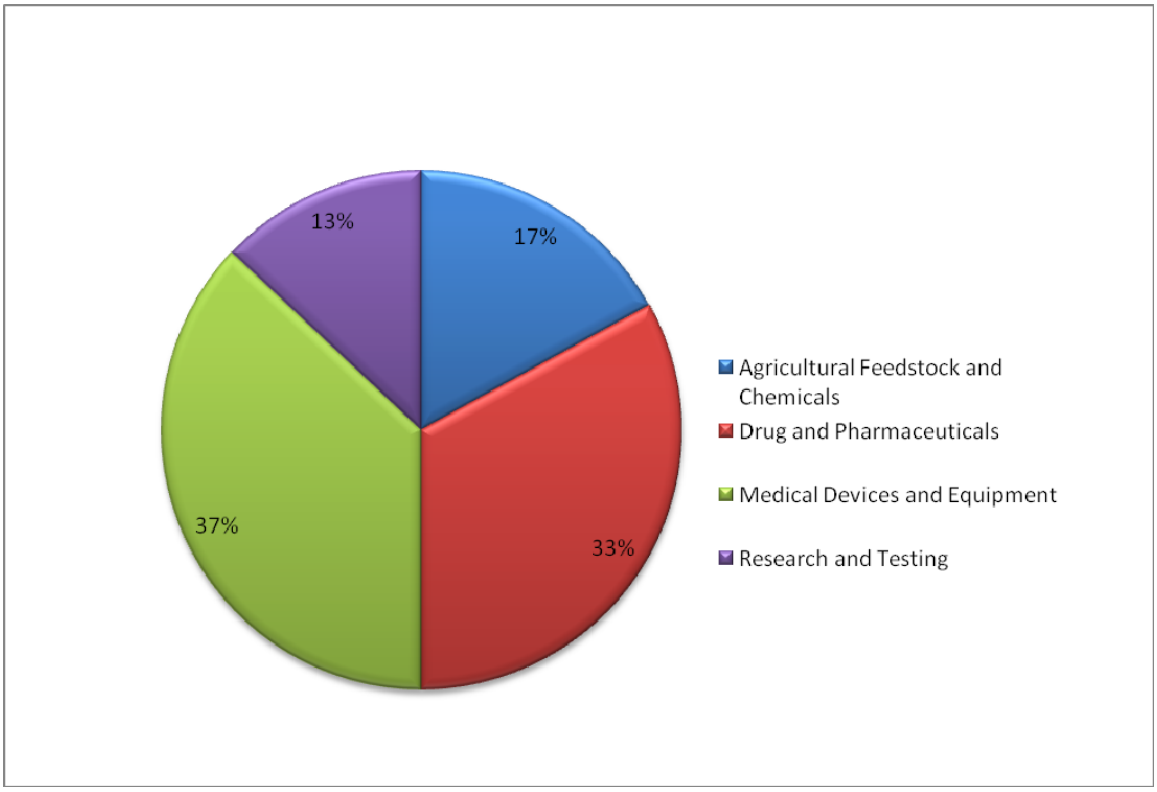
3.3 Data

The biotech industry is a composition of numerous manufacturing, R&D, and services industries. Consequently, it does not have a separate NAICS code since different subsectors are involved in the production of biotech products. However, Battelle Technology Partnership Practice and State Science and Technology Institute (SSTI) (2004) classified the bioscience[1] into five major subsectors as follows: 1. Agricultural Feedstock and Chemicals (NAICS: 311221, 311222, 311223, 325193, 325199, 325221, 325222, 325311, 325312, 325314, 325320, 424910), 2. Drugs and Pharmaceuticals (NAICS: 325411, 325412, 325413, 325414), 3. Medicinal Devices and Equipment (NAICS: 339111, 339112, 339113, 339114, 334510, 334516, 334517), 4. Research and Testing (NAICS: 541380, 541710), and 5. Academic Health Centers, Research Hospitals, and Research Institutes. The first four subsectors include twenty five industries that are involved in biotechnology activities, with total employment of 885,368 jobs across 17,207 establishments (Battelle and SSTI, 2004). Figure 3.2 illustrates the U.S. employment distribution across the bioscience subsectors.

The present study analyzed several categories of variables that are considered to affect the location of biotech firms, such as agglomeration factors, infrastructure factors, and local economic and socioeconomic factors. County-level data was obtained from the 2003 county business patterns (U.S. Census Bureau), Economic Research Service,

[1] "The biosciences are not just biotechnology but rather a range of industry sectors relying on insights into the way living organisms function."(Battelle and SSTI, 2004).

National Agricultural Statistics Service, Battelle Technology Partnership Practice and SSTI (state-level data), and U.S. Dept. of Labor. The dependent variables considered in the study are: county level number of biotech establishments in 2003 (firms belonging to the aforementioned NAICS codes) and the change in county level biotech establishments' from1998 to 2003. The study also analyzed the regional differences in factors affecting the location of new biotech establishments by employing a separate model for each of the four regions (South, Northeast, Midwest and West).



Source: Battelle Technology Partnership Practice and State Science and Technology Institute. 2004.

Figure 3.2. U.S. Employment Distribution Across the Bioscience Subsectors.

Economies of scale associated with agglomeration factors are believed to be one of the driving forces in the geographical distribution of the biotech industry (Pisano, Shan, and Teece, 1988; Gray and Parker, 1998; Goetz and Rupasingha, 2002; Munroe,

Craft, and Hutton, 2002; McCandless, 2005). Agglomeration economies indicate that performance of one biotech firm is influenced by the other biotech firms located nearby. The resulting spillovers may be due to an already existing industry-specific infrastructure, which is associated with lower transaction costs, proximity to research institutions and specialized intermediate industries, good transportation facilities, and availability of skilled labor pool and financial resources. The research oriented biotech firms generate externalities and spillovers, which tend to be spatially proximate to where they were created, resulting in positive economies of scale for firms located in that region (Jaffe, Trajtenberg, and Henderson, 1993; Dahlander and McKelvey, 2003). Zeller (2001) analyzed the spatial clustering of biotech firms in Germany and argued that, even though knowledge and technology transfer often happens on a global scale, the exchange of tacit knowledge, however, is facilitated by spatial proximity. This approach of industrial cluster analysis is used by many economic development agencies to enhance regional growth. Akundi (2003) conducted a survey of state cluster initiatives and found that as many as 40 states in the U.S. consider industrial cluster analysis as their critical strategy to promote economic development. In this study, we include a spatial lag variable as a proxy for agglomeration economies that accounts for the biotech establishment counts in neighboring counties. The variable is hypothesized to have a positive effect on the location of biotech firms.

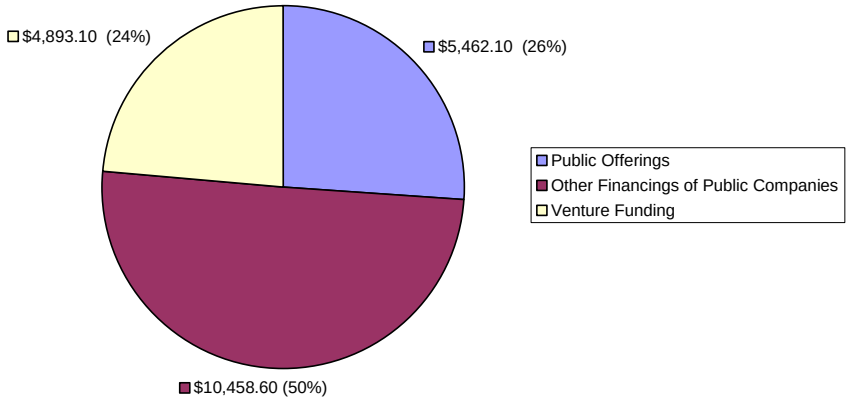
A factor that is considered to be a prerequisite for attracting a biotech firm is proximity to research institutions. Several studies have analyzed the role of research institutes in the development and commercialization of biotechnology (Powell and Brantley, 1992; Darby and Zucker, 1996; Zucker, Darby, and Brewer, 1998; Prevezer, 1998; Zucker, Darby, and Armstrong, 2003; Dahlander and McKelvey, 2003; Xia and

Buccola, 2005). Industry funded university research increased from $630 million in 1985 to $2.1 billion in 2004 (National Science Foundation, 2006), indicating an increasing affiliation between university and industry in technology advancement. Some of the primary reasons for this collaboration include access to complementary research activity and human capital, increasing commercial opportunities, stringency of patent law and federal policies, and the relative decline of public research funding (Santoro and Alok, 1999; Yang and Buccola, 2003). This study includes county-level number of colleges, universities, and professional schools (*colleges*) as a proxy for the proximity to research institutions and assumes it will have a positive effect on the location of biotech industry. County-level data for this variable was collected from county business patterns (U.S. Census Bureau, 1998 and 2003) with 611310 NAICS classification, which includes number of colleges, universities, and professional schools.

Ongoing research intensity in life sciences at research institutions of a particular state is also considered to be a critical factor in the location decision of a biotech firm (Zucker, Darby, and Brewer, 1998; Munroe, Craft, and Hutton, 2002; Xia and Buccola, 2005). State level university life science Research &Development expenditures (*R&D*), National Institute of Health support for institutions (*NIH*), higher education degrees in biological science (*Biological Degrees*), and average biological scientists in the workforce 2000-2002 (*Biological Scientists*), are included as proxies for ongoing research intensity. All these variables are assumed to have a positive effect impact on the site-selection of a biotech firm. State level data for R&D expenditures was collected from National Science Foundation, whereas state level data for other three variables was collected from Battelle and SSTI (2004).

Businesses that provide venture capital are considered to be an important source of capital, especially, for new and small firms (Powell, Koput, Bowie, and Smith-Doerr, 2002). For a small biotech firm, availability of venture capital in a particular region is as important as the strong research capacity of that region. During 2004, venture capital accounted for approximately 23.5% of the total financing for the biotech industry as indicated in figure 3.3 (BIO, 2005; BioWorld, 2005). Most of the biotech firms are small and operate at a loss, spending large amount of money on research and development for several years, before earning a profit (Cortright and Mayer, 2002). For example, only 1 in 5,000 potential new medicines reach the pharmacy shelf, and that is after 12 to 15 years of R&D with an average expenditure of $500 million (California Trade and Commerce Agency, 2001). As a result, most of the small biotech firms depend on venture capital funds, on research contracts and equity investment from large biotech firms, and on sales of their company stock in public markets (Cortright and Mayer, 2002). Therefore, the availability of local venture capital firms (*Venture Capital*) is hypothesized to have a positive effect on the location of a biotech firm in that region (data for county-level number of venture capital establishments was collected from county business patterns with 523910 NAICS classification).

Agriculture is an important component of the biotech industry. Some studies have analyzed issues related exclusively to the spatial distribution of agricultural biotech firms and its relationship with research institutions (Kalaitzandonakes and Bjornson, 1997; Graff, 1997; Begemann 1997; Brennan, Pray, and Courtmanche, 1999; Sporleder, Moss, and Nickles, 2002; Yang and Buccola, 2003; Sporleder and Moss, 2004; Xia and Buccola, 2005). Around 13 percent of firms in biotechnology are primarily involved in



Source: Biotechnology Industry Organization, 2005; BioWorld, 2005.

Figure 3.3. Biotech Industry Financing, 2004.

agriculture (Dibner 1995; Graff 1997). As illustrated in figure 2, agriculture related biotech activities account for 17% of total U.S. employment across bioscience subsectors. According to Ernst and Young (2000), in 1999, agricultural biotech firms employed 21,900 workers, generated $2.3 billion in revenues and $1.4 billion in personal income for employees and owners. The primary goal of agricultural biotechnology is to develop high yielding varieties with improved resistance to natural enemies (e.g. pest, diseases, weeds, and adverse growing conditions), and better quality and longer shelf life for fruits and vegetables (Ernst and Young, 2005). Since some of the biotech firms seek applications directed toward agricultural production, it is hypothesized that, in order to gain positive external economies of scale (low transaction costs), biotech firms prefer to

locate in regions with significant agricultural production (*Farmland*) (county level data for farmland was collected from U.S. Census of Agriculture). Similarly, since the biotech industry involves drugs and pharmaceutical firms, and medicinal devices and equipment firms, we hypothesize that a county with more hospitals (*Hospitals*) will have a positive effect on the location of the biotech industry. County-level data for this variable was collected from county business patterns (U.S. Census Bureau, 1998 and 2003) with 622110 NAICS classification, which includes number of general medical and surgical hospitals.

In terms of conventional location theory, local property taxes (*Property Tax*) may discourage new investment by increasing the costs of production. However, in case of high-tech firms (such as the biotech industry), high property taxes are considered to be proxies for greater availability or higher quality of local public goods (Goetz and Rupasingha), which in turn reflects high standards of living of the local community. Therefore, we assume that property taxes are positively correlated with the location of biotech firms (county-level data for property tax collections was obtained from the U.S. Census Bureau). Similarly, counties with high unemployment (*unemployment*) and poverty rates (*poverty*), which reflect low standards of living of the local community, are considered to have a negative effect on the location-decision of biotech firms (county-level data for unemployment and poverty rate are collected from Economic Research Service USDA and U.S. Bureau of Census, respectively). Similarly, counties with higher crime rates (*Crime Index*) are also considered to have a negative impact on the location of biotech firms. County-level data for crime index was obtained from Geospatial and Statistical Data Center, University of Virginia Library. Urban economies are measured using the rural-urban continuum codes for U.S. counties, which range from 1 to 9, where

1 represents extremely urban and 9 represents extremely rural. They were further aggregated into two groups: Metro (counties belonging to rural-urban continuum codes of 1-3) and Non-metro (counties belonging to rural-urban continuum codes of 4-9). The study assigns a value of 1 for non-metro counties and 0 for metro counties. Since, most of the existing biotech firms are located in the metropolitan cities; we hypothesize that the variable (*Metro-Nonmetro*) will have a negative impact on the location-decision of biotech establishments.

The impact of labor quality on the location decision of biotech industries is measured by county-level average wage per job (*Wage*) and percentage of persons with a college degree (*Education*) (Zucker, Darby, and Brewer, 1998). Both variables are considered to have a positive relationship with the site-selection of the biotech industry. County-level data for *wage* and *education* are obtained from the Bureau of Economic Analysis and Economic Research Service USDA, respectively. Biotech firms prefer to locate in highly populated centers (*Population*) as it provides appropriate services such as contracting for site building, major equipment, and availability of housing (McCandless, 2005). Counties with high median household income (*Income*), which represents a high standard of living, are considered to have a favorable impact on the biotech firms' site-selection. County-level data for population and median household income was obtained from U.S. Census Bureau and Economic Research Service USDA, respectively. Similarly, median housing values (*Housing Value*) are used as a proxy for the quality of housing in a given county. It is expected to have a positive impact on the location-decision. County-level data for median housing value was obtained from U.S. Census Bureau. Table 3.1 and 3.2 presents the descriptive statistics of all the variables included

Table 3.1. Summary Statistics of Variables Employed in the Spatial Tobit Model for U.S. Biotech Industry Location.

Variable	Mean	Std. Dev.
Total Number of Biotech Establishments (number)[1]	11.50	41.64
Poverty Rate (percent)	13.36	4.89
Population (number in 1000s)	92.92	303.47
Owner-Occupied Housing Units: Median Value (in $1000)	80.93	41.94
Unemployment Rate (percent)	5.97	1.96
Median Household Income (in $1000)	36.73	9.28
Total Number of Venture Capital Firms (number)	1.89	11.45
Number of Colleges, Universities and Professional Schools (number)	1.08	4.86
Number of General Medical and Surgical Hospitals (number)	1.75	3.70
Average Wage per Job (in $1000)	27.05	5.79
Percent of Persons with a College Degree (percent)	16.51	7.80
Property Tax (in $1000)	22.74	23.41
Crime Index (index in 100s)	35.95	139.73
Land in Farm Acres (in 1000 acres)	301.14	385.21
Metro-Nonmetro (dummy)	0.65	0.48
University Life Science R&D Expenditures (in $1000)	563.61	573.18
NIH Support for Institutions (in $1000)	427.71	516.33
Higher Education Degrees in Biological Science (number in 100s)	27.26	21.44
Biological Scientists in Workforce 2000-2002 Avg (number in 100s)	118.38	106.01

Note: [1] text in parenthesis indicates units of measurement.

Table 3.2. Summary Statistics of Variables Employed in the Spatial Tobit Model for Change in U.S. Biotech Industry Location.

Variable	Mean	Std. Dev.
Change in Biotech Establishments (number)[1]	1.75	6.30
Change in Poverty Rate (percent)	-1.32	1.96
Change in Population (number)	6581.22	277954.70
Change (1990-2000)[2] in Owner-Occupied Housing Units: Median Value (dollars)	27172.60	19233.77
Change in Unemployment Rate (percent)	0.81	1.98
Change in Median Household Income (dollars)	2786.50	2449.81
Change in Number of Venture Capital Firms (number)	-0.31	3.22
Change in Number of Colleges, Univ. and Professional Schools (number)	0.13	1.10
Change in Number of General Medical and Surgical Hospitals (number)	-0.05	0.91
Change in Average Wage per Job (dollars)	4101.71	1779.04
Change (1990-2000) in Percent of Persons with a College Degree (percent)	3.03	2.21
Change (1997-2002) in Property Tax (in $1000)	-43909.30	272672.30
Change (1998-2002) in Crime Index (index)	-71.99	2474.28
Change(1997-2002) in Land in Farm (acres)	163.96	113160.50
Metro-Nonmetro (dummy)	0.65	0.48
Change(1997-2002) in University Life Science R&D Expenditures (in $1000)	304.93	315.40

Note: [1] text in parenthesis indicates units of measurement.
[2] change in variables are measured from 1998 to 2003, unless mentioned otherwise.

in the spatial tobit model for the location of U.S. biotech industry and for change in the location of U.S. biotech industry, respectively.

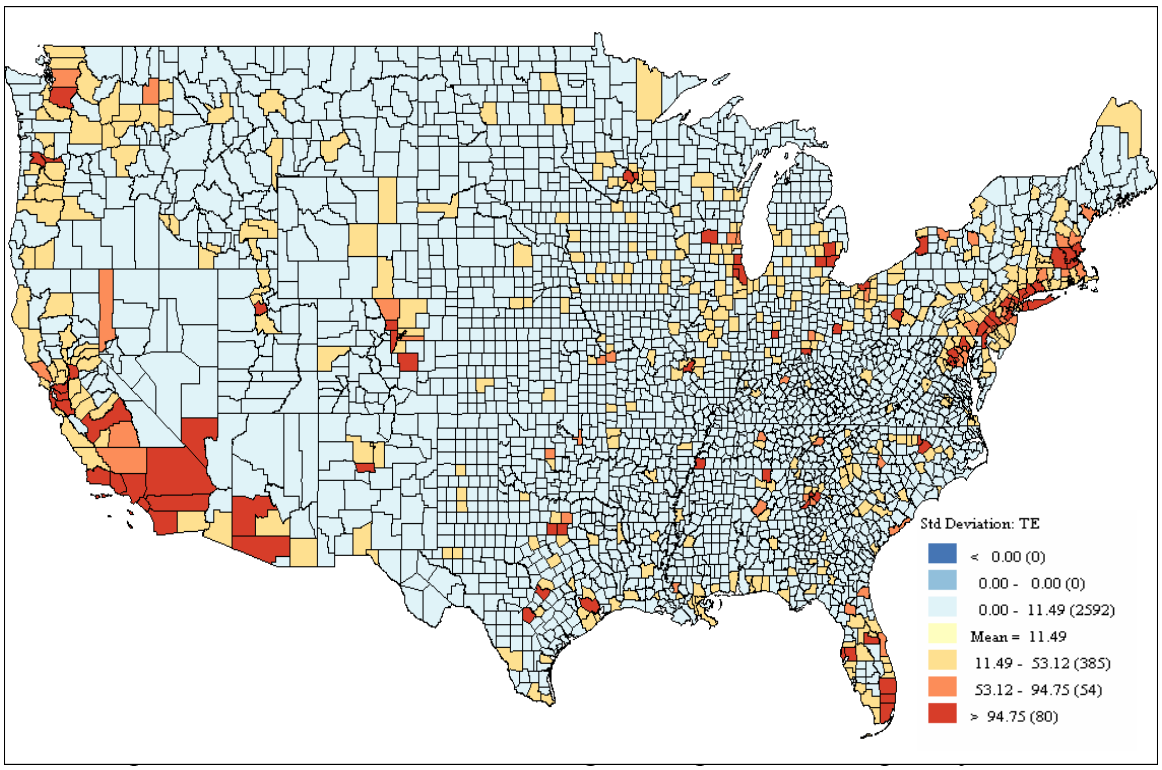
3.4 Spatial Exploratory Analysis

The spatial distribution of biotech firms based on 2003 county business pattern data is presented in figure 3.4. The figure illustrates standard deviations of biotech establishments with the mean of biotech establishments equal to 11.49. A high concentration of firms is seen in the Northeast and West, as well as in major metropolitan cities, which involve 519 counties, accounting for 16.68 percent of the total observations. Most counties that are without or not adjacent to a major metropolitan city are also without a biotech firm. These regions compose the remaining 2,592 counties, accounting in terms of the number of biotech establishments, where each of the top thirty counties include at least one major city in 2003. Figure 3.6 indicates the mean change in the number of biotech establishments (1998-2003) equal to 1.15. Most of the new biotech firms established between 1998 and 2003 are located in counties associated with major metropolitan cities. This implies that, the U.S. biotech industry exhibits a spatial pattern and it is not independently distributed over space.

The spatial association of biotech firms is tested using a global Moran's I, which measures similarities and dissimilarities in biotech establishments across neighboring counties (Anselin, 1995). For the number of biotech establishments, *y*, Morans'I is:

$$I = \left(\frac{n}{\sum_i \sum_j w_{ij}} \right) \times \left(\frac{\sum_i \sum_j (y_i - \mu)(y_j - \mu)}{\sum_i (y_i - \mu)^2} \right) \tag{3.1}$$

where wij indicates elements of the spatial weight matrix W (Rook contiguity weight matrix) between two points (i & j), μ the mean of all y observations, and i, j=1,…,n. A positive and significant value for Moran's I indicate positive spatial correlation,



Note: TE represents total number of biotech establishments in each county

Figure 3.4. Spatial Distribution of the U.S. Biotech Establishments, 2003.

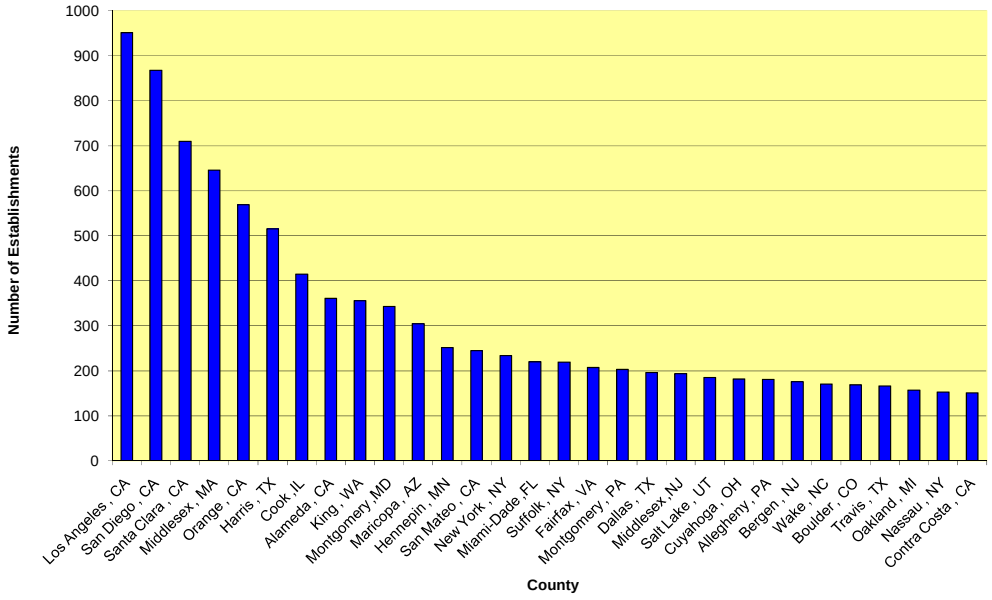


Figure 3.5. U.S. Biotech Establishments by County, 2003.

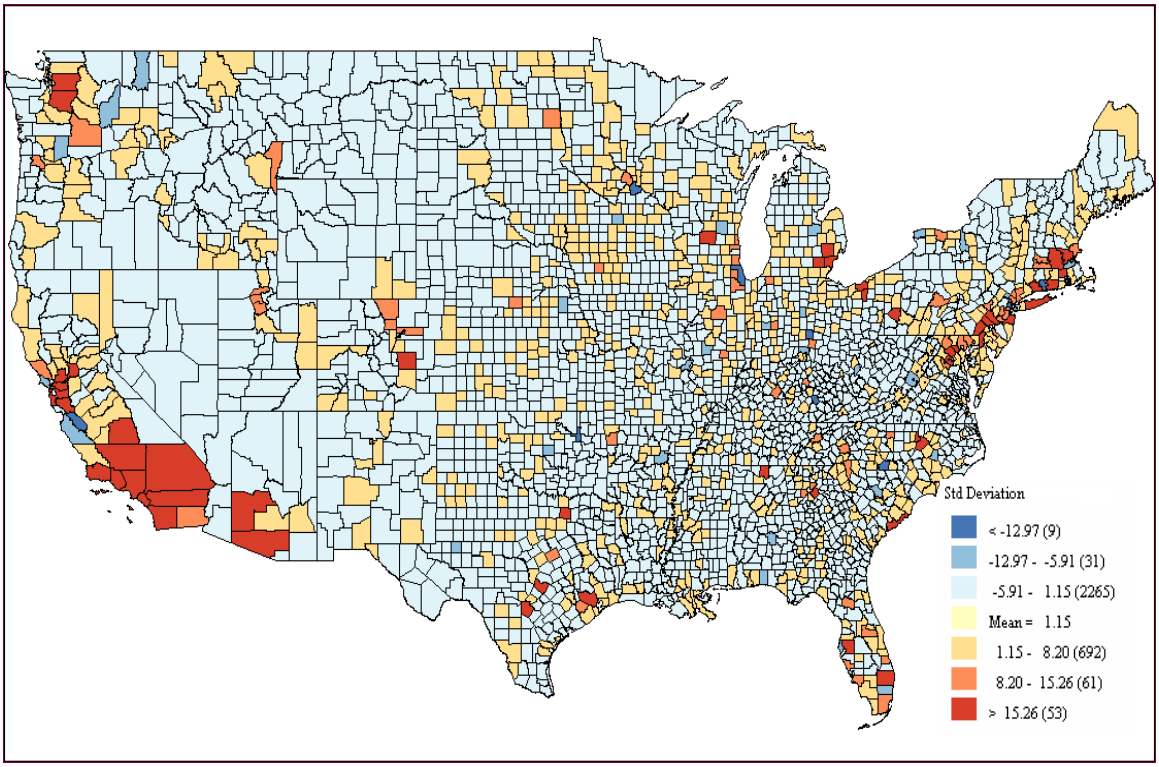


Figure 3.6. Spatial Distribution of Change in the Number of U.S. Biotech Establishments from 1998-2003.

showing that counties with a high or low number of establishments are similar to their neighboring counties. Conversely, a negative and significant value for Moran's I indicates negative spatial correlation, showing that counties with a high or low number of establishments are unlike their neighboring counties (Pacheco and Tyrrell, 2002). We calculate Moran's I for the 2003 number of biotech establishments across all contiguous U.S. counties, employing GeoDa, spatial data analysis software. The Moran's I statistic is equal to 0.3058, indicating a significant strong positive spatial relationship. However, in the case of uneven spatial clustering, global spatial indicators such as Moran's I are found to be less useful. This resulted in a new general class of local spatial indicators such as Local Indicators of Spatial Association (LISA, also known as Local Moran), which measures the contribution of individual counties to the global Moran's I statistic (Anselin, 1995). The LISA statistic is calculated for the i^{th} county as:

(3.2) $$I_i = z_i \sum_j w_{ij} z_j$$

where w_{ij} indicates elements of the spatial weight matrix $\boldsymbol{W}$ (Rook contiguity weight matrix) between two points (i & j), and z_i and z_j indicates the standardized number of establishments for county i and j, respectively. The sum of LISAs ($\sum_i I_i$) for all observations is proportional to global Moran's I, implying that LISA statistic can be interpreted as indicators of local spatial clusters and as diagnostics for local instability (spatial outliers) (Anselin, 1995).

Figure 3.7 illustrates the biotech industry clusters produced by LISA. It indicates the locations with a significant Local Moran statistic classified by type of spatial correlation: (a) high-high association (HH), a county with many biotech firms has neighboring counties with many biotech firms; (b) low-low association (LL), a county with few biotech firms has neighboring counties with few biotech firms; (c) low-high association (LH), a county with few biotech firms has neighboring counties with many biotech firms; and (d) high-low association (HL) , a county with many biotech firms has neighboring counties with few biotech firms. The HH and LL locations suggests clustering of similar values (positive spatial correlation), whereas the HL and LH locations indicate spatial outliers (negative spatial correlation) (Anselin, 1995). A positive and high autocorrelation is found in California, the Northeast, as well as in major cities such as Seattle, Portland, Salt Lake City, Phoenix, Denver, Houston, Dallas, Minneapolis, Chicago, Detroit, Cleveland, Cincinnati, Atlanta, Miami, Orlando, Tampa, and Raleigh. Figure 3.8 indicates locations with a significant Local Moran statistic. It illustrates that Imperial County (CA), Orange County (CA), Riverside County (CA), San Mateo County (CA), Lee County (KY), and Grant County (WV) have the most

significant biotech clusters among all U.S. counties, which is indicated by a p-value of 0.0001.

In addition to the spatial autocorrelation, multivariate spatial correlation is also analyzed employing a multivariate Moran's I statistic. The multivariate spatial correlation "centers on the extent to which values of one variable (z_k) observed at a given location show a systematic (more than likely under spatial randomness) association with another variable (z_l) observed at the neighboring locations." (Anselin, Syabri, and Smirnov, 2002). The multivariate Moran's I is as follows:

(3.3) $$I_{kl} = \frac{z_k^{'} W z_l}{n}$$

Where n indicates the number of observations, $\boldsymbol{W}$ indicates rook contiguity weight matrix, and z_k and z_l indicate standardized variables with mean zero and standard

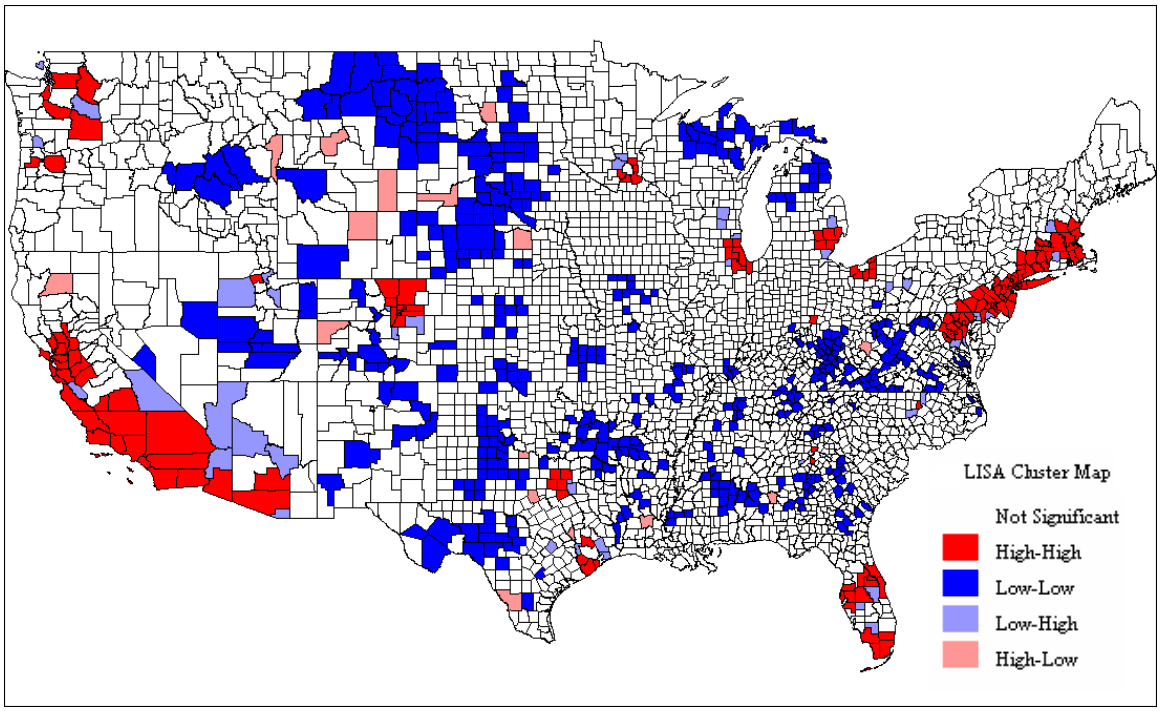


Figure 3.7. Local Indicator of Spatial Association (LISA) Cluster Map for Biotech Establishments.

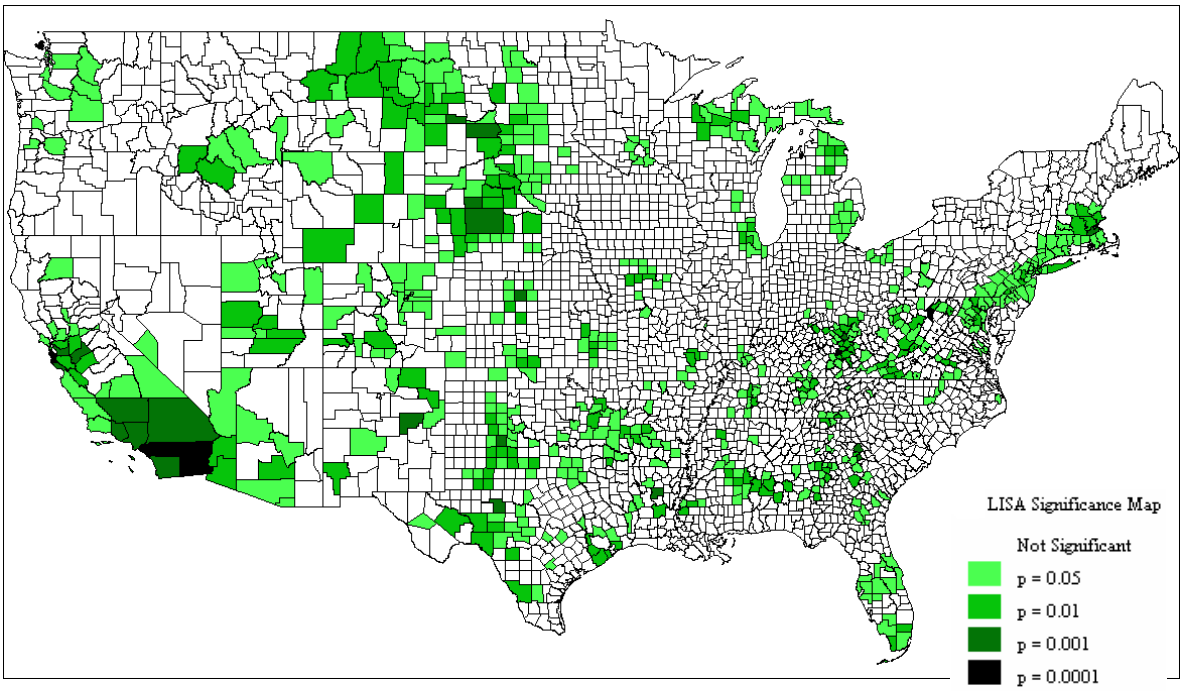


Figure 3.8. Local Indicator of Spatial Association (LISA) Significance Map for Biotech Establishments.

deviation equal to one (Anselin, Syabri, and Smirnov, 2002). Using a similar rationale as in the development of LISA, a Multivariate Local Moran Statistic (MLMS) was developed by Anselin, Syabri, and Smirnov (2002). This is defined as follows:

(3.4) $$I_{kl}^{i} = z_{k}^{i} \sum_{j} w_{ij} z_{l}^{j}$$

where w_{ij} indicates elements of the spatial weight matrix, $\boldsymbol{W}$ (Rook contiguity weight matrix) between points (i & j), and z_k^i and z_l^j indicates the standardized variables for county *i* and *j*, respectively. The MLMS "gives an indication of the degree of linear association (positive or negative) between the value for one variable at a given location *i* and the average of *another* variable at neighboring locations." (Anselin, Syabri, and Smirnov, 2002). Similar to LISA, MLMS suggests two classes of positive spatial correlation, or spatial clusters (HH and LL), and two classes of negative spatial correlation, or spatial outliers (HL and LH) (Anselin, Syabri, and Smirnov, 2002).

Since the affect of agricultural production on the location of the agricultural biotech industry is of potential interest to the agricultural economists, this study analyzes spatial correlation between *Farmland* and *spatial lag* of the number of agricultural biotech establishments (agricultural feedstock subsector). The Multivariate Global Moran's I statistic is equal to 0.0218, indicating a significant positive spatial relationship between agricultural production and the location of the agricultural biotech industry. The MLMSs cluster map indicates a positive spatial correlation in California, and parts of Midwest, West, Texas and Florida, whereas a negative spatial correlation is seen mainly in the Northeast and Central U.S., and parts of Midwest, West and Florida (figure 3.9). The significance of the MLMSs spatial clusters is illustrated in figure 3.10. This indicates

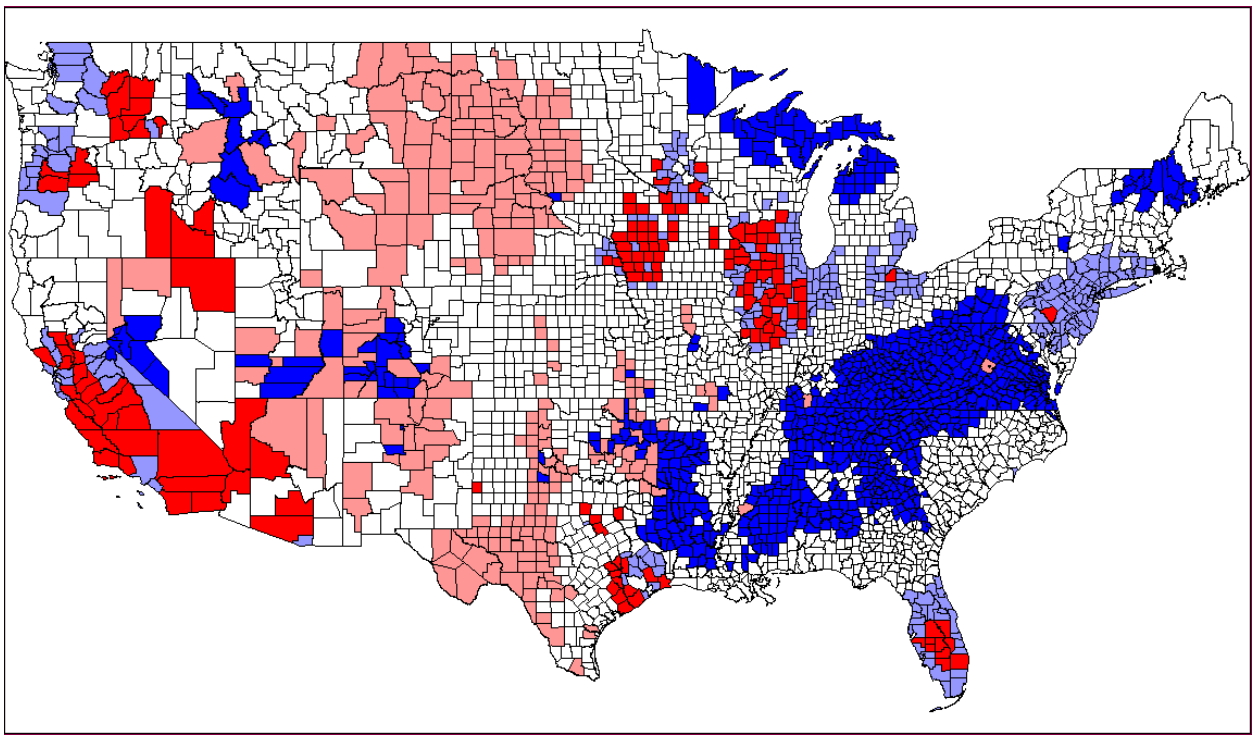


Figure 3.9. Bivariate LISA Cluster Map for Farmland and Spatial Lag of Biotech Establishments.

that globally, agricultural production and the agricultural biotech industry have a positive relation; moreover, locally there are certain counties where spatial association is very significant.

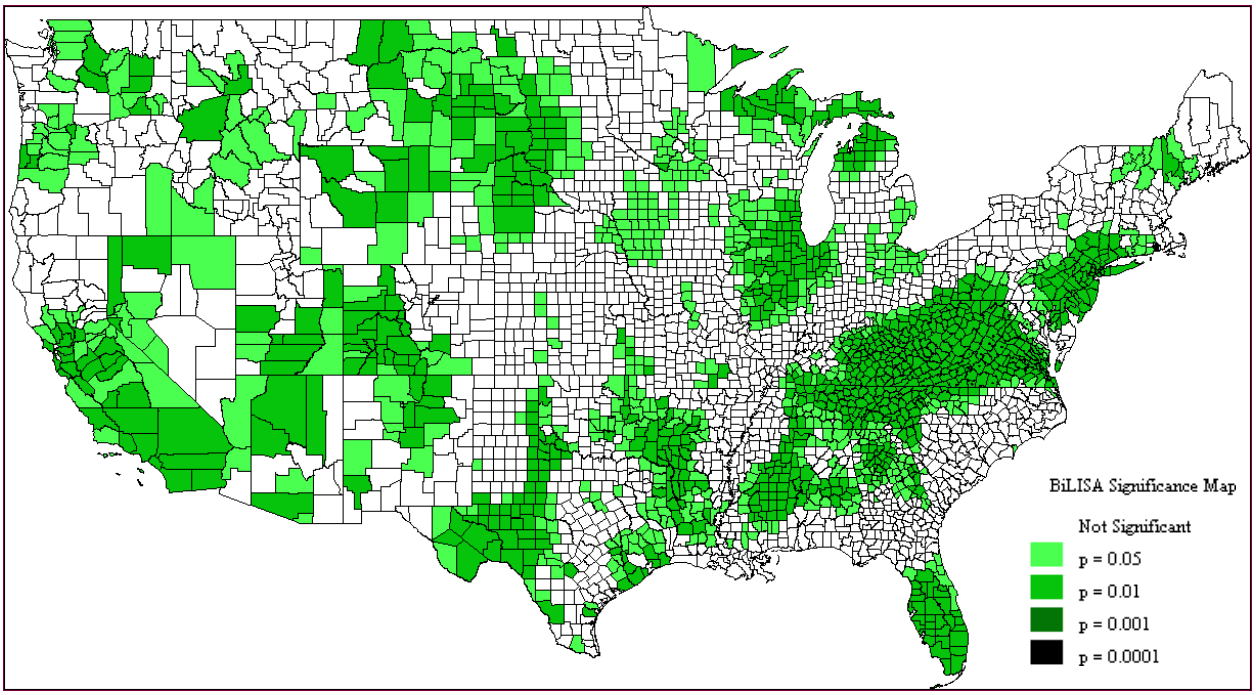


Figure 3.10. Bivariate LISA Significance Map for Farmland and Spatial Lag of Biotech Establishments.

3.5 Econometric Model

Most of the previous empirical studies on industry location have employed non-spatial econometric models, such as Ordinary Least Squares (OLS), Poisson, Negative Binomial, and Tobit (for recent exception see Roe, Irwin, and Sharp,2002; Goetz and Rupasingha, 2002; Isik, 2004; and Sambidi and Harrison, 2005). Since the data are collected with reference to points in space, employing OLS (and models mentioned above) as an econometric tool will produce spatially autocorrelated residuals, resulting in biased estimates and all inferences based on the model may be incorrect (Anselin, 1988; LeSage, 1999).

The spatial correlation of OLS residuals (without spatial component) is formally tested employing different spatial correlation indices (Morans-I, Wald, Lagrange multiplier and the Likelihood ratio test statistic), as suggested by LeSage (1999). All the tests indicated the presence of spatially correlated residuals in the regression model (Table 3.3). To overcome this problem of spatial autocorrelation, different approaches were undertaken, which involved spatial weights matrices (for example, spatial autoregressive model (SAR) and Spatial Error Model (SEM)). However, spatial correlation of data involving discrete dependent variables, have received little attention in the literature. In contrast to general spatial models, estimation of spatial discrete models yields a non-spherical variance covariance matrix, resulting in a heteroskedastic error term (Anselin 2002; Fiva and Rattsø, 2005). To solve this problem, McMillen (1992) employed an error model (EM) algorithm approach to estimate the SAR and SEM probit models containing spatial heteroskedasticity. However, McMillen's EM estimator is associated with certain drawbacks (LeSage, 2000), which were overcome by LeSage (2000), who developed a Gibbs sampling approach to estimate heteroskedastic spatial autoregressive and spatial error probit and tobit models.

Table 3.3. Tests for Spatial Correlation in Residuals of a Regression Model.

Test Statistics	Value
Moran's I Statistic	9.112***
Lagrange Multiplier	77.370***
Likelihood Ratio	69.765***
Wald	826.400***

Note: *** indicates significance at the 1% level

The dependent variable (number of biotech establishments) contains 561 observations with zeros, and a data range of 0 to 951. In this study the number of biotech establishments in a given county is assumed to reflect the strength of that county in

attracting new biotech establishments. Counties without biotech establishments are considered to be undesirable for biotech firm location, but we are not certain how much undesirable those counties are. This introduces a censor data problem into the estimation and thus more traditional methods (such as count data models) cannot be used for the estimation (Long 1997). Therefore, the study uses spatial tobit model over spatial count data models for the estimation. The dependent variable accounting for the change in the number of biotech establishments ranged from -66 to +207. During 1998 to 2003, while some (887) counties lost biotech establishments, some (900) had zero biotech establishments in 1998 and 2003 or the numbers did not change over the period. Goetz and Rupasingha (2002) indicated that the local factors explaining the growth of an industry are different from those explaining decline. Therefore, negative values of this dependent variable are recoded to zero because we are focusing on determining the factors affecting the location of new biotech establishments. Theoretically the tobit method is considered to be the most suitable approach when the underlying dependent variable contains negative values that have been censored (clustered) to zero in the empirical apprehension of the variable (Sigelman and Zeng 1999). The spatial tobit model with a spatial lag variable is as follows:

$$y = \rho \mathbf{W} y + \mathbf{X}\beta + \mu$$

$$\mu \sim N(0, \sigma^2 I_n) \tag{3.5}$$

where y is a n x 1 vector of the dependent variable (either the number of establishments in 2003 or change in the number of establishments between 1998 and 2003), ρ is the scalar for spatial lag coefficient, $\mathbf{W}$ is the n x n spatial weigh matrix, β is the k x 1 parameter vector, $\mathbf{X}$ is the n x k matrix of exogenous explanatory variables, $\boldsymbol{\mu}$ is an n x 1

vector of normally distributed error terms with zero mean and variance σ^2. The model is estimated employing a Bayesian estimation method provided by LeSage's econometric toolbox (2005). The Bayesian approach is a Gibbs sampling (Markov Chain Monte Carlo (MCMC)) method, which allows prior knowledge to be introduced when available, or implements diffuse priors in model estimation (LeSage 2000). The study refers to LeSage (2000) and Casella and George (1992) for detailed presentation of Bayesian Spatial Tobit model and Gibbs sampling MCMC method, respectively. The Gibbs sampling approach introduces a conditional distribution for the censored dependent variable conditional on all other parameters in the model (Fiva and Rattsø 2005). Once the unobserved latent variables are generated they are used in place of the censored observations of the dependent variable in the tobit model (LeSage 2000). The Bayesian approach of the spatial tobit model relaxes the assumption of constant variance of error terms made by the maximum likelihood estimation even after controlling for spatial dependence (LeSage 2000; Fleming 2004).

3.6 Results

Table 3.4 presents the parameters estimates and marginal effects of the spatial and standard tobit models of the biotech establishments along with the marginal effects. Tables 3.5-3.9 presents the parameter estimates and marginal effects of the spatial and standard tobit model of the change in the number of biotech establishments for the U.S., South, Northeast, Midwest, and the West, respectively. Relatively small standard errors for the spatial tobit model, compared to the standard tobit, indicated that the former is a better fit. Moreover, the highly significant spatial lag parameter (ρ) (except the Northeast equation) suggests that inference based on the standard tobit specification without a spatial correction, is not valid for the data under consideration.

Table 3.4. Estimates of Factors Affecting the Location of U.S. Biotech Industry.

	Coefficients		Marginal Effects	
Variable	**Spatial Tobit**	**Tobit**	**Spatial Tobit**	**Tobit**
Constant	-2.6763***	-42.2665***	-0.4902	-27.0928
	(0.9169)	(5.3442)		
Poverty	0.0099	0.3575*	0.0018	0.2291
	(0.0241)	(0.1655)		
Population	2.90E-05***	0.0970***	5.31E-06	0.0622
	(9.31E-06)	(0.0056)		
Median Housing Value	-1.68E-05***	-0.0033	-3.08E-06	-0.0021
	(3.67E-06)	(0.0197)		
Unemployment	-0.0536*	-0.2853	-0.0098	-0.1829
	(0.0366)	(0.2878)		
Median Household Income	0.0001***	0.2934**	1.63E-05	0.1881
	(2.22E-05)	(0.1189)		
Venture Capital	0.4466***	0.4667***	0.0818	0.2991
	(0.0849)	(0.0572)		
Colleges	0.5343***	1.5797***	0.0979	1.0126
	(0.1074)	(0.2166)		
Hospitals	0.4801***	0.2760	0.0879	0.1769
	(0.0876)	(0.3258)		
Average Wage Per job	1.41E-05	0.5906***	2.57E-06	0.3786
	(1.46E-05)	(0.1113)		
Education	0.0504***	0.4267***	0.0092	0.2735
	(0.0144)	(0.0880)		
Property Tax	2.33E-05***	0.0078	4.26E-06	0.0050
	(7.39E-06)	(0.0221)		
Crime Index	0.0001	0.0032***	1.23E-05	0.0020
	(0.0001)	(0.0011)		
Farm Land	1.66E-07	-0.0006***	3.05E-08	-0.0004
	(1.64E-07)	(0.0001)		
Metro-NonMetro	0.2688*	2.1332*	0.0492	1.3674
	(0.1767)	(1.0991)		
Life Science R&D	-0.0010	-0.0192***	-0.0002	-0.0123
	(0.0008)	(0.0047)		
NIH	0.0009**	0.0157***	0.0002	0.0101
	(0.0005)	(0.0027)		
Degrees in Biological Science	0.0043	0.0355	0.0008	0.0228
	(0.0126)	(0.0870)		
Biological Scientists in workforce	-0.0005	0.0381*	-0.0001	0.0244
	(0.0031)	(0.0200)		
rho	0.0242***			
	(0.0078)			

Note: *,**,*** Statistical significance at the 10, 5, and 1% levels, respectively.
Values in the parenthesis indicate standard errors.

The spatial lag coefficient (ρ) is positive and significant at the 1% level, except for the change in biotech establishments' equation of the West, where it was found to be significant at the 10% level. This result indicates the presence of spatial agglomeration

economies for the spatial structure of the biotech industry. The positive sign indicates that the spatial distribution of biotech firms is positively correlated across counties. The spatial agglomeration parameter being positive and significant in the change in biotech establishments' equation indicates the willingness of new biotech firms to locate in regions with existing biotech companies. As mentioned earlier, agglomeration factors result in economies of scale, which create a favorable infrastructure for new and existing biotech firms. Hence, counties producing biotech products tend to be concentrated across regions in order to utilize positive externalities associated with agglomeration economies.

Most of the county-level variables in the spatial tobit model of the biotech establishments have the expected signs and are significantly different from zero. A county's population had an expected positive sign, and was found to be significant at the 1% level. This is in accordance with the present spatial distribution of biotech firms, which are located mainly in major metropolitan cities that are highly populated. However, change in population was found to be insignificant in the spatial tobit models of the change in biotech establishments, except the South where it was found to be positive and significant at the 5% level.

The median housing value had a negative and significant impact on the location of biotech firms, indicating that biotech firms avoid locating in a county with high housing values. Among the models accounting for the change in biotech establishments, the change in median housing value variable was found to be negative and significant (at the 10% level) in the Northeast and West. This result is some what surprising since most of the biotech firms are located in the urban areas of the Northeast and West, where housing costs are considered to be high. However, increasing housing values might be considered as one of the primary reasons for the emergence of new biotech regions such

as, Dallas, Houston, San Antonio, Memphis, Richmond, and Miami-Dade (McCandless 2005). Moreover, Goetz and Rupasingha (2002) also found that the median housing values have a negative and significant impact on the growth of U.S. high-tech firms, which prefer to locate in metropolitan areas.

Table 3.5. Estimates of Factors Affecting the Change in New U.S. Biotech Establishments.

	Coefficients		Marginal Effects	
Variable	Spatial Tobit	Tobit	Spatial Tobit	Tobit
Constant	-0.4813***	-7.6844***		
	(0.1957)	(0.8019)	-0.2612	-2.6687
Poverty	0.0969***	0.4296***		
	(0.0301)	(0.1305)	0.0526	0.1492
Population	1.26E-07	2.17E-06***		
	(2.41E-07)	(8.19E-07)	6.82E-08	7.54E-07
Median Housing Value	-1.88E-06	5.00E-06		
	(3.14E-06)	(1.20E-05)	-1.02E-06	1.74E-06
Unemployment	0.0248	0.1080		
	(0.0268)	(0.1160)	0.0135	0.0375
Median Household Income	-1.15E-05	0.0001		
	(2.31E-05)	(0.0001)	-6.24E-06	0.0000
Venture Capital	0.0359	0.4638***		
	(0.0328)	(0.0662)	0.0195	0.1611
Colleges	0.2169***	1.6237***		
	(0.0921)	(0.1847)	0.1177	0.5639
Hospitals	0.0608	0.1853		
	(0.0765)	(0.2090)	0.033	0.0644
Average Wage Per job	0.0001***	0.0006***		
	(3.11E-05)	(0.0001)	4.01E-05	0.0002
Education	0.1199***	0.3665***		
	(0.0274)	(0.1078)	0.0651	0.1273
Property Tax	-8.13E-07**	-7.73E-06***		
	(5.00E-07)	(9.04E-07)	-4.41E-07	-2.68E-06
Crime Index	1.73E-05	0.0003***		
	(3.78E-05)	(0.0001)	9.41E-06	0.0001
Farm Land	6.68E-08	-9.71E-07		
	(4.88E-07)	(1.81E-06)	3.63E-08	-3.37E-07
Metro-Nonmetro	-0.3989***	-1.6831***		
	(0.1189)	(0.4825)	-0.2165	-0.5845
Life Science R&D	0.0002	0.0039***		
	(0.0002)	(0.0006)	8.31E-05	0.0014
rho	0.2141***			
	(0.0287)			

Note: *,**,*** Statistical significance at the 10, 5, and 1% levels, respectively.
Values in the parenthesis indicate standard errors.

The unemployment rate had an expected negative sign and was found to be significant (at the 10% level) in the spatial tobit model of the biotech establishments. This result is in accordance with the fact that high unemployment rate reflects a lower local quality of life or a weak economy, which is not generally preferred by biotech firms (Goetz and

Table 3.6. Estimates of Factors Affecting the Change in New Biotech Establishments in the South.

	Coefficients		Marginal Effects	
Variable	**Spatial Tobit**	**Tobit**	**Spatial Tobit**	**Tobit**
Constant	-0.3353*	-6.0870***		
	(0.2544)	(0.9855)	-0.1933	-0.0006
Poverty	0.0782**	0.3817**		
	(0.0400)	(0.1530)	0.0451	3.82E-05
Population	9.20E-07**	6.00E-06***		
	(3.99E-07)	(1.56E-06)	5.30E-07	6.00E-10
Median Housing Value	-1.97E-06	3.05E-06		
	(7.61E-06)	(2.86E-05)	-1.14E-06	3.05E-10
Unemployment	-0.0493**	-0.1098		
	(0.0306)	(0.1251)	-0.0284	-1.10E-05
Median Household Income	-1.50E-06	0.0001		
	(3.41E-05)	(0.0001)	-8.67E-07	8.18E-09
Venture Capital	-0.0149	0.1454		
	(0.0409)	(0.0952)	-0.0086	1.45E-05
Colleges	0.1141	1.2735***		
	(0.1122)	(0.2954)	0.0657	0.0001
Hospitals	0.0774	0.3476		
	(0.1034)	(0.3033)	0.0446	3.48E-05
Average Wage Per job	3.73E-05	0.0004***		
	(3.52E-05)	(0.0001)	2.15E-05	4.11E-08
Education	0.1051***	0.4605***		
	(0.0389)	(0.1447)	0.0606	4.60E-05
Property Tax	-2.97E-07	-4.99E-06**		
	(8.41E-07)	(2.36E-06)	-1.71E-07	-4.99E-10
Crime Index	-0.0001*	3.08E-05		
	(0.0001)	(1.19E-04)	-4.78E-05	3.08E-09
Farm Land	1.05E-06	2.84E-06		
	(9.80E-07)	(4.27E-06)	6.05E-07	2.84E-10
Metro-Nonmetro	-0.5040***	-1.4942***		
	(0.1508)	(0.3985)	-0.2905	-0.0001
Life Science R&D	0.0002	0.0016*		
	(0.0002)	(0.0009)	0.0001	1.56E-07
rho	0.2110***			
	(0.0399)			

Note: *,**,*** Statistical significance at the 10, 5, and 1% levels, respectively.
Values in the parenthesis indicate standard errors.

Rupasingha, 2002). The results indicate that, as the unemployment rate in a particular county increases by 1%, possibility of locating a biotech firm in that county decreases by 0.0098 units. Conversely, the variable was found to be insignificant in the standard tobit model, which failed to account for spatial autocorrelation. Among the change in biotech

Table 3.7. Estimates of Factors Affecting the Change in Biotech Establishments in the West.

	Coefficients		Marginal Effects	
Variable	**Spatial Tobit**	**Tobit**	**Spatial Tobit**	**Tobit**
Constant	-0.4925	-4.6392		
	(0.6648)	(3.4836)	-0.2639	-1.5575
Poverty	0.0063	0.2808		
	(0.0892)	(0.5223)	0.0034	0.0943
Population	8.75E-07	1.66E-05**		
	(1.44E-06)	(6.85E-06)	4.69E-07	5.58E-06
Median Housing Value	-1.29E-05*	-1.03E-04**		
	(8.24E-06)	(4.70E-05)	-6.94E-06	-3.47E-05
Unemployment	0.2398***	1.2523**		
	(0.0903)	(0.5172)	0.1285	0.4204
Median Household Income	0.0001	0.0007		
	(0.0001)	(0.0004)	4.61E-05	0.0002
Venture Capital	0.0661	1.2022***		
	(0.1084)	(0.2153)	0.0354	0.4036
Colleges	0.7215***	0.4966		
	(0.3027)	(0.5778)	0.3866	0.1667
Hospitals	0.0837	-0.2694		
	(0.2925)	(0.8160)	0.0449	-0.0905
Average Wage Per job	0.0002**	0.0003		
	(0.0001)	(0.0005)	0.0001	0.0001
Education	0.0241	0.0825		
	(0.0768)	(0.4556)	0.0129	0.0277
Property Tax	-5.77E-06***	-2.61E-05***		
	(2.73E-06)	(4.02E-06)	-3.09E-06	-8.75E-06
Crime Index	-0.0002*	0.0006		
	(0.0002)	(0.0004)	-1.24E-04	0.0002
Farm Land	-4.50E-07	-4.92E-06		
	(7.76E-07)	(3.34E-06)	-2.41E-07	-1.65E-06
Metro-Nonmetro	-1.0068**	-4.7887*		
	(0.4748)	(2.4826)	-0.5395	-1.6077
Life Science R&D	0.0006*	0.0043**		
	(0.0004)	(0.0019)	0.0003	0.0014
rho	0.0900*			
	(0.0545)			

Note: *,**,*** Statistical significance at the 10, 5, and 1% levels, respectively.
Values in the parenthesis indicate standard errors.

establishment models, the change in unemployment variable was found to be negative and significant (at the 5% level) in the South, whereas it was found to be positive and significant (at the 1% level) in the West. The reason for this can be attributed to the fact that unemployment rate is high in the South compared to the West. High unemployment rate in the South indicates high poverty and low standards of living, which is not

Table 3.8. Estimates of Factors Affecting the Change in Biotech Establishments in the Midwest.

	Coefficients		Marginal Effects	
Variable	**Spatial Tobit**	**Tobit**	**Spatial Tobit**	**Tobit**
Constant	-0.5450	-3.2227***		
	(0.4836)	(0.8961)	-0.2932	-1.3707
Poverty	0.1185**	0.1639		
	(0.0630)	(0.1242)	0.0638	0.0697
Population	7.09E-08	-5.41E-07		
	(3.73E-07)	(6.34E-07)	3.81E-08	-2.30E-07
Median Housing Value	2.42E-06	3.90E-05**		
	(9.02E-06)	(1.68E-05)	1.30E-06	1.66E-05
Unemployment	0.0611	0.1619		
	(0.0697)	(0.1375)	0.0329	0.0689
Median Household Income	-8.91E-05**	-0.0003***		
	(4.37E-05)	(0.0001)	-4.79E-05	-0.0001
Venture Capital	0.1439*	0.3045**		
	(0.1015)	(0.1277)	0.0774	0.1295
Colleges	0.0290	0.9265***		
	(0.2104)	(0.2749)	0.0156	0.3940
Hospitals	-0.0392	0.2563*		
	(0.1222)	(0.1518)	-0.0211	0.1090
Average Wage Per job	0.0001*	0.0003***		
	(0.0001)	(0.0001)	4.62E-05	0.0001
Education	0.1552***	0.3204***		
	(0.0460)	(0.0852)	0.0835	0.1363
Property Tax	-8.97E-07	-2.15E-06**		
	(9.49E-07)	(1.09E-06)	-4.83E-07	-9.16E-07
Crime Index	3.02E-05	-1.19E-05		
	(5.95E-05)	(6.58E-05)	1.62E-05	-5.05E-06
Farm Land	-6.53E-07	-2.39E-07		
	(1.65E-06)	(2.88E-06)	-3.52E-07	0.00
Metro-Nonmetro	0.0243	-0.4070		
	(0.2167)	(0.3935)	0.0131	-0.1731
Life Science R&D	-0.0008	-0.0003		
	(0.0006)	(0.0012)	-0.0004	-0.0001
rho	0.3983***			
	(0.0585)			

Note: *,**,*** Statistical significance at the 10, 5, and 1% levels, respectively.
Values in the parenthesis indicate standard errors.

preferred by biotech firms. Conversely, low unemployment rate in the West indicates lack of skilled labor. Therefore, biotech firms locating in the West prefer to locate in counties with high unemployment rate. Thus, change in unemployment rate is considered to have a parabolic effect on the location of biotech industry.

Table 3.9. Estimates of Factors Affecting the Change in Biotech Establishments I in the Northeast.

	Coefficients		Marginal Effects	
Variable	**Spatial Tobit**	**Tobit**	**Spatial Tobit**	**Tobit**
Constant	-1.3988	-2.0580		
	(1.6577)	(3.5053)	-0.5508	-0.0002
Poverty	1.0563***	2.0950***		
	(0.3624)	(0.7616)	0.4160	0.0002
Population	3.62E-07	3.61E-06**		
	(8.12E-07)	(1.81E-06)	1.42E-07	3.61E-10
Median Housing Value	-2.39E-05*	-2.92E-05		
	(1.56E-05)	(3.73E-05)	-9.42E-06	-2.92E-09
Unemployment	0.1131	-0.0412		
	(0.2787)	(0.6718)	0.0445	-4.12E-06
Median Household Income	0.0004*	0.0009**		
	(0.0002)	(0.0004)	0.0001	8.91E-08
Venture Capital	0.4703***	1.2762***		
	(0.1469)	(0.2389)	0.1852	0.0001
Colleges	-0.4130	1.0210		
	(0.3787)	(0.6454)	-0.1627	0.0001
Hospitals	0.4907*	1.2000*		
	(0.3841)	(0.6919)	0.1932	0.0001
Average Wage Per job	-3.19E-05	-0.0005		
	(0.0003)	(0.0006)	-1.26E-05	-5.20E-08
Education	0.2650	0.3274		
	(0.2489)	(0.5413)	0.1044	3.27E-05
Property Tax	-3.45E-06***	-9.29E-06***		
	(1.17E-06)	(2.03E-06)	-1.36E-06	-9.29E-10
Crime Index	0.0001	0.0004		
	(0.0002)	(0.0003)	3.66E-05	3.50E-08
Farm Land	8.86E-06	2.54E-05		
	(2.55E-05)	(0.0001)	3.49E-06	2.54E-09
Metro-Nonmetro	0.0754	-1.8444**		
	(0.7469)	(0.9432)	0.0297	-0.0002
Life Science R&D	0.0021**	0.0027		
	(0.0011)	(0.0024)	0.0008	2.70E-07
rho	0.0100			
	(0.1073)			

Note: *,**,*** Statistical significance at the 10, 5, and 1% levels, respectively.
Values in the parenthesis indicate standard errors.

A county's median household income, which reflects the local standard of living, was found to have a positive and significant impact on the location of biotech firms. This

result is consistent with previous literature, which indicates biotech firms' preference for locating in regions with a high standard of living and well developed infrastructure. The change in median household income variable was found to have a positive and significant (at the 10% level) impact on change in biotech firms in the Northeast, conversely, it had an unexpected significant negative sign in the model for the Midwest. Availability of local venture capital was also found to be positive and significant, indicating biotech firms' dependency on local financial sources. As the number of venture capital firms in a given county increases by one unit, the chances of locating a biotech firm in that county increases by 0.082 units. The change in number of local venture capital firms' variable was found to be critical mainly in the Northeast and Midwest, where it was found to be positive and significant at the 1% and 10% level, respectively.

The number of colleges and hospitals in a given county were also found to have a positive and significant impact on the location of the biotech industry. As the number of colleges and hospitals in a particular county increase by one unit, the number of biotech firms in that county increases by 0.098 and 0.088 units, respectively. Also noteworthy is that coefficient of hospitals in spatial tobit model is high compared to the one in the standard tobit model, which indicated a downward bias of the estimate. In the model accounting for the change in biotech establishments in the U.S., the change in number of colleges' variable was found to be positive and significant at the 1% level, whereas, the change in number of hospitals variable was found to be insignificant. The change in number of colleges variable was found to be positive and significant (at the 1% level) in the West (insignificant in the rest of the regions), whereas, the change in number of hospitals variable was found to be positive and significant in the Northeast (insignificant in the rest of the regions). This result is in accordance with the fact that biotech firms in

the Northeast are mainly involved in the manufacturing and marketing of new drugs; conversely, firms in the West are mainly engaged in R&D activities, which involve research institutes.

In the model for the biotech establishments, the education variable which reflects the labor quality in a given county was positive, as expected, and was found to be significant at the 1% level. This result indicates biotech companies' preference for counties with a skilled labor pool. However, the average wage per job variable was found to be insignificant in the location of biotech industry. In the model for change in biotech establishment (U.S.), both change in education and average wage per job variables were found to be positive and significant at the 1% level. Among the regional models, the change in average wage per job variable was found to be positive and significant in the West and Midwest regions, conversely, the change in education variable was found to be positive and significant in South and Midwest.

The property tax variable, which was used as a proxy for high standard of living, was found to have an expected positive sign and was significant at the 1% level. Conversely, the change in property tax variable was found to have a negative and significant impact in the models for the change in biotech establishments in the U.S., Northeast, and West regions. This result indicates that increase in local property taxes especially, in the Northeast, might be one of the reasons for the emergence of new biotech regions other than the West and Northeast. The urbanization economy (Metro-Nonmetro) variable had an unexpected positive sign in the biotech establishments' model and was found to be significant at the 10% level. Conversely, Goetz and Rupasingha (2002) indicated that rural counties have a negative and significant impact on the location of high-tech firms, which includes drugs and pharmaceuticals and R&D services. The

reason for this may be attributed to the fact that, biotech establishments in our model include firms that are related to agriculture (agricultural feedstock and chemicals), and are assumed to be located in close proximity to rural areas, where farm production is high. Therefore, the sign of urbanization economies variable may be changed by the inclusion of agricultural biotechnology firms. However, the urbanization economy variable (*Metro-Nonmetro*) was found to have a negative and significant impact in the change in biotech establishment models of the U.S., South, and West, indicating the new biotech firms' preference to locate in metro counties rather than non-metro counties.

A county's poverty rate was found to have a insignificant impact on the location of the biotech industry. Conversely, the change in poverty rate variable was found to be positive and significant in all the change in biotech establishment models, except the West. A county's crime index was found to be insignificant in the site-selection of biotech firms, which is consistent with Goetz and Rupasingha's (2002) findings. However, the variable was found to be negative and highly significant in the standard tobit model, indicating a upward bias of the estimate that failed to correct for spatial autocorrelation. Among the spatial tobit models of the change in biotech establishments' the change in crime index variable was found to be negative and significant in the South and West. This indicates that the new biotech firms' willing to locate in the South or West, prefer to locate in counties that have low crime rate.

The amount of farmland in a given county was found to be insignificant in the location of biotech industry. In addition, the change in farmland variable was also found to be insignificant in all the change in biotech establishments' models. The state level variables (*R&D, NIH, Biological Degree, and Biological Scientists*) were found to be insignificant, except for *NIH*, which was found to be positive and significant at the 5%

level. The reason for the insignificance of the rest of the variables may be attributed to the fact that they are all measured at the state level and were not able to capture the county level effects. The change in *R&D* expenditures variable was found to be positive and significant in the change in biotech establishment models of the West and Northeast. This result is in accordance with the current spatial concentration of biotech firms in the West and Northeast, which are involve states spending large amount of money on research and development.

3.7 Conclusions

Over the past two decades the U.S. biotech industry has experienced significant growth, resulting in an increase in size and number of establishments. Currently, several state and local economic development agencies are designing and implementing strategies to attract new biotech firms, resulting in stiff competition among and within states. As a result of this increasing competition, the U.S. biotech industry is experiencing some changes in its geographical distribution. However, only some new state/regions are likely to attract biotech firms, as most biotech firms are tending to cluster along existing biotech regions. Several studies have analyzed the location aspects of the biotech industry, however, our understanding of the spatial influence on the regional distribution of biotech establishments, is anecdotal. This study employs a Bayesian spatial tobit model that analyzes factors affecting site-selection of the U.S. biotech industry taking the spatial affect into consideration. The study examines the impacts of agglomeration factors, infrastructure factors, and local economic and socioeconomic factors on the county-level biotech establishments. A total of twenty five biotech related industries were analyzed in the study.

The hypothesis of spatial agglomeration economies is confirmed for the spatial structure of the biotech industry, indicating that biotech firms are positively correlated

across counties, resulting in clustering of biotech production. Availability of venture capital firms, research institutions, and hospitals were found to have the most significant impact on the location of biotech firms. This indicates that the biotech firms prefers to locate in regions where they have a source for financing their business , access to research institutes to collaborate with skilled labor and obtain new technology, and access to hospitals for research, testing and marketing of new biotech products.

Biotech companies also prefer to locate in counties with a well developed infrastructure. This is indicated by the positive and significant estimates of median household income, education, population, and property tax, and a negative and significant estimate of the unemployment rate. In terms of the theory of industry location, firms should prefer counties with low wages, low property taxes and high unemployment rate; however, the preference of biotech firms seen here is different. In the case of biotech industry location, these three variables are assumed to proxy the standard of living of a given county, thus, indicating their preference to locate in counties with a high standard of living.

The above findings may hinder rural areas hopes of attracting biotech firms; however, they are capable of attracting at least one category of the biotech industry (Agricultural Feedstock and Chemicals), which is involved in the agricultural and biotech activities. The rural areas may also want to target the biotech firms that are involved in the manufacturing of intermediate products and drugs that have achieved commercial scale. These types of biotech firms are found to operate in locations that are associated with low costs of production, availability of space for expansion, low median housing values and good incentives (Gray and Parker, 1998).Thus, the state and local economic development agencies should design strategies based on the type of biotech firm they want to attract.

The regional models in this study also shredded some light on the regional differences in factors affecting the location of new biotech establishments. Prospective biotech firms willing to locate in the West prefer to establish in metro-counties with easy access to research institutes and skilled labor pool, and that are associated with low median housing values, property taxes, and crime rate. Since the existing biotech firms in the Northeast are mainly associated with manufacturing and marketing of biotech products, biotech firms preferences are different in the Northeast compared to other regions. Biotech firms that are willing to locate in the Northeast prefer counties with easy access to funding sources (venture capital firms and state incentives), hospitals for research, testing and marketing of new biotech products, and the ones that are associated with low median housing value and property tax.

Future research is directed toward a separate analysis of factors affecting the location of agricultural and non-agricultural biotech firms. Including county-level variables related to the state and local economic development incentives, R&D expenditures, and environmental constraints may further enlighten our understanding of the biotech industry location.

3.8 References

Akundi, K. 2003 Cluster-based Economic Development, Part 1: A Survey of State Initiatives. Texas Economic Development, Business and Industry Data Center, www.bidc.state.tx.us.

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

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

Anselin, L., I. Syabri, and O. Smirnov. 2002. "Visualizing Multivariate Spatial Correlation with Dynamically Linked Windows." In: Proceedings of the SCISS Specialist Meeting "New Tools for Spatial Data Analysis". Santa Barbara, California, USA. May 10-11, 2002.

Battelle Technology Partnership Practice and State Science and Technology Institute. 2004. " Laboratories of Innovation: State Bioscience Initiatives 2004." Biotechnology Industry Organization, Washington D.C.

Begemann BD. 1997. "Competitive strategies of biotechnology firms: Implication for US Agriculture." Journal of Agricultural and Applied Economics 29:117-22.

Brennan, M.F., Pray, C.E., & Courtmanche, A. 1999. "Impact of industry concentration on innovation in the U.S. plant biotech industry." Paper presented at the Transitions in Agbiotech: Economics of Strategy and Policy conference, Washington, DC.

Biotechnology Industry Organization. 2005-2006. "Guide to Biotechnology." Biotechnology Industry Facts. BIO, Washington D.C.

BioWorld Publishing Group. 2005. BioWorld Today.

California Trade and Commerce Agency, Office of Economic Research. 2001. Biotechnology. May 2001.

Casella, G. and George, E. (1992) "Explaining the Gibbs Sampler," *Amer. Statistician* 46, 167-174.

Cortright, J., & Mayer, H. (2002). "Signs of Life: The Growth of Biotechnology Centers in the U.S." Portland: The Brookings Institution Center on Urban and Metropolitan Policy.

Dahlander, L. and McKelvey, M. (2003): Revisting Frequency and Spatial Distribution:Innovation Collaboration for Biotech Firms. Paper presented at DRUID's Summer 2003 Conference, Helsingore, June 12-14.

Darby M.R. and L.G. Zucker, 1996. "Star Scientists, Institutions, and the Entry of Japanese Biotechnology Enterprises," NBER Working Papers 5795, National Bureau of Economic Research, Inc.

Dibner, Mark D. 1995. Biotechnology Guide U.S.A., 3rd ed., Research Triangle Park, NC: Institute for Biotechnology Information (IBI).

Ernst & Young. 2000. "The Economic Contributions of the Biotechnology Industry to the U.S. Economy." Prepared for Biotechnology Industry Organization. Ernst & Young Economics Consulting and Quantitative Analysis.

Ernst & Young. 2004. "Resurgence: The Americas Perspective–Global Biotechnology Report 2004 Ernst & Young LLP.

Ernst & Young. 2005. "Beyond Borders: The Global Biotechnology Report." Ernst & Young LLP.

Feldman, M., 2003. “The locational dynamics of the US biotech industry: Knowledge externalities and the anchor hypothesis.” Industry & Innovation, 10(3), 311-328.

Fiva J.H. and J. Rattsø. 2005. “Decentralization with Property Taxation to Improve Incentives: Evidence from Local Governments’ Discrete Choice.” Working Paper Series No. 6/2005. Department of Economics, Norwegian University of Science and Technology.

Fleming, M. (2004) "Techniques for Estimating Spatially Dependent Discrete Choice Models". In Advances in Spatial Econometrics, L. Anselin and R. Florax (eds), Berlin: Springer-Verlag.

Geospatial and Statistical Data Center. ‘Crimes reported 1994-2002’. University of Virginia Library. http://fisher.lib.virginia.edu/collections/stats/crime/crimes94.html

Goetz, S. J. and A. Rupasingha. 2002. “High-Tech Industry Clustering: Implications for Rural Areas.” American Journal of Agricultural Economics 84(5).

Goetz, S. and S. Morgan. 1995. “State-Level Locational Determinants of Biotechnology Firms”. Economic Development Quarterly, (9):174-184.

Graff D. G. 1997. “The Agricultural Biotechnology Industry in Overview.” Agricultural and Resource Economics, University of California, Berkeley.

Gray M. and Parker E. 1998. “Industrial Change and Regional Development: The Case of the US Biotechnology and Pharmaceutical Industries.” ESRC Centre for Business Research, University of Cambridge, Working Paper No. 95.

Grudkova, V. (2001). The Technology Economy: Why do Tech Companies Go Where They Go? EDA National Forum, Washington, DC (May 30, 2001).

Isik, M.2004. “Environmental Regulation and the Spatial Structure of the U.S. Dairy Sector.” American Journal of Agricultural Economics 86: 949-962.

Jaffe, A., M. Trajtenberg, and R. Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. The Quarterly Journal of Economics 108 (3):577-598.

Kalaitzandonakes, N. and Bjornson, B. (1997). "Vertical and Horizontal Coordination in the Agro-biotechnology Industry: Evidence and Implications" Journal of Agricultural and Applied Economics 29: 129-139.

Lerner, J. and Merges, R. 1998. “The Control of Strategic Alliances: An Empirical Analysis of the Biotechnology Industry.” Journal of Industrial Economics, Vol. 46. pp. 125–156.

LeSage, J.P. 1999. "The Theory and Practice of Spatial Econometrics." Unpublished, Dept. of Econ., University of Toledo.

LeSage, J. P. 2000. "Bayesian Estimation of Limited Dependent variable Spatial Autoregressive Models", Geographical Analysis, Volume 32, number 1, p. 19-35.

LeSage, J. P. (2005), "Econometrics Toolbox", MatLab programs. Downloaded from http://www.spatial-econometrics.com.

McCandless M.E. 2005. "Biotechnology Locations: Hitting the Mark." Business Facilities. http://www.businessfacilities.com/bf_05_05_cover.asp (Date last visited 14rth May 2006).

McMillen, D. P. (1992), 'Probit with spatial autocorrelation', Journal of Regional Science 32-3, 335–348.

Munroe T., G. Craft, and D. Hutton. 2002. "A Critical Analysis of the Local biotechnology Industry Cluster — Counties of Alameda, Contra Costa, & Solano in California." Volume III, Appendix. Munroe Consulting Inc., Craft Consulting and Hutton Associates. Research Monograph prepared for a Consortium of Bay Area Organizations, May 2002.

National Science Foundation. 2006. Industrial Funding of Academic R&D Continues to Decline in FY 2004. NSF 06-315.

Pacheco, Andrada I., and Timothy J. Tyrrell. 2002. "Testing Spatial Patterns and Growth Spillover Effects in Clusters of Cities." Journal of Geographical Systems 4:275-285.

Pisano, G. P., W. Shan, and D. J. Teece. 1988. Joint ventures and collaboration in the biotechnology industry. In International Collaborative Ventures in US Manufacturing, edited by D. C. Mowery. Cambridge: Ballinger.

Prevezer, M. 1998. Clustering in biotechnology in the USA, pp. 124–193 in Swan, G. M. P., Prevezer, M. and Stout, D. K. (eds), The Dynamics of Industrial Clustering. International Comparisons in Computing and Biotechnology, Oxford, Oxford University Press.

Powell, W. W., and P. Brantley. 1992. Competitive cooperation in biotechnology:Learning through networks? In Networks and Organizations. Structure, Form andAction, edited by R. Eccles. Boston: Harvard Business School Press.

Powell W. W., K.W. Koput, J. Bowie,. and L. Smith-Doerr. 2002. "The Spatial Clustering of Science and Capital. Journal of Regional Studies, 36 (3).

Roe, B., E. Irwin, and J. S. Sharp. 2002. "Pigs in Space: Modeling the Spatial Structure of Hog Production in Traditional and Nontraditional Production Regions." American Journal of Agricultural Economics 84: 259-278.

Sambidi P.R. and W. Harrison. 2005. "Spatial Dependency of the Geographically Concentrated U.S. Broiler Industry." Paper Presented at the American Agricultural Economics Association Annual Meeting, Providence, Rhode Island.

Santoro, Michael D., and Alok K. Chakrabarti. "Building Industry-University Research Centers: Some Strategic Considerations." International Journal of Management Reviews (September 1999): 225 - 245.

Sigelman, Lee, and Langche Zeng. 1999. Analyzing censored and sample-selected data with Tobit and Heckit models. *Political Analysis* 8:167-82.

Sporleder, T.L., L.E. Moss, and L.A. Nickles. 2002. "Agricultural Biotechnology Start-Ups: Does Venture Capital Matter?" Prepared for presentation at the Third Annual Agricultural Biotechnology Conference, Ohio's Future in Functional Foods, Reynoldsburg, OH.

Sporleder, T.L. and L.E. Moss. 2004. "Knowledge Capital, Intangible Assets, and Leverage: Evidence from U.S. Agricultural Biotechnology Firms." Paper Presented at the 2004 Symposium of the International Food and Agribusiness Management Association Meetings, Switzerland, June 12-13, 2004.

U.S. Census Bureau. 2003. County Business Pattern. Data Available at http://censtats.census.gov/cbpnaic/cbpnaic.shtml

U.S. Department of Agriculture. 2003. Economic Research Service. Data Available at http://www.ers.usda.gov/Data/

___. National Agricultural Statistics Service. Historical Data. Available at http://www.nass.usda.gov/Data_and_Statistics/index.asp

U.S. Department of Labor. Bureau of Labor Statistics. Available at http://www.bls.gov/home.htm

Xia, Y., and Buccola, S. 2005. University life science programs and agricultural biotechnology. American Journal of Agricultural Economics, 87(1), 229-243.

Yang H. and S. Buccola. 2003. "University-Industry Relationships and the Design of Biotechnology Research." Paper Presented at the American Agricultural Economics Association Annual Meeting, Montreal, Canada, July 2003.

Zeller, C., 2001, Clustering biotech: A recipe for success? Spatial patterns of growth of biotechnology in Munich, Rhineland and Hamburg, Small Business Economics, 17, 123-141.

Zucker, L. G., M. R. Darby, and M. B. Brewer. 1998. Intellectual human capital and the birth of U.S. biotechnology enterprises. The American Economic Review 88:290-306.

Zucker L. G. & M.R. Darby & J.S. Armstrong, 2003. "Commercializing knowledge: university science, knowledge capture and firm performance in biotechnology," Proceedings, Federal Reserve Bank of Dallas, issue Sep, pages 149-170.

CHAPTER 4

SPATIAL CLUSTERING OF INNOVATIVE ACTIVITIES: A CASE OF U.S. BIOTECH RELATED RESEARCH AND DEVELOPMENT ACTIVITIES

4.1 Background

Over the past few decades there has been considerable scientific interest in understanding the forces that contribute to the clustering of innovative activities and growth potential of firms and regions. Similar to other high-tech firms, Research and Development (R&D) firms involved in biotech activities are receiving a wide spread attention based on their potential for exceptional growth. Compared to other traditional industries, R&D firms are considered to agglomerate and form clusters in specific areas. Knowledge spillovers are considered to be one of the main reasons for this geographical concentration of R&D firms (Acs et al. 1992, 1994; Feldman 1994a, b; and Audretsch and Feldman 1996). The role of these spillovers, which generate increasing returns and eventually economic growth, is the central theme of the new theory of endogenous growth (Romer 1986, 1990; Grossman and Helpman 1991).

A recent stream of literature involving economic geographers and location theorists have focused on analyzing different forms of agglomeration economies to better understand the theory behind industry location. The concept of agglomeration economies indicate that the performance of one firm is influenced by other firms located nearby. If a firm benefits by locating near an existing firm, it indicates positive economies of scale. Conversely, if a firm is deterred by locating near an existing firm, it indicates negative economies of scale. Agglomeration economies are further divided into Localization economies and Urbanization economies. Localization economies involve technical externalities and knowledge spillovers (Marshall-Arrow-Romer (MAR) externalities) that

are specific to an industry whereby the productivity or growth of a firm in a given industry in a given region is assumed to increase the performance of others firms in that industry (van Oort, 2004). The externalities and spillovers include: the formation of a skilled labor pool and the production of new ideas (based on accumulation of human capital and face to face communication) and the availability of specialized input services (Marshall, 1920; Ottaviano and Thisse, forthcoming). This type of economies results in specialization of a region with respect to a specific industry (Paci and Usai, 1999). Urbanization economies reflect economic externalities that are transferred to firms of different industries as a result of savings from large scale operation of a city as a whole (Jacobs 1969; van Oort, 2004). It indicates economies of scale associated with generalized location factors such as good infrastructure, favorable community attitude, tax credits and subsidies, and favorable socioeconomic factors. These factors are not specific to a particular industry, but, favor any kind of industry. This in turn results in sectoral diversity, and that is why we see a wide variety of industries in major metropolitan areas. This type of economy is considered to be more significant for high-tech industries, which are involved in innovative activities and depend on knowledge spillovers from outside the core industry (Henderson et al., 1995; Paci and Usai 1999).

Generally speaking, there are two sources of spillovers: one is private-sector spillovers, which include existing R&D facilities, while the other is public-sector spillovers, which include public research institutions, such as universities, medical centers, and government laboratories (Kyle 2004; Furman et al. 2004). Krugman (1991) argue that knowledge spillovers especially, tacit knowledge from public and private R&D facilities to third-party firms are restricted by the distance between them. In fact, despite the technology advancement, knowledge is still considered to be costly and difficult to

transfer (Jaffe et al., 1993; von Hippel, 1994). This indicates that proximity and location play a key role in the success of a R&D firm (Audretsch and Feldman 1996). However, Furman et al. (2006) argue that the effectiveness of spillovers is not only dependent on spatial proximity, but also on the types of investment made by the recipient firm (Cohen and Levinthal, 1998) and on the nature of contractual agreement between the knowledge source and recipient. For instance, Zucker, Darby and Armstrong (1998) in their study indicated that the success of biotech firms is not driven by simple geographic proximity to research institutes with "star scientists', but rather by the contractual agreements between particular firms and "stars" (Furman et al. 2006). Moreover, Toole (2003) argue that tacit knowledge can only be restricted for a short period of time since valuable thoughts and methods spread quickly (Feldman 2003). In this paper we try to incorporate these spatial issues in econometric modeling and estimate the influence of localization and urbanization economies on the spatial distribution of biotech related R&D and testing facilities.

Our objectives in this paper are twofold. First, we examine the inter- and intra-industry spatial association of the biotech related R&D and testing facilities across all contiguous U.S. counties. Second, we employ spatial econometrics to analyze the extent to which numerous firm-specific, location-specific, and inter- and intra-industry spatial agglomeration factors affect the location, movement and concentration of biotech related R&D and testing facilities. This type of econometric modeling will explicitly consider the potential for spatial effects such as spatial autocorrelation and spatial heterogeneity that may invalidate the interpretation of standard econometric analyses based on contiguous cross-sectional data (Anselin et al. 2000). In most of the previous literature directed toward location of high-tech industries, these effects are typically excluded or treated

inappropriately (for recent exceptions, see Paci and Usai 1999; Anselin et al. 2000; Goetz and Rupasingha 2002).

4.2 Location Choices for R&D Firms

Gray and Parker (1998) argue that there are two choices for emerging R&D firms to locate. One is to locate in new regions without industrial history and where they can develop their own production unit and maximize their profits by restricting the spread of their knowledge. These firms can also collaborate with local research institutes; however, they may have to set some contractual agreements to prevent the spread of knowledge to other emerging R&D firms that may wish to grasp that knowledge and develop their own facility in that region. Research institutes play a critical role in the spatial distribution of R&D firms. They are considered to be the source of basic research and high skilled labor (Anselin et al. 2000) for the high-tech industries. Knowledge created from basic research at universities is considered to be a public good; therefore, the resulting positive externalities are not locally bounded, but can freely move across borders to private sector in the form of spillovers (Anselin et al. 1997).

The second choice is to locate in regions with well established firms, which create an infrastructure suitable for the new firms. In the case when emerging R&D firms believe incoming spillovers from existing firms are significant, they may tend to engage in cooperative R&D agreements (Cassiman and Veugelers 2002). Feldman et al. (2004) indicate that co-location of R&D firms facilitates knowledge spillovers by providing opportunities for both planned and unanticipated interactions. This type of knowledge spillovers are considered to be less costly compared to the ones generated internally, or sourced externally through contractual agreement (Feldman et al. 2004). Aharonson et al. (2004) found that a biotech firm R&D investments is more productive when it is located

close to other biotech firms that are working on similar issues. The authors also found that the productivity of a biotech firm is enhanced by its own R&D alliance and also R&D alliances of other firms located within a cluster. However, if the existing firms avoid cooperation, this may hinder the productivity of new firms. Kyle (2004) argue that proximity to R&D facilities of competing biotech firms that are publishing many scientific papers has some negative impact on R&D activities of a new or existing biotech firm. The author indicates two reasons for this negative impact. First, the firm, which is lagging behind, may cut back on its patenting efforts and reallocate its drug discovery expenditures from patenting to other functions. The second reason could be that a competitor may lock the access to research institutes for other firms, by making contractual agreements with researchers of that institute (Kyle 2004 and Furman et al. 2004).

4.3 Data

The present study analyzed several categories of variables that are considered to affect the location of biotech related R&D and testing facilities, such as agglomeration factors, infrastructure factors, and local economic and socioeconomic factors. County-level data was obtained from the 2003 county business patterns (U.S. Census Bureau), Economic Research Service, National Agricultural Statistics Service, and U.S. Dept. of Labor. The dependent variable considered in the study is county level number of R&D and testing facilities (NAICS codes: 541380, 541710) in 2003.

Economies of scale associated with agglomeration factors are believed to be one of the driving forces in the geographical distribution of innovative activities, such as R&D and testing facilities (Pisano, Shan, and Teece, 1988; Audretsch and Feldman 1996; Gray and Parker, 1998; Paci and Usai 1999; Goetz and Rupasingha, 2002; Furman et al.

2005). Agglomeration economies indicate that performance of one R&D firm is influenced by the other R&D firm located nearby. The resulting spillovers may be due to an already existing industry specific infrastructure, which is associated with lower transaction costs, proximity to research institutions and specialized intermediate industries, good transportation facilities, and availability of skilled labor pool and financial resources. As mentioned earlier, R&D firms generate externalities and spillovers, which tend to be spatially proximate to where they were created, resulting in positive economies of scale for firms located in that region (Jaffe, Trajtenberg, and Henderson, 1993; Dahlander and McKelvey, 2003). Zeller (2001) analyzed the spatial clustering of biotech firms in Germany and argued that while codified knowledge and technology transfer often transmitted at low cost on a global scale, the exchange of tacit knowledge is facilitated more effectively by spatial proximity. In this study, we include a spatial lag variable as a proxy for agglomeration economies that accounts for R&D and testing facilities in neighboring counties. The variable is hypothesized to have a positive effect on the location of R&D and testing facilities.

A factor that is considered to be a prerequisite for attracting a R&D firm is proximity to research universities, which are responsible for knowledge spillovers and the supply of skilled labor pool. Industry funded university research increased from $630 million in 1985 to $2.1 billion in 2004 (National Science Foundation, 2006), indicating an increasing affiliation between university and industry in technology advancement. Several studies have analyzed the relationship between university research and high-tech innovations (Audretsch and Feldman 1996; Anselin et al. 1997; Zucker, Darby, and Brewer, 1998; Prevezer, 1998; Anselin et al. 2000; Zucker, Darby, and Armstrong, 2003; Dahlander and McKelvey, 2003; Kyle 2004; Furman et al. 2005; Xia and Buccola, 2005).

Zucker et al. (1998) have noted that growth in biotechnology firms is largely influenced by the presence of "star" scientists at universities. Furman et al. (2006) argue that the spillovers from research universities (public spillovers) may play a major role in driving private sector productivity, as compared to the spillovers from other R&D firms (private spillovers). Generally speaking, R&D firms prefer to locate in close proximity of research institutes, because universities do not generally seek to secure the full value of the knowledge they create, and are likely to promote the spread (public spillovers) of that knowledge (Kyle 2004). The reason is that researchers at universities are generally rewarded for publishing their work and for the influence that work has on future research in that area (Kyle 2004). However, Anselin et al. (2000) argue that public spillovers are specific to certain industries. According to the authors, public spillovers are not significant in the Drugs and Chemicals (SIC28) and in the Machinery (SIC35) sectors, where as, they are evident in the Electronics (SIC36) and the Instruments (SIC38) sectors. In order to analyze the influence of research universities on the location of R&D firms, this study includes county-level number of colleges, universities, and professional schools (*colleges*) as a proxy for the proximity to research universities. We assume that it will have a positive effect on the location of R&D and testing firms. County-level data for this variable was collected from county business patterns (U.S. Census Bureau 2003) with 611310 NAICS classification, which includes number of colleges, universities, and professional schools.

Biotech related R&D and testing firms are mainly involved in the development of a new product, once the new product is developed and tested, it is produced in relatively large batches at regional manufacturing firms (pilot plants). These manufacturing firms are generally located in close proximity of R&D and testing firms, because

manufacturing the new biotech products is increasingly complex and difficult, and require a set of highly skilled workforce (Pisano, 1997; Lee and Burrill, 1995; Gray and Parker, 1998). Once the new biotech product gets FDA approval its production shifts from pilot plants to the commercial plants, where it is produced on a large scale. Commercial production of the new biotech product involves two stages: 1) production of high-value intermediate active ingredient, and 2) formulation and packaging of the final product (Gray and Parker 1998). Generally, the first stage is performed at manufacturing plants that are nearby R&D firms, whereas the second stage is performed at manufacturing plants located in low-cost regions (Gray and Parker 1998). In this study, we include county-level number of biotech related manufacturing firms (*Manufacturing*)[1] as one of the explanatory variables and hypothesize that it will have a positive effect on the location of R&D and testing firms in that region.

Businesses that provide venture capital are considered to be an important source of capital, especially, for new and small firms (Powell, Koput, Bowie, and Smith-Doerr, 2002). For a small R&D firm, availability of venture capital in a particular region is as important as the strong research capacity of that region. During 2004, venture capital accounted for approximately 23.5% of the total finance for the biotech industry (BIO, 2005; BioWorld, 2005). Most of the biotech related R&D firms are small and operate at a loss, spending large amount of money on research and development for several years, before earning a profit (Cortright and Mayer, 2002). For example, only 1 in 5,000 potential new medicines reach the pharmacy shelf, and that is after 12 to 15 years of R&D with an average expenditure of $500 million (California Trade and Commerce

[1]the biotech related manufacturing firms belong to the following NAICS codes: 311221, 311222, 311223, 325193, 325199, 325221, 325222, 325311, 325312, 325314, 325320, 424910, 325411, 325412, 325413, 325414,339111, 339112, 339113, 339114, 334510, 334516, and 334517.

Agency, 2001). As a result, most of the biotech related R&D firms depend on venture capital funds, on research contracts and equity investment from large R&D firms, and on sales of their company stock in public markets (Cortright and Mayer, 2002). Therefore, the availability of local venture capital firms (*Venture Capital*) is hypothesized to have a positive effect on the location of R&D and testing firms in that region (data for county-level number of venture capital establishments was collected from county business patterns with 523910 NAICS classification).

. Agriculture is an important component of the biotech industry. Some studies have analyzed issues related exclusively to the spatial distribution of agricultural biotech R&D firms and its relationship with research institutions (Kalaitzandonakes and Bjornson, 1997; Graff, 1997; Begemann 1997; Brennan, Pray, and Courtmanche, 1999; Sporleder, Moss, and Nickles, 2002; Yang and Buccola, 2003; Sporleder and Moss, 2004; Xia and Buccola, 2005). The primary goal of agricultural biotech related R&D firms is to develop high yielding varieties with improved resistance to natural enemies (e.g. pest, diseases, weeds, and adverse growing conditions), and better quality and longer shelf life for fruits and vegetables. Since some of the biotech related R&D firms seek applications directed toward agricultural production, it is hypothesized that, in order to gain positive economies of scale (low transaction costs), R&D firms prefer to locate in regions with significant agricultural production (*Farmland*) (county level data for farmland was collected from U.S. Census of Agriculture). Similarly, since the biotech industry involves drugs and pharmaceutical firms, and medicinal devices and equipment firms, we hypothesize that a county with more hospitals (*Hospitals*) will have a positive effect on the location of R&D and testing firms (U.S. Census Bureau, 1998 and 2003) with 622110 NAICS classification, which includes number of general medical and surgical hospitals.

. In terms of conventional location theory, local property taxes (*Property Tax*) may discourage new investment by increasing the costs of production. However, in case of high-tech firms (such as R&D firms), high property taxes are considered to be proxies for greater availability or higher quality of local public goods (Goetz and Rupasingha 2002), which in turn reflects a high standards of living of the local community. Therefore, we assume that property taxes are positively correlated with the location of R&D firms (county-level data for property tax collections was obtained from the U.S. Census Bureau). Similarly, counties with high unemployment rate (*unemployment*) and poverty rate (*poverty*), which reflect a low standard of living of the local community, are considered to have a negative effect on the location-decision of R&D firms(county-level data for unemployment and poverty rate are collected from Economic Research Service USDA and U.S. Bureau of Census, respectively). Similarly, counties with higher crime rates (*Crime Index*) are also considered to have a negative impact on the location of R&D firms. As indicated earlier, the R&D industry is mainly concentrated in the metropolitan areas, which account for 92% of the total number of R&D and testing facilities in 2003. County-level data for crime index was obtained from Geospatial and Statistical Data Center, University of Virginia Library. In order to measure the impact of economies on the location of R&D firms, we group U.S. counties into Metro (counties belonging to rural-urban continuum codes of 1-3) and Non-metro (counties belonging to rural-urban continuum codes of 4-9). The study assigns a value of 1 for non-metro counties and 0 for metro counties. Since most of the existing R&D firms are located in the metropolitan cities, the study hypothesize that the variable (*Metro-Nonmetro*) will have a negative impact on the location-decision of the new R&D firms.

The impact of labor quality on the location decision of R&D firms is measured by county-level average wage per job (*Wage*) and percentage of persons with a college degree (*Education*) (Zucker, Darby, and Brewer, 1998). Both variables are considered to have a positive relationship with the site-selection of the R&D industry. County-level data for *wage* and *education* are obtained from the Bureau of Economic Analysis and Economic Research Service USDA, respectively. R&D firms prefer to locate in highly populated centers (*Population*) as it provides appropriate services, such as contracting for site building, major equipment, and availability of housing (McCandless, 2005). Counties with median household incomes (*Income*), which represents a high standard of living, are considered to have a favorable impact on R&D firms' site-selection (county-level data for population and median household income was obtained from U.S. Census Bureau and Economic Research Service USDA, respectively). Similarly, median housing values (*Housing Value*) are used as a proxy for the quality of housing in a give county. County-level data for median housing value was obtained from U.S. Census Bureau.It is expected to have a positive impact on the location-decision. County-level data for median housing value was obtained from U.S. Census Bureau. Table 4.1 presents the descriptive statistics of all the variables included in the model.

4.4 Spatial Exploratory Analysis

The spatial distribution of R&D firms based on 2003 county business pattern data is presented in figure 4.1. The figure illustrates standard deviations of R&D and testing establishments with the mean equal to 6.05. A high concentration of firms is seen in the Northeast and West, as well as in major metropolitan cities, which involve 408 counties, accounting for 13 percent of the total observations. Most counties that are without or not adjacent to a major metropolitan city, are also without R&D and testing firms. Figure 4.2

presents the top twenty U.S. counties in terms of the number of R&D and testing establishments, where each of the top twenty counties included at least one major city. This implies that, the U.S. R&D and testing industry exhibit's a spatial pattern, and it is not independently distributed over space.

Table 4.1. Summary Statistics of Variables.

Variable	Mean	Std. Dev.
Total Number of R&D and Testing Establishments (number)[1]	6.06	28.41
Nonmetro-Metro (dummy)	0.65	0.48
Poverty Rate (percent)	13.36	4.89
Population (number in 1000s)	92.92	303.47
Owner-Occupied Housing Units: Median Value (in $1000)	80.93	41.94
Unemployment Rate (percent)	5.97	1.96
Median Household Income (in $1000)	36.73	9.28
Total Number of Venture Capital Firms (number)	1.89	11.45
Number of Colleges, Universities and Professional Schools (number)	1.08	4.86
Number of General Medical and Surgical Hospitals (number)	1.75	3.70
Average Wage per Job (in $1000)	27.05	5.79
Percent of Persons with a College Degree (percent)	16.51	7.80
Property Tax (in $1000)	22.74	23.41
Crime Index (index in 100)	35.95	139.73
Land in Farm Acres (in 1000 acres)	301.14	385.21

Note: [1] text in parenthesis indicates units of measurement.

The spatial association of R&D firms is tested using a global Moran's I, which measures similarities and dissimilarities in R&D establishments across neighboring counties (Anselin, 1995). For the number of R&D establishments, *y*, Morans'I is:

$$I = \left(\frac{n}{\sum_i \sum_j w_{ij}} \right) \times \left(\frac{\sum_i \sum_j (y_i - \mu)(y_j - \mu)}{\sum_i (y_i - \mu)^2} \right) \tag{4.1}$$

where w_{ij} indicates elements of the spatial weight matrix, $\boldsymbol{W}$ (Rook contiguity weight matrix) between points *i* and *j*, μ the mean of all *y* observations, and i, j=1,…,n. A positive and significant value for Moran's I indicate positive spatial correlation, showing that counties with a high or low number of establishments are similar to their neighboring counties. Conversely, a negative and significant value for Moran's I indicates negative

spatial correlation, showing that counties with a high or low number of establishments are unlike their neighboring counties (Pacheco and Tyrrell, 2002).

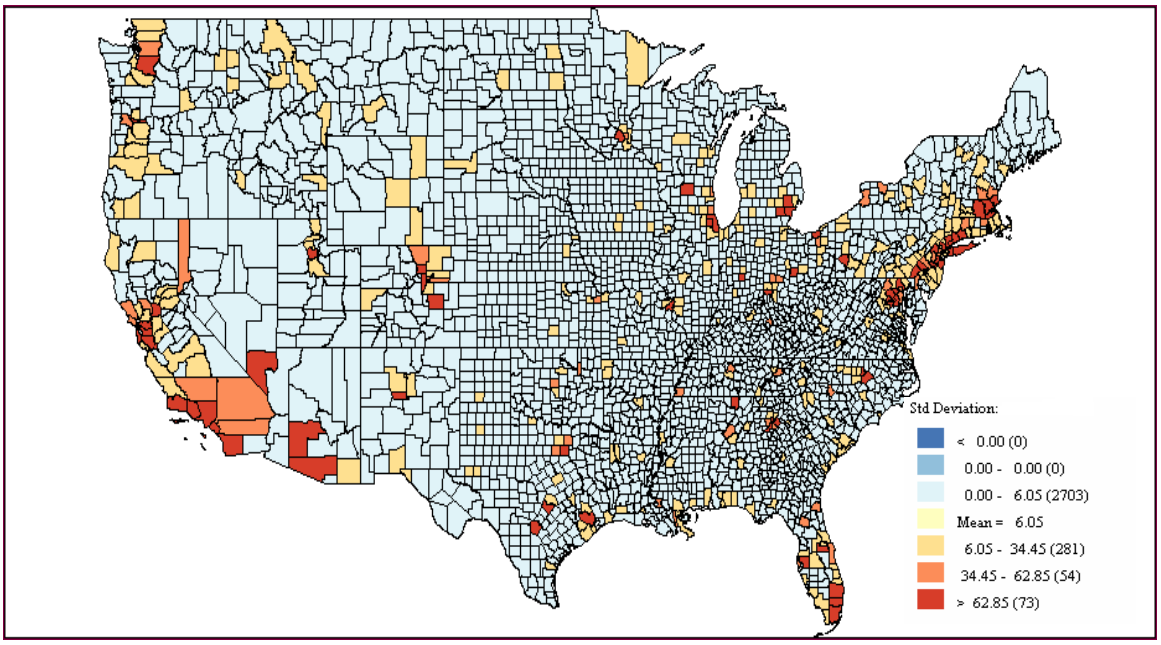


Figure 4.1. Spatial Distribution of R&D and Testing Laboratories, 2003.

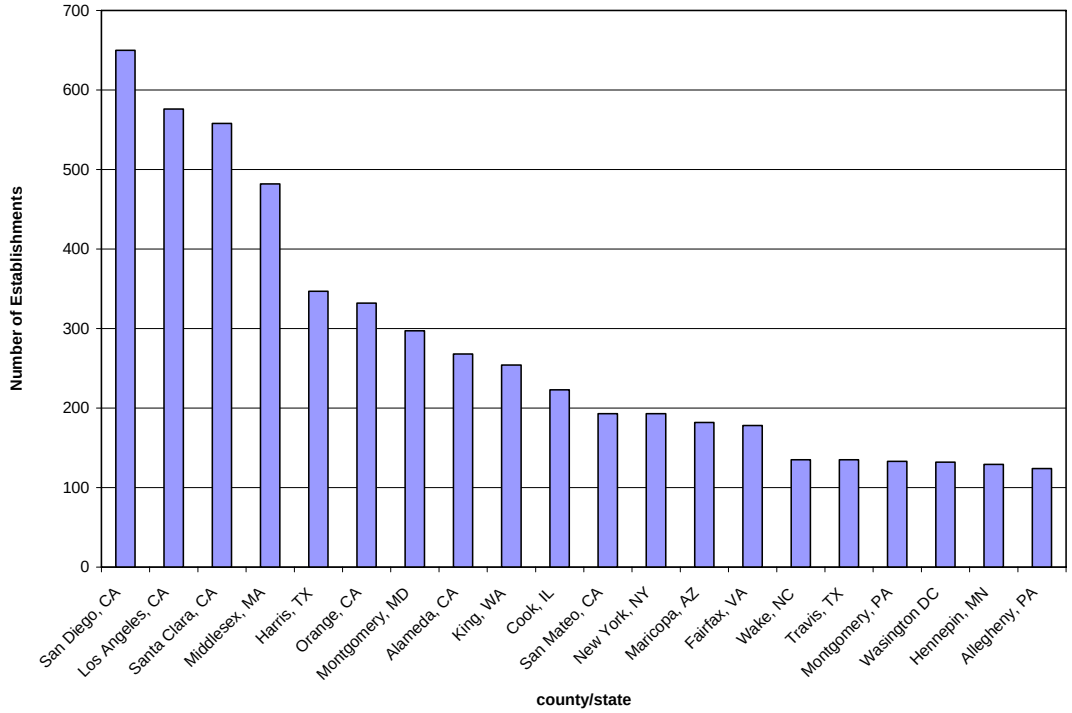


Figure 4.2. Top 20 U.S. Counties with R&D and Testing Laboratories, 2003.

We calculate Moran's I for the 2003 number of R&D and testing establishments across all contiguous U.S. counties, employing GeoDa, a spatial data analysis software. The Moran's I statistic is equal to 0.2528, indicating a significant strong positive spatial relationship. However, in the case of uneven spatial clustering, global spatial indicators such as Moran's I are found to be less useful. This resulted in a new general class of local spatial indicators such as Local Indicators of Spatial Association (LISA, also known as Local Moran), which measures the contribution of individual counties to the global Moran's I statistic (Anselin, 1995). The LISA statistic is calculated for the i^{th} county as:

$$I_i = z_i \sum_j w_{ij} z_j \tag{4.2}$$

where w_{ij} indicates elements of the spatial weight matrix, $\boldsymbol{W}$ (Rook contiguity weight matrix), and z_i and z_j indicates the standardized number of establishments for county i and j, respectively. The sum of LISAs ($\sum_i I_i$) for all observations is proportional to global Moran's I, implying that LISA statistic can be interpreted as indicators of local spatial clusters and as diagnostics for local instability (spatial outliers) (Anselin, 1995).

Figure 4.3 illustrates the R&D and testing industry clusters produced by LISA. It indicates the locations with a significant Local Moran statistic classified by type of spatial correlation: (a) high-high association (HH), a county with many R&D and testing firms has neighboring counties with many R&D and testing firms; (b) low-low association (LL), a county with few R&D and testing firms has neighboring counties with few R&D and testing firms; (c) low-high association (LH), a county with few R&D and testing firms has neighboring counties with many R&D and testing firms; and (d) high-low association (HL) , a county with many R&D and testing firms has neighboring counties with few R&D and testing firms. The HH and LL locations suggests clustering

of similar values (positive spatial correlation), whereas the HL and LH locations indicate spatial outliers (negative spatial correlation) (Anselin, 1995). A positive and high autocorrelation is found in California, the Northeast, as well as in some of the major cities such as Seattle, Portland, Salt Lake City, Phoenix, Denver, Houston, Chicago, Detroit, and Miami.

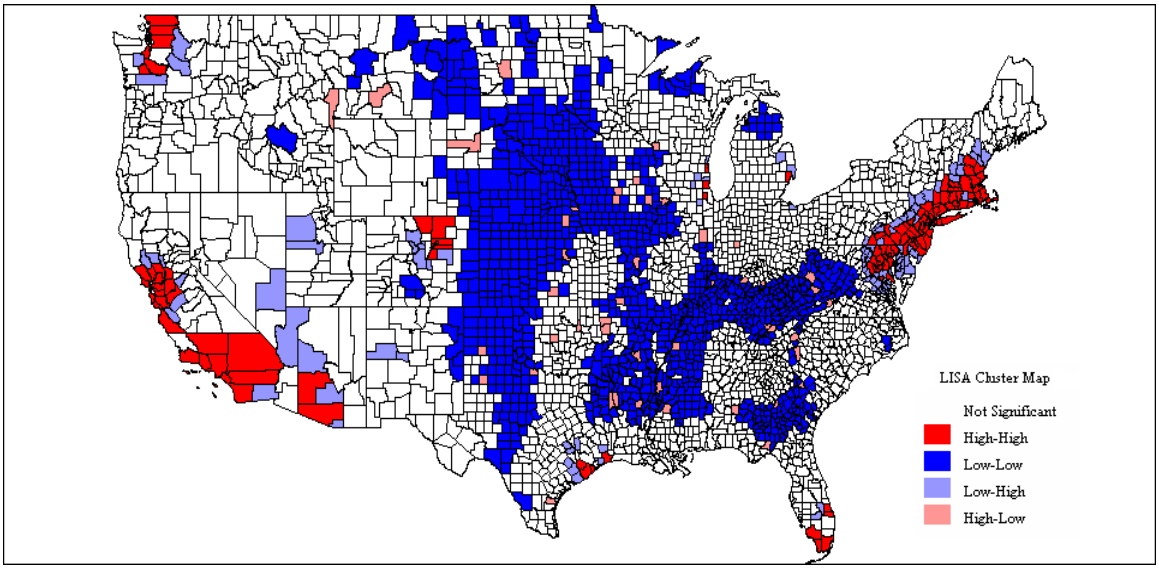


Figure 4.3. Local Indicator of Spatial Association (LISA) Cluster Map of R&D and Testing Establishments, 2003.

In addition to the spatial autocorrelation, multivariate spatial correlation is also analyzed employing a multivariate Moran's I statistic. The multivariate spatial correlation "centers on the extent to which values of one variable (z_k) observed at a given location show a systematic (more than likely under spatial randomness) association with another variable (z_l) observed at the neighboring locations." (Anselin, Syabri, and Smirnov, 2002). The multivariate Moran's I is as follows:

(4.3) $$I_{kl} = \frac{z_k' W z_l}{n}$$

Where n indicates the number of observations, $\boldsymbol{W}$ indicates rook contiguity weight matrix, and z_k and z_l indicate standardized variables with mean zero and standard deviation equal to one (Anselin, Syabri, and Smirnov, 2002). Using a similar rationale as in the development of LISA, a Multivariate Local Moran Statistic (MLMS) was developed by Anselin, Syabri, and Smirnov (2002). This is defined as follows:

$$I_{kl}^{i} = z_k^i \sum_j w_{ij} z_l^j \tag{4.4}$$

where w_{ij} indicates elements of the spatial weight matrix, $\boldsymbol{W}$ (Rook contiguity weight matrix), and z_k^i and z_l^j indicates the standardized variables for county *i* and *j*, respectively. The MLMS "gives an indication of the degree of linear association (positive or negative) between the value for one variable at a given location *i* and the average of *another* variable at neighboring locations." (Anselin, Syabri, and Smirnov, 2002). Similar to LISA, MLMS suggests two classes of positive spatial correlation, or spatial clusters (HH and LL), and two classes of negative spatial correlation, or spatial outliers (HL and LH) (Anselin, Syabri, and Smirnov, 2002).

Since the influence of research universities on the location of the R&D industry is of potential interest to the spatial and economic scientists, this study analyzes spatial correlation between *colleges* and *spatial lag* of the dependent variable (number of R&D and testing establishments). The Multivariate Global Moran's I statistic is equal to 0.2366, indicating a significant positive spatial relationship between research institutes and the location of the R&D industry. The MLMSs cluster map indicates a positive spatial correlation mainly in the Northeast, California and some of the major metropolitan cities, whereas a negative spatial correlation is indicated in some counties of the West and Northeast, and parts of Texas (figure 4.4). Similarly, we analyze the spatial

correlation between biotech related manufacturing and R&D firms. The Multivariate Global Moran's I statistic is equal to 0.3055, indicating a significant positive spatial relationship between biotech related manufacturing firms and the location of the R&D industry. The MLMSs cluster map indicates a positive spatial correlation mainly in the Northeast, California and major metropolitan cities (figure 4.5). We further analyze this relationship along with the other factors discussed earlier, employing a spatial econometric model.

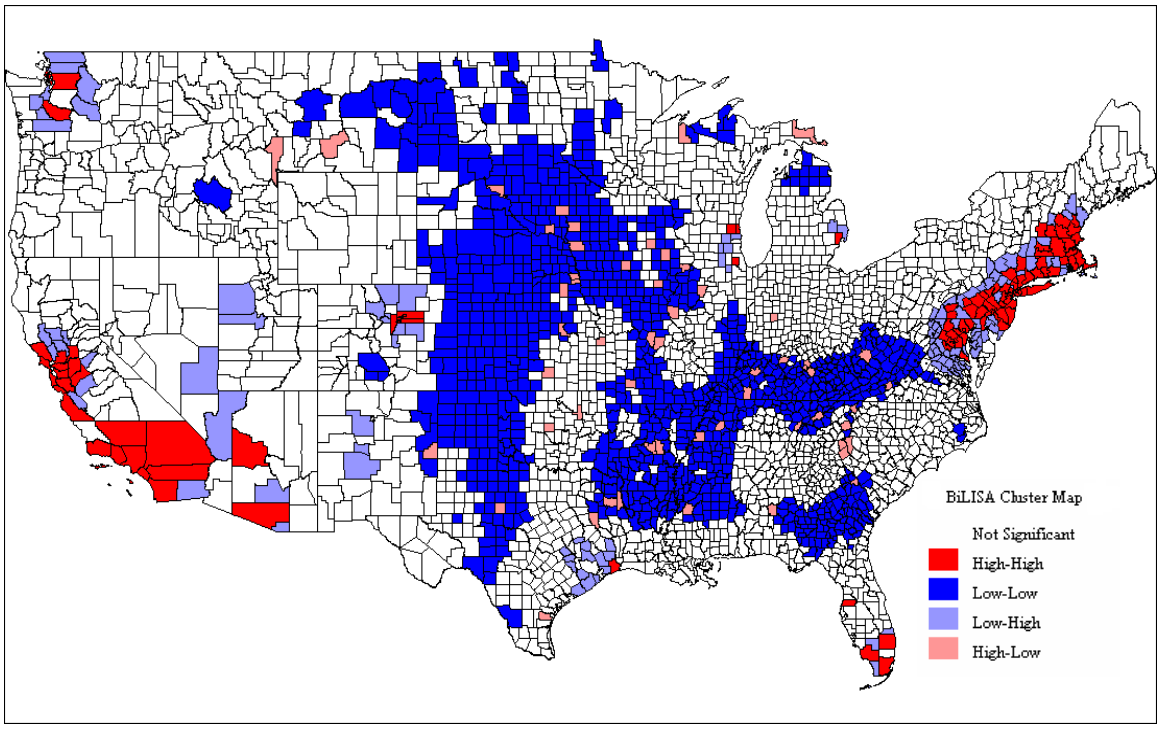


Figure 4.4. Bivariate LISA Cluster Map of Research Institutes and Spatial Lag of R&D and Testing Establishments, 2003.

4.5 Econometric Model

Most of the previous empirical studies on industry location have employed non-spatial econometric models, such as the Ordinary Least Squares (OLS), Poisson, Negative Binomial, and Tobit (for recent exception see Roe, Irwin, and Sharp,2002; Goetz and

Rupasingha, 2002; Isik, 2004; and Sambidi and Harrison, 2005). Since the data are collected with reference to points in space, employing OLS (and models mentioned above) as an econometric tool will produce spatially autocorrelated residuals, resulting in biased estimates and all inferences based on the model may be incorrect (Anselin, 1988; LeSage, 1999).

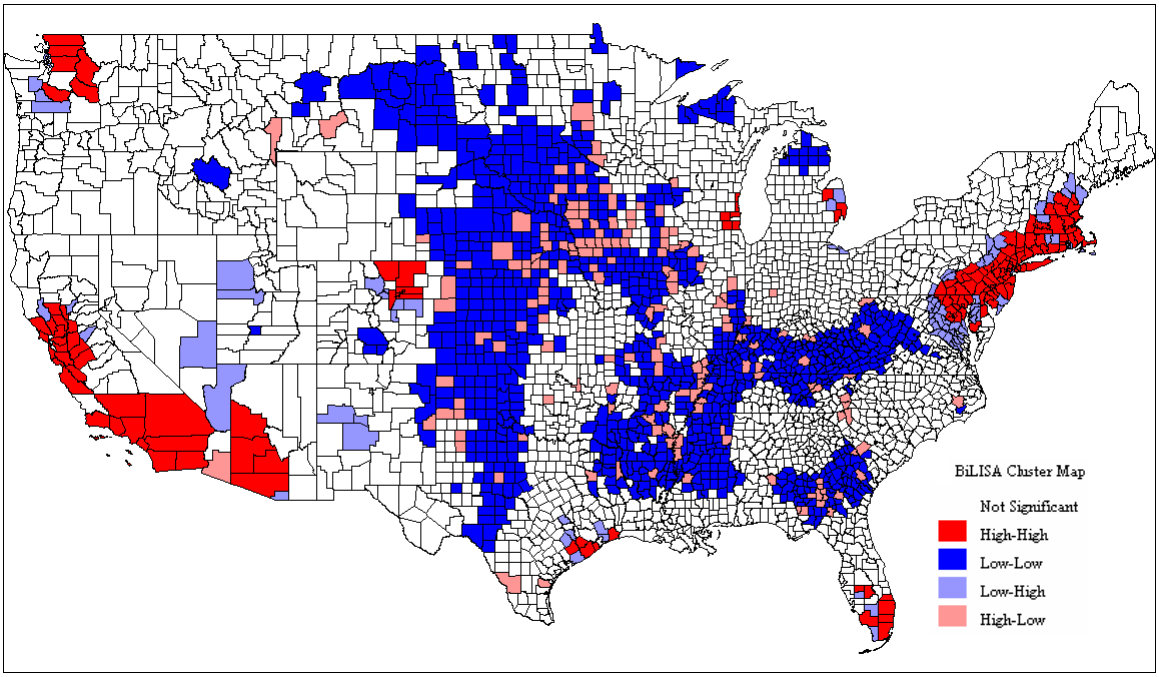


Figure 4.5. Bivariate LISA Cluster Map of Manufacturing Firms and Spatial Lag of R&D and Testing Establishments, 2003.

The spatial correlation of OLS residuals (without spatial component) is formally tested employing different spatial correlation indices (Morans-I, Wald, Lagrange multiplier and the Likelihood ratio test statistic), as suggested by LeSage (1999). All the tests indicated the presence of spatially correlated residuals in the regression model (Table 4.2). To overcome this problem of spatial autocorrelation, two different specifications are available. The two forms of specifications are: 1) a spatial autoregressive model (SAR), which is appropriate when spatial dependence exists in the

form of a spatially lagged dependent variable, and 2) a spatial error model (SEM), which is relevant when the spatial dependence operates through the disturbance term (Anselin 1988). The former can be expressed as:

$$y = \rho \mathbf{W} y + \mathbf{X}\beta + \mu$$

(4.5)

$$\mu \sim N(0, \sigma^2 I_n)$$

where y is a n x 1 vector of the dependent variable (the number of R&D and Testing laboratories), $\mathbf{W}y$ is a spatially lagged dependent variable for spatial weights matrix $\mathbf{W}$, ρ is the scalar for spatial lag coefficient, β is the k x 1 parameter vector, $\mathbf{X}$ is the n x k matrix of exogenous explanatory variables, $\boldsymbol{\mu}$ is an n x 1 vector of normally distributed error terms with zero mean and variance σ^2. The spatial lag $\mathbf{W}y$ can be considered as a spatially weighted average of the R&D and Testing facilities at neighboring counties. Ignoring a spatially lagged dependent variable yields inconsistent and biased estimates for the parameter coefficients in the model (Anselin et al. 2000).

Table 4.2. Tests for Spatial Correlation in Residuals of a Regression Model.

Test Statistics	Value
Moran's I Statistic	9.815***
Lagrange Multiplier	92.232***
Likelihood Ratio	81.521***
Wald	1072.053***

Note: *** indicates significance at the 1% level

The second form of spatial dependence, which is often expressed as a spatial autoregressive process for the disturbance term is expressed as:

$$y = \mathbf{X}\beta + \varepsilon$$

(4.6)

$$\varepsilon = \lambda \mathbf{W} \varepsilon + \mu$$

$$\mu \sim N(0, \sigma^2 I_n)$$

where λ is a spatial autoregressive coefficient and μ is is a standard spherical error term. Ignoring spatial dependence in the error term does not lead to biased estimates, but the estimate of their variance will be biased, leading to erroneous interpretations and wrong conclusions (Anselin et al. 2000).

If there was evidence indicating spatial dependence in both forms, spatial lag and error terms, a more general specification can be employed. This general spatial specification is called general spatial model (SAC), which accounts for both the spatially lagged term as well as a spatial error structure.

$$y = \rho \mathbf{W} y + \mathbf{X}\beta + \varepsilon$$

(4.7)
$$\varepsilon = \lambda \mathbf{W} \varepsilon + \mu$$

$$\mu \sim N(0, \sigma^2 I_n)$$

This specification is employed if there is evidence that spatial dependence existed in the error structure from a SAR estimation, which is tested by employing a LM-test. This study employs a spatial two-stage least squares (spatial 2SLS) estimator to examine the factors affecting the location of R&D and testing facilities in the U.S. We use the spatial 2SLS estimator in the estimation of this model, since it is robust to possible misspecifications and it appropriately accounts for the endogenity of the spatial lag term (Anselin 1988; Isik 2004). Moreover, there is a possibility that some of the other variables employed in the model may be influenced by R&D and testing laboratories. For example, the existence of a R&D and testing laboratory at a particular region may influence the location-decision of a biotech related manufacturing firm in that particular region. The spatial 2SLS also accounts for the possible endogenity of the explanatory variables in determining the location of R&D firms. We follow Kelejian and Prucha's

generalized spatial 2SLS procedure for estimating the spatial autoregressive model with autoregressive disturbances. The model is estimated employing the MATLAB routines provided by LeSage's econometric toolbox (2005).

4.6 Results

Results of the empirical estimation are presented in table 4.3. We follow Goetz et al. (2004), in employing the criteria outlined by LeSage (1999) to select the appropriate spatial specification for our data. Since the general spatial model (SAC) model accounts

Table 4.3. Estimates of Factors Affecting the Location of U.S. R&D and Testing Facilities.

Variable	Coefficient	t-stat
Constant	-14.2770***	-4.1780
Manufacturing firms	1.4832***	38.8986
Poverty	0.1528	1.3899
Population	1.32E-07	0.0350
Median Housing Value	3.60E-05***	3.0323
Unemployment	-0.1095	-0.6014
Median Household Income	-0.0001	-0.9406
Venture Capital	0.3961***	12.0509
Colleges	0.9629***	7.5752
Hospitals	-0.6684***	-3.4984
Average Wage Per job	0.0003***	5.0635
Education	0.2000***	3.7645
Property Tax	-0.0001***	-4.9191
Crime Index	-0.0002***	-4.2593
Farm Land	-1.54E-06**	-1.9380
Nonmetro-metro	1.0086	1.5784
lamda	0.4138***	14.2923

Note: *,**,*** Statistical significance at the 10, 5, and 1% levels, respectively.

for both the substantive and nuisance dependence (Anselin et al. 2000), we first estimate the SAC model. The estimated SAC model resulted in a negative value for the spatial agglomeration parameter (ρ), indicating that there are no spatial effects in the dependent variable. However, the SAC model resulted in positive and significant spatial autoregressive coefficient (λ), indicating that the spatial dependence exists in the error term. Therefore, we employ the SEM specification to estimate the spatial 2SLS model to

analyze the factors affecting the location-decision of R&D and testing firms. The spatial 2SLS model produced an R-square equal to 0.80 indicating a good-fit for these data. Since the spatial autoregressive coefficient (λ) is found to be positive and significant at the 1% level, this suggests that inference based on the standard 2SLS specification without a spatial correction, is not valid for the data under consideration. This result indicates the presence of spatial agglomeration economies for the spatial structure of the R&D industry. The positive sign indicates that the spatial distribution of R&D firms is positively correlated across counties.

Most of the regressors in the estimated model had expected signs and are found to be significantly different from zero. The number of manufacturing firms in a given county was found to be positive and significant at the 1% level, indicating R&D and manufacturing firms' preference to locate in close proximity of each other. This result is in accordance with Gray and Parker (1998) who argued that manufacturing firms that produce high-value intermediate active ingredient are located near by R&D firms, since the production requires significant input from high-skilled labor. Similarly, the number of research institutes in a given county was also found to be positive and significant at the 1% level, indicating the significant relationship between R&D firms and research institutes. This result is broadly consistent with earlier findings in Audretsch and Feldman (1996), Zucker, Darby, and Brewer (1998), and Furman et al. (2006) that knowledge spillovers from university research to innovative firms are at work in the biotech related R&D industry. Conversely, the number of hospitals in a given county had an unexpected negative sign and was found to be significant.

The median housing value had a positive and significant impact on the location of R&D firms, indicating that firms involved in research activities prefer to locate in

counties with well-developed infrastructure and high-standard of living. This result is in accordance with the current spatial distribution of R&D firms, which are mainly located in the urban areas that are associated with high cost-of-living. Similarly, a county's average hourly wage and education variable, which are considered to be proxies for availability of high-skilled labor, are found to be positive and significant at the 1% level. A county's crime rate was found to have negative impact on the location of R&D firms. This result indicates that within the metropolitan cities (where most R&D firms are located), R&D firms prefer to locate in areas with low crime rate.

The property tax variable was found to be negative and significant at the 1% level, indicating that R&D firms prefer to locate in counties with low property tax. Since the cost associated with the development of an innovative new drug is more than $800 million (DiMasi, Hansen, and Grabowski 2003; Congressional Budget Office 2006), R&D firms prefer to locate in regions that provide R&D tax credits and funding sources for their facilities. Similarly, availability of local venture capital was also found to be positive and significant, indicating R&D firms' dependency on local financial sources. This result is consistent with earlier findings of Powell, Koput, Bowie, and Smith-Doerr (2002) that venture capital firms are an important source of capital for new and small R&D firms.

The amount of farmland in a given county was found to be negative and significant at the 5% level, indicating R&D firms' preference to locate in urban regions with well-developed infrastructure. Generally, counties with large farmland are considered to be rural, less populated, with a poor infrastructure, compared to the urban counterpart (especially in terms of research universities, good transportation facilities, financial resources, specialized intermediate industries, and availability of skilled labor

pool). However, the urbanization economies variable (*Nonmetro-Metro*) was found to be insignificant in the location of R&D firms. The reason for this can be that this variable is correlated with other explanatory variables, which are picking up its effects.

4.7 Conclusion

The hypothesis of spatial agglomeration economies is confirmed for the spatial structure of the R&D industry, indicating that R&D firms are positively correlated across counties, resulting in clustering of research activities. Proximity to manufacturing firms and research universities, and availability of venture capital firms were found to have the most significant impact on the location of R&D firms. This indicates that the R&D firms prefer to locate in regions where they have a source for financing their business , access to research institutes to collaborate with skilled labor and obtain new technology, and access to manufacturing firms to produce intermediate active ingredient in large batches before it gets commercialized . The significance of both spatial autoregressive parameter and the *college* variable indicate that public as well as private spillovers are at work in the R&D industry, resulting in their spatial clustering.

R&D firms also prefer to locate in counties with a well developed infrastructure. This is indicated by the positive and significant estimates of median housing value, average wage per job, and education, and a negative and significant estimate of the farmland. The significant negative sign associated with property tax estimate indicates that in order to attract new R&D firms and develop the existing ones, the state and local economic development agencies should provide certain tax credits and business incentives.

These findings may hinder rural areas hopes of attracting R&D firms. However, rural areas that are adjacent to major metropolitan cities are still capable of attracting manufacturing firms that produce intermediate products and drugs that have achieved commercial scale. In order to attract these firms, the rural areas should develop their

infrastructure, especially, highway access, housing facilities, and business incentives. Thus, the state and local economic development agencies should design strategies based on the type of firm they want to attract.

4.8 References

Acs, Z, D. Audretsch and M. Feldman .1992. "Real Effects of Academic Research: A Comment," American Economic Review, 82: 363–367.

Acs, Z. J., Audretsch, D. B., Feldman, M. P. (1994) "R&D Spillovers and Innovative Activity", Managerial and Decision Economics, Vol. 15, No. 2 (Mar., 1994), pp. 131-138

Aharonson, S. B., Baum, J.A.C. and Feldman, M.P., 2004. Industrial Clustering and the Returns to Inventive Activity: Canadian Biotechnology Firms, 1991-2000, Danish research Unit of Industrial Dynamics (DRUID) Working Paper, 04-03.

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

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

Anselin, L., A. Varga and Z. Acs. 1997. "Local Geographic Spillovers Between University Research and High Technology Innovations," Journal of Urban Economics 42: 422–448.

Anselin, L., A. Varga and Z. Acs. 2000. "Geographical Spillovers and University Research: A Spatial Econometric Approach." Growth and Change 31: 501-515.

Anselin, L., I. Syabri, and O. Smirnov. 2002. "Visualizing Multivariate Spatial Correlation with Dynamically Linked Windows." In: Proceedings of the SCISS Specialist Meeting "New Tools for Spatial Data Analysis". Santa Barbara, California, USA. May 10-11, 2002.

Audretsch, D. B. and M. P. Feldman. (1996). "Knowledge Spillovers and The Geography Innovation and Production." American Economic Review, 86: 630-640.

Battelle Technology Partnership Practice and State Science and Technology Institute. 2004. " Laboratories of Innovation: State Bioscience Initiatives 2004." Biotechnology Industry Organization, Washington D.C.

Begemann BD. 1997. "Competitive strategies of biotechnology firms: Implication for US Agriculture." Journal of Agricultural and Applied Economics 29:117-22.

Brennan, M.F., Pray, C.E., & Courtmanche, A. 1999. “Impact of industry concentration on innovation in the U.S. plant biotech industry.” Paper presented at the Transitions in Agbiotech: Economics of Strategy and Policy conference, Washington, DC.

Biotechnology Industry Organization. 2005-2006. “Guide to Biotechnology.” Biotechnology Industry Facts. BIO, Washington D.C.

BioWorld Publishing Group. 2005. BioWorld Today.

California Trade and Commerce Agency, Office of Economic Research. 2001. Biotechnology. May 2001.

Cassiman, B. and R. Veugelers. 2002. R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium. American Economic Review September 2002: 1169-1184.

Cohen, W.M. and D.A Levinthal. 1989. Innovation and Learning: The Two Faces of R&D. Economic Journal. 99: 569-596.

Congressional Budget Office. 2006. Research and Development in the Pharmaceutical Industry. A CBO Study, Congress of the United States. October 2006.

Cortright, J., & Mayer, H. (2002). “Signs of Life: The Growth of Biotechnology Centers in the U.S.” Portland: The Brookings Institution Center on Urban and Metropolitan Policy.

Dahlander, L. and McKelvey, M. (2003): Revisting Frequency and Spatial Distribution:Innovation Collaboration for Biotech Firms. Paper presented at DRUID's Summer 2003 Conference, Helsingore, June 12-14.

Darby M.R. and L.G. Zucker, 1996. "Star Scientists, Institutions, and the Entry of Japanese Biotechnology Enterprises," NBER Working Papers 5795, National Bureau of Economic Research, Inc.

DiMasi, J. A., R. W. Hansen, and H. Grabowski. 2003. The price of innovation: new estimates of drug development costs. Journal Of Health Economics 22:151-185.

Feldman, M., 2003. “The locational dynamics of the US biotech industry: Knowledge externalities and the anchor hypothesis.” Industry & Innovation, 10(3), 311-328.

Feldman, M. P. 1994a. “Knowledge Complementarity and Innovation.” Small Business Economics, 6: 363-372.

Feldman, M. P. 1994b. The Geography of Innovation. Boston: Kluwer Academic Publishers, Second Printing. (1999).

Furman, Jeffrey L. and Cockburn, Iain M., "Public & Private Spillovers, Location and the Productivity of Pharmaceutical Research" (September 2006). NBER Working Paper No. W12509

Furman, J.L., M.K. Kyle, I. Cockburn, and R. Henderson. 2004. "Public and Private Spillovers, Location, and the Productivity of Pharmaceutical Research." Mimeo. Duke University.

Geospatial and Statistical Data Center. 'Crimes reported 1994-2002'. University of Virginia Library. http://fisher.lib.virginia.edu/collections/stats/crime/crimes94.html

Goetz, S. J. and A. Rupasingha. 2002. "High-Tech Industry Clustering: Implications for Rural Areas." American Journal of Agricultural Economics 84(5).

Goetz, Stephan J., Anil Rupasingha and Julie Zimmerman, "Spatial Food Stamp Program Participation Dynamics in US Counties," Review of Regional Studies, 34, 3 (2004):53-71.

Graff D. G. 1997. "The Agricultural Biotechnology Industry in Overview." Agricultural and Resource Economics, University of California, Berkeley.

Gray M. and Parker E. 1998. "Industrial Change and Regional Development: The Case of the US Biotechnology and Pharmaceutical Industries." ESRC Centre for Business Research, University of Cambridge, Working Paper No. 95.

Grossman, G., and E. Helpman. 1991. Innovation and Growth in the Global Economy. Cambridge, MA: MIT Press.

Henderson, J. Vernon, Kuncoro, Ari, and Turner, Matt. "Industrial Development in Cities." Journal of Political Economy, October 1995, 103(5), pp. 1067-90.

Isik, M.2004. "Environmental Regulation and the Spatial Structure of the U.S. Dairy Sector." American Journal of Agricultural Economics 86: 949-962.

Jacobs, J. (1969), The economy of cities. Random House: New York.

Jaffe, A., M. Trajtenberg, and R. Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. The Quarterly Journal of Economics 108 (3):577-598.

Kalaitzandonakes, N. and Bjornson, B. (1997). "Vertical and Horizontal Coordination in the Agro-biotechnology Industry: Evidence and Implications" Journal of Agricultural and Applied Economics 29: 129-139.

Kelejian, H. and I. Prucha (1998). A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. Journal of Real EstateFinance and Economics 17, 99–121.

Krugman P. 1991. Geography and Trade. MIT Press, Cambridge, MA

Kyle M. 2004. Does Locale Affect R&D Productivity? The Case of Pharmaceuticals. FRBSF Economic Letter. Number 2004-32, November.

Lee, K.B. Jr., and Burrill, S.G. 1995. Biotech 96: Pursuing Sustainability. The Tenth Industry Annual Report. Ernst & Young LLP, London.

LeSage, J.P. 1999. "The Theory and Practice of Spatial Econometrics." Unpublished, Dept. of Econ., University of Toledo.

LeSage, J. P. (2005), "Econometrics Toolbox", MatLab programs. Downloaded from http://www.spatial-econometrics.com.

Marshall, A. 1920. Principles of Economics. Eighth edition. London: Macmillan.

McCandless M.E. 2005. "Biotechnology Locations: Hitting the Mark." Business Facilities. http://www.businessfacilities.com/bf_05_05_cover.asp (Date last visited 14rth May 2006).

National Science Foundation. 2006. Industrial Funding of Academic R&D Continues to Decline in FY 2004. NSF 06-315.

van Oort G. F. 2004. "Urban Growth and Innovation: Spatially Bounded Externalities in the Netherlands." Published by Ashgate Publishing Company, USA.

Ottaviano, G. and J.-F. Thisse. New Economic Geography: what about the N? Environment and Planning A, forthcoming.

Pacheco, Andrada I., and Timothy J. Tyrrell. 2002. "Testing Spatial Patterns and Growth Spillover Effects in Clusters of Cities." Journal of Geographical Systems 4:275-285.

Paci, Rosario and Stefano Usai (1999), 'Externalities, Knowledge Spillovers and the Spatial Distribution of Innovation', GeoJournal, 49(4): 381–390.

Pisano, G. P. (1997). The Development Factory: Unlocking the Potential of Process Innovation. Boston, MA, Harvard Business School Press.

Pisano, G. P., W. Shan, and D. J. Teece. 1988. Joint ventures and collaboration in the biotechnology industry. In International Collaborative Ventures in US Manufacturing, edited by D. C. Mowery. Cambridge: Ballinger.

Prevezer, M. 1998. Clustering in biotechnology in the USA, pp. 124–193 in Swan, G. M. P., Prevezer, M. and Stout, D. K. (eds), The Dynamics of Industrial Clustering. International Comparisons in Computing and Biotechnology, Oxford, Oxford University Press.

Powell, W. W., and P. Brantley. 1992. Competitive cooperation in biotechnology:Learning through networks? In Networks and Organizations. Structure, Form andAction, edited by R. Eccles. Boston: Harvard Business School Press.

Powell W. W., K.W. Koput, J. Bowie,. and L. Smith-Doerr. 2002. "The Spatial Clustering of Science and Capital. Journal of Regional Studies, 36 (3).

Roe, B., E. Irwin, and J. S. Sharp. 2002. "Pigs in Space: Modeling the Spatial Structure of Hog Production in Traditional and Nontraditional Production Regions." American Journal of Agricultural Economics 84: 259-278.

Romer, P. 1986. Increasing Returns and Long-Run Growth. Journal of Political Economy. 94(5): 1002-1037.

Romer, P. 1990. "Endogenous Technological Change," Journal of Political Economy, 98: S71–S102.

Sambidi P.R. and W. Harrison. 2005. "Spatial Dependency of the Geographically Concentrated U.S. Broiler Industry." Paper Presented at the American Agricultural Economics Association Annual Meeting, Providence, Rhode Island.

Sporleder, T.L., L.E. Moss, and L.A. Nickles. 2002. "Agricultural Biotechnology Start-Ups: Does Venture Capital Matter?" Prepared for presentation at the Third Annual Agricultural Biotechnology Conference, Ohio's Future in Functional Foods, Reynoldsburg, OH.

Sporleder, T.L. and L.E. Moss. 2004. "Knowledge Capital, Intangible Assets, and Leverage: Evidence from U.S. Agricultural Biotechnology Firms." Paper Presented at the 2004 Symposium of the International Food and Agribusiness Management Association Meetings, Switzerland, June 12-13, 2004.

Toole, A. 2003: Understanding entrepreneurship in U.S. biotechnology: what are its characteristics, facilitating factors, and policy challenges?, Chapter 9, in D. Hart (ed.), The Emergence of Entrepreneurship Policy: Government, Start-ups, and Growth in the Knowledge Economy. Cambridge, UK: Cambridge University Press.

U.S. Census Bureau. 2003. County Business Pattern. Data Available at http://censtats.census.gov/cbpnaic/cbpnaic.shtml

U.S. Department of Agriculture. 2003. Economic Research Service. Data Available at http://www.ers.usda.gov/Data/

___. National Agricultural Statistics Service. Historical Data. Available at http://www.nass.usda.gov/Data_and_Statistics/index.asp

U.S. Department of Labor. Bureau of Labor Statistics. Available at http://www.bls.gov/home.htm

von Hippel, Eric (1994) "Sticky Information" and the Locus of Problem Solving: Implications for Innovation" Management Science 40, no.4 (April): 429-439

Xia, Y., and Buccola, S. 2005. University life science programs and agricultural biotechnology. American Journal of Agricultural Economics, 87(1), 229-243.

Yang H. and S. Buccola. 2003. "University-Industry Relationships and the Design of Biotechnology Research." Paper Presented at the American Agricultural Economics Association Annual Meeting, Montreal, Canada, July 2003.

Zeller, C., 2001, Clustering biotech: A recipe for success? Spatial patterns of growth of biotechnology in Munich, Rhineland and Hamburg, Small Business Economics, 17, 123-141.

Zucker, L. G., M. R. Darby, and M. B. Brewer. 1998. Intellectual human capital and the birth of U.S. biotechnology enterprises. The American Economic Review 88:290-306.

Zucker L. G. & M.R. Darby & J.S. Armstrong, 2003. "Commercializing knowledge: university science, knowledge capture and firm performance in biotechnology," Proceedings, Federal Reserve Bank of Dallas, issue Sep, pages 149-170.

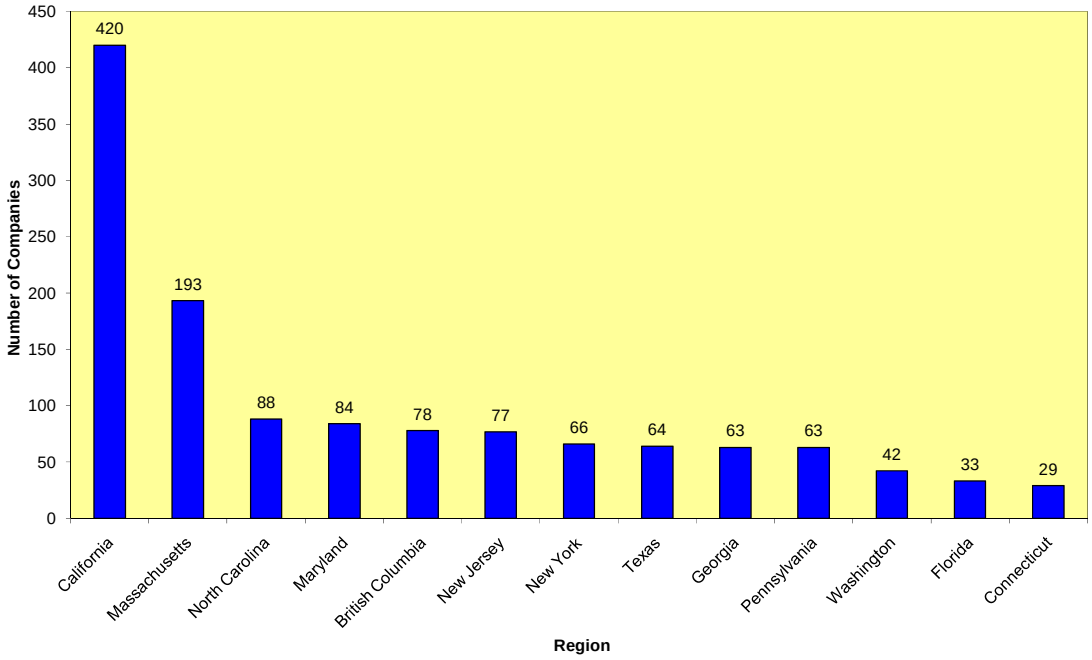
CHAPTER 5

A CRITICAL ANALYSIS OF THE GEOGRAPHICAL DISTRIBUTION OF BIOTECH RELATED MANUFACTURING AND R&D FACILITIES IN THE UNITED STATES

5.1 Background

The biotech industry is one of the fastest growing industries in the U.S., increasing sales from $7.7 billion in 1994 to $33.3 billion in 2004 (Ernst and Young 2005). According to Biotechnology Industry Organization (BIO), in 2003 there were 1,473 biotechnology companies in the U.S., employing 198,300 people and spending $17.9 billion on research and development. The success of biotech industry is also indicated by the fact that in 2004, 40 states have adopted strategies to stimulate the growth of biotechnology and 50 states have technology based economic development initiatives that are available for biotech companies (Battelle and State Science and Technology Institute (SSTI) 2004). A survey administered by Grudkova (2001) on 77 local and 36 state economic development agencies indicated that 83 percent of the respondents considered biotechnology as one of their top two targets for industrial development (Cortright and Mayer 2002). The top five states in terms of number of biotech companies are: California (420), Massachusetts (193), North Carolina (88), Maryland (84), and New Jersey (77) (figure 5.1; Ernst and Young, 2004). The biotech industry is mainly concentrated in nine cities/regions (San Francisco Bay Area, Boston/Cambridge, San Diego, Los Angeles, New York, Philadelphia, Raleigh-Durham, Seattle and Washington, DC), which account for three-fourths of the nation's largest biotechnology firms and for three-fourths of the biotech firms formed in the past decade (Cortright and Mayer, 2002).

Biotech firms are mainly research and development (R&D) oriented and operate in collaboration with research-oriented universities, biomedical research centers, and



Source: Ernst & Young LLP, America's Biotechnology Report: Resurgence, 2004.

Figure 5.1 U.S. Biotech Companies by State and Province.

other diversified companies that aid in production and distribution of biotech products. The biotech products are related to drugs and pharmaceuticals, agriculture, and environment, which aid in improving the quality of health, increasing agricultural production, improving the food quality, minimizing environmental hazards and providing a cleaner environment. Since different subsectors are involved in the production of biotech products, the biotech industry in U.S. doesn't have a separate NAICS or SIC code. However, Battelle and SSTI classified the bioscience into five major subsectors as follows: 1. Agricultural Feedstock and Chemicals (AF&C; NAICS: 311221, 311222, 311223, 325193, 325199, 325221, 325222, 325311, 325312, 325314, 325320, 424910), 2. Drugs and Pharmaceuticals (D&P; NAICS: 325411, 325412, 325413, 325414), 3.

Medicinal Devices and Equipment (MD&E; NAICS: 339111, 339112, 339113, 339114, 334510, 334516, 334517), 4. Research and Testing (R&T; NAICS: 541380, 541710), and 5. Academic Health Centers, Research Hospitals, and Research Institutes (Battelle and SSTI 2004). In accordance with Battelle and SSTI (2004), the last subsector is not included in the study because it cannot be separated from the overall hospital sector. The first four subsectors include twenty five industries that are involved in biotechnology activities, with total employment of 885,368 jobs across 17,207 establishments (Battelle and SSTI 2004). Examples of companies and products produced by each subsector are presented in table 1. For example, some of the products produced by the AF&C subsector are: 1. genetically modified high-yield and disease-resistant varieties, 2. ethanol and biodiesel fuels, 3. biodegradable materials synthesized from plant-based feedstock (Battelle and SSTI 2004).

Since biotech activities are performed by several industries belonging to different sectors/ subsectors, it is important to analyze location factors specific to a given subsector. For example, industries belonging to the R&T subsector may wish to locate in close proximity to research institutes, whereas industries belonging to the AF&C subsector may wish to locate in areas with access to farmland. Moreover, variation exists in the geographical distribution of the four subsectors. For example, the AF&C subsector is found to be most dispersed subsector with 21 states having a specialization[1], whereas only 8 states have specialization in the D&P subsector (Battelle and SSTI 2004). Several studies have empirically examined the location aspects of the U.S. biotech industry; however, none of them have simultaneously analyzed location factors specific to a given subsector.

[1]Battelle and SSTI (2004) define regional specialization as regions with Location Quotients (LQ) of 1.2 or greater. LQ measures the level of employment concentration for a given subsector within a state relative to the country (Battelle and SSTI, 2004)

Table 5.1. Industries in the Bioscience Subsectors

Categories	NAICS Codes	Examples of Products	Examples of Companies
Agricultural Feedstock and Chemicals			
Wet Corn Milling	311221	i) Nutritionally enhanced, genetically engineered, insect-resistant crops ii) Ethanol and biodiesel fuels iii) Biodegradable materials synthesized from plant-based feedstock	BASF, Bayer Crop Science, Cargill-Dow LLC, Dow Agrosciences, Genencor International (Bioproducts Division), Monsanto, and Pioneer Hi-Bred International.
Soybean Processing	311222		
Other Oilseed Processing	311223		
Ethyl Alcohol Manufacturing	325193		
All Other Basic Organic Chemical Manufacturing	325199		
Cellulosic Organic Fiber Manufacturing	325221		
Noncellulosic Organic Fiber Manufacturing	325222		
Nitrogenous Fertilizer Manufacturing	325311		
Phosphatic Fertilizer Manufacturing	325312		
Fertilizer (Mixing Only) Manufacturing	325314		
Pesticide and Other Agricultural Chemical Manufacturing	325320		
Farm Supplies Merchant Wholesalers (only seeds)	424910		
Drugs and Pharmaceuticals			
Medicinal and Botanical Manufacturing	325411	Vaccines, Cancer Treatments, Herbal Supplements and Vitamins, Tissue and Cell Culture Media, and Delivery Platforms	Aventis, Biogen IDEC, Genentech, Novartis Animal Health Products, Ortho Clinical Diagnostics, Pfizer, and Roche Centralized Diagnostics.
Pharmaceutical Preparation Manufacturing	325412		
In-Vitro Diagnostic Substance Manufacturing	325413		
Biological Producing (except Diagnostic) Manufacturing	325414		
Medical Devices and Equipment			
Electromedical Apparatus Manufacturing	334510	Minimally invasive surgical equipment Systems manufactured from biomaterials Therapeutic implantable devices	GE health care, Medtronic, Smith and Nephew, and W.L. Gore Medical
Analytical Laboratory Instrument Manufacturing	334516		
Irradiation Apparatus Manufacturing	334517		
Laboratory Apparatus and Furniture Manufacturing	339111		
Surgical and Medical Instrument Manufacturing	339112		
Surgical Appliance and Supplies Manufacturing	339113		
Dental Equipment and Supplies Manufacturing	339114		
Research and Testing			
Testing Laboratories	541380	Preclinical drug therapeutics Human Growth Hormones, Monoclonal antibodies, protein receptors, drug discovery techniques, and drug delivery technology	Applied Molecular Evolution, Inc, Charles River Laboratories, Isis Pharmaceuticals, Inc., Sirna Therapeutics.
R&D in Physical, Engineering and Life Sciences	541710		

Source: Battelle and SSTI, 2004.

The primary objective of this paper is to identify county level determinants of the geographical distribution of the U.S. biotech industry. Specifically, this study analyzes the numerous firm-specific, location-specific, and agglomeration factors that affect the location, movement and concentration of biotech firms, based on their biotech activities. The study utilizes a seemingly unrelated regression model that captures the likelihood of contemporaneous correlation between the disturbance terms across equations. Analyzing these factors will aid state and local economic development agencies in designing specific strategies to attract biotech firms, based on the functions they perform.

5.2 Literature Review

The Biotech industry is defined as "the application of biological knowledge and techniques to develop products and services" (BIO, 2005). Goetz and Morgan (1995) defined it as "any technique that uses living organisms to make/modify products, improve plants or animals, or develop microorganisms for a specific use." The most critical application of biotechnology is seen in the field of medicine, where it is making a continuous progress in improving human health and quality of life. One of the important applications of biotechnology in medicine is gene therapy, which offers "the potential to supply a patient's cells with healthy copies of missing or defective genes to prevent cancer and other diseases" (Ernst and Young 2000).

Another critical area of biotechnology is agriculture, where it is mainly focused on producing genetically modified crop varieties, which are considered to be high-yielding, disease and environmental stress resistant, with better quality and longer shelf life. These genetically modified crop varieties are considered to be environmental friendly as they reduce the need for pesticides and herbicides. According to the National Center for Food and Agricultural Policy's (NCFAP) report (Sankula and Blumenthal

2004), the U.S. farmers who adopted the 11 biotech crop varieties increased crop yields by 5.3 billion pounds, lowered pesticide use by 46.4 million pounds, and saved $1.5 billion by lowering production costs, which resulted in a net economic impact or savings of $1.9 billion (Biotechnology Industry Organization 2005). The U.S. is considered to be the global leader in the production of biotech crops, accounting for 58.8% (117.6 million acres) of the global area (200 million acres) of biotech crops in 2004 (Council for Biotechnology Information 2005). Biotechnology is also widely employed in the livestock industry, where it provides tools for improving animal health and increase livestock and poultry productivity (BIO 2005). In the case of food processing industry, biotechnology is mainly employed in enzyme production to enhance nutritional value of food products.

Similar to the differences in functions performed, the four biotech subsectors also indicate some differences in their geographical distribution. Stewart (2004) utilized the information provided by Battelle and SSTI (2004) and ranked the U.S. states for each of the four subsectors based on the LQ (see footnote 1). The top five states for each of the four subsectors are as follows: Iowa, South Carolina, Tennessee, West Virginia, and Idaho (AF&C); New Jersey, Delaware, Indiana, Connecticut, and North Carolina (D&P); Utah, Minnesota, Massachusetts, Delaware, and Indiana (MD&E); and, Delaware, Massachusetts, Idaho, Maryland, and California (R&T). Delaware was found to be the only state with specialization in all subsectors. Figures 5.2-5.5 presents the top twenty U.S. counties with respect to the number of establishments in each of the four subsectors. Only five counties (Cook IL, Los Angeles CA, Maricopa AZ, Orange CA, and San Diego CA) are represented among the top twenty counties list for all subsectors. Figures 5.6-5.9 presents counties with number of establishments greater than the national average

for each of the four subsectors, respectively. As illustrated in the figures, the AF&C subsector is widely distributed (861 counties), compared to the D&P (542 counties), MD&E (545 counties), and R&T (408 counties) subsectors. All subsectors, except the AF&C, are found to be concentrated in the Northeast, California, and other major metropolitan cities.

Several studies have empirically examined the location aspects of the U.S. biotech industry (Goetz and Morgan, 1995; Darby and Zucker, 1996; Gray and Parker, 1998; Prevezer, 1998; Lerner and Merges, 1998;; Zucker, Darby, and Brewer, 1998; Brennan, Pray and Courtmanche, 1999; Zucker, Darby, and Armstrong, 2003; Xia and Buccola, 2005).The rationale for concentration of the U.S. biotech industry in California and the Northeast has been attributed to proximity to highly research-oriented universities, research parks and laboratories, and well-developed infrastructure. Gray and Parker (1998) examined the theoretical arguments surrounding the location and organization of biotech firms and analyzed the prospects for industrial renewal and regional transformation. The authors segregate the U.S. biotech industry into three different categories/regions based on the functions performed by biotech firms in those regions.

The first category includes mature drug producing regions, such as New York, New Jersey, Pennsylvania, Delaware, Illinois, and Indiana. These regions include mature pharmaceutical firms that were established prior to the commercialization of biotechnology (before 1970s), and are now primarily involved in the manufacturing (53 %) and marketing (72%) of new drugs. Another category includes emerging drug-producing regions, such as San Francisco, San Diego, Los Angeles, Seattle, and Boston. Firms in these regions were established mainly during and after the commercialization of biotechnology, and are primarily involved in R&D activities (82%) for new drugs. A

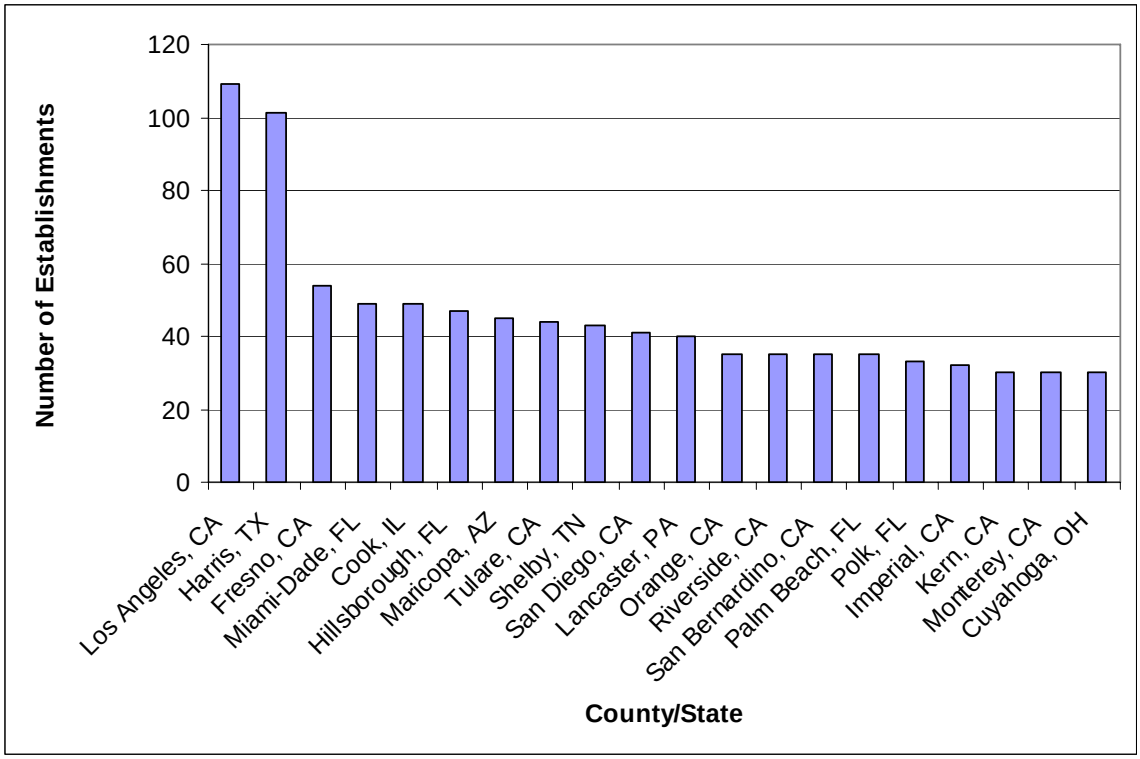


Figure 5.2. Top 20 U.S. Counties with Agricultural Feedstock and Chemical Establishments in 2003.

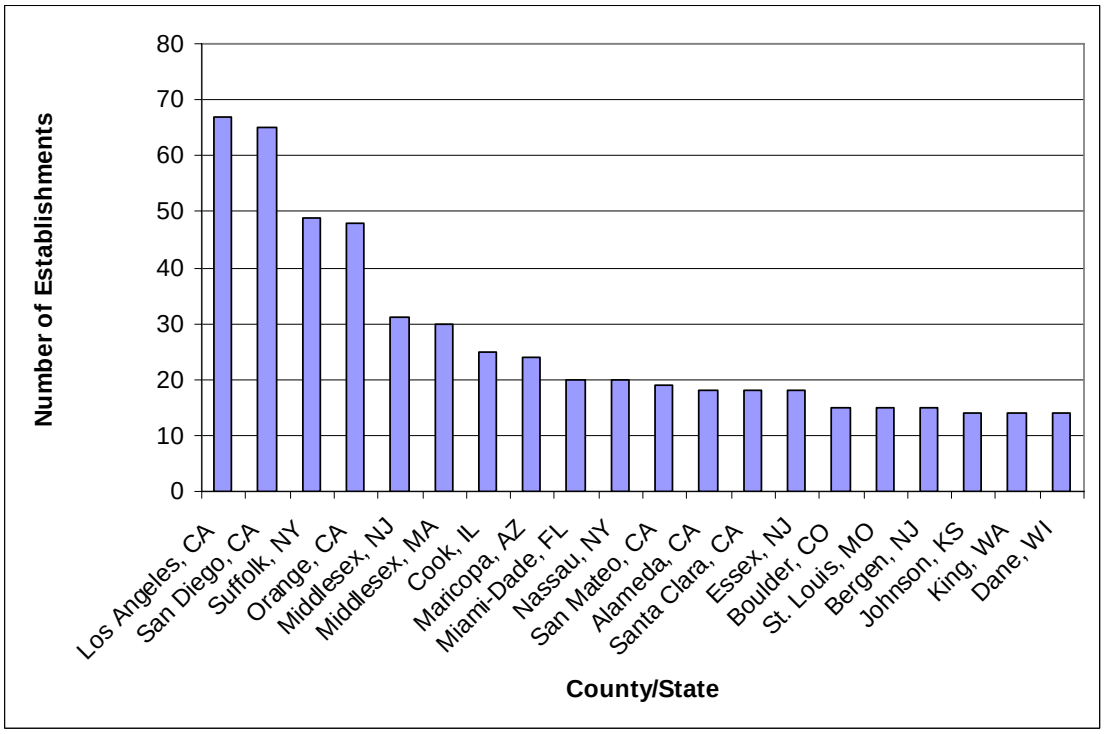


Figure 5.3. Top 20 U.S. Counties with Drug and Pharmaceutical Establishments in 2003.

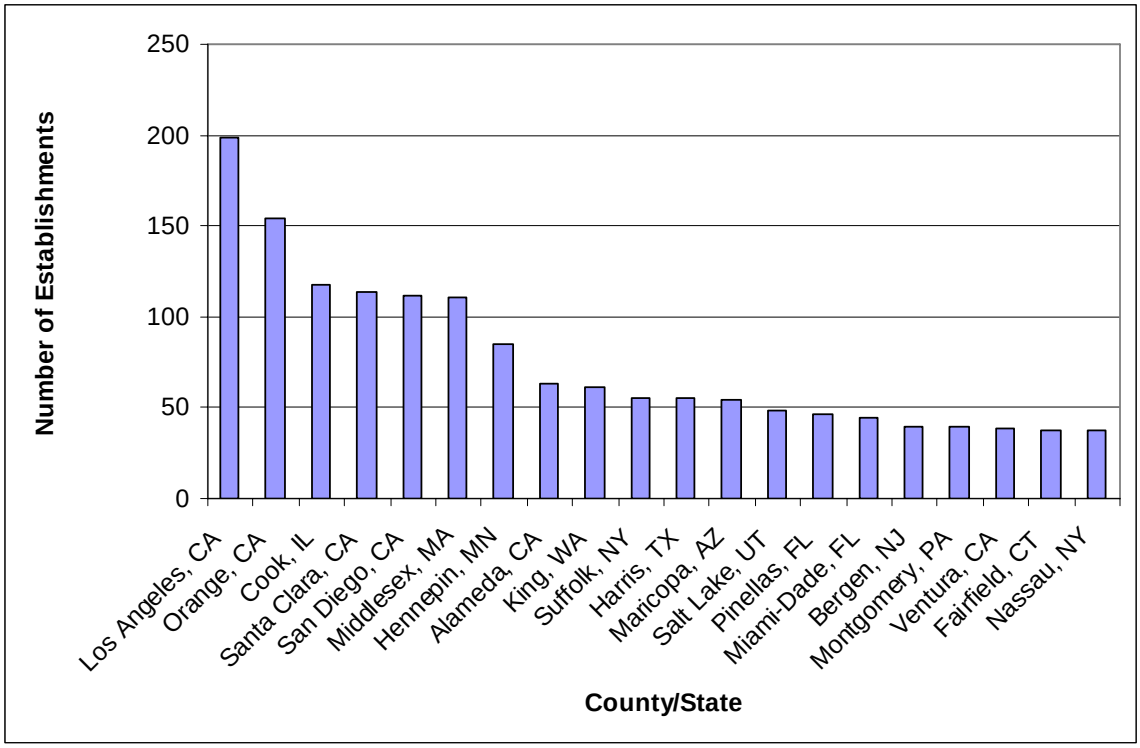


Figure 5. 4. Top 20 U.S. Counties with Medical Devices and Equipment Establishments in 2003.

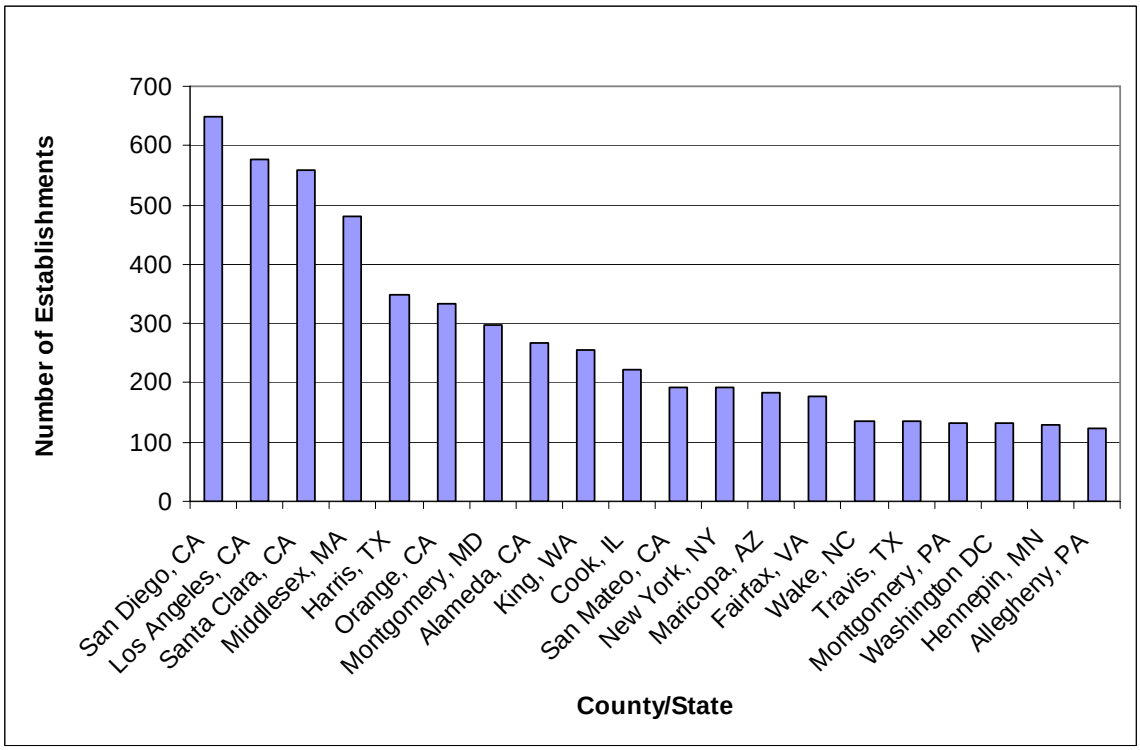


Figure 5.5. Top 20 U.S. Counties with R&D and Testing Establishments in 2003.

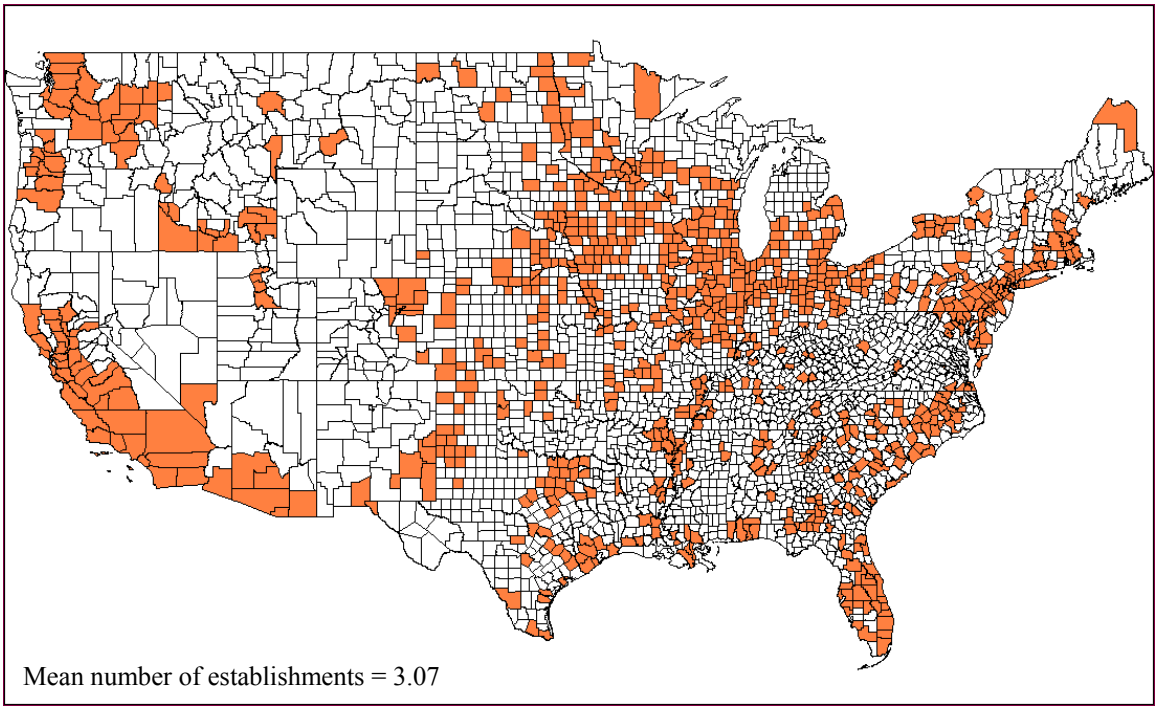


Figure 5.6. Geographical Distribution of Agricultural Feedstock and Chemicals Subsector.

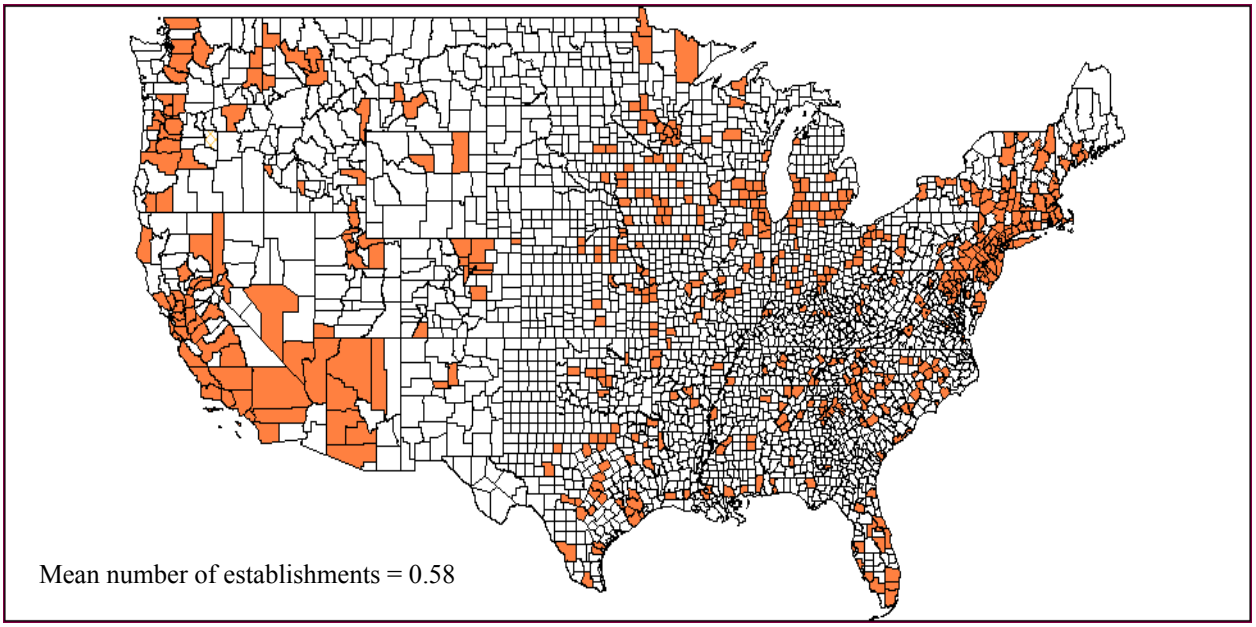


Figure 5.7. Geographical Distribution of Drugs and Pharmaceuticals Subsector.

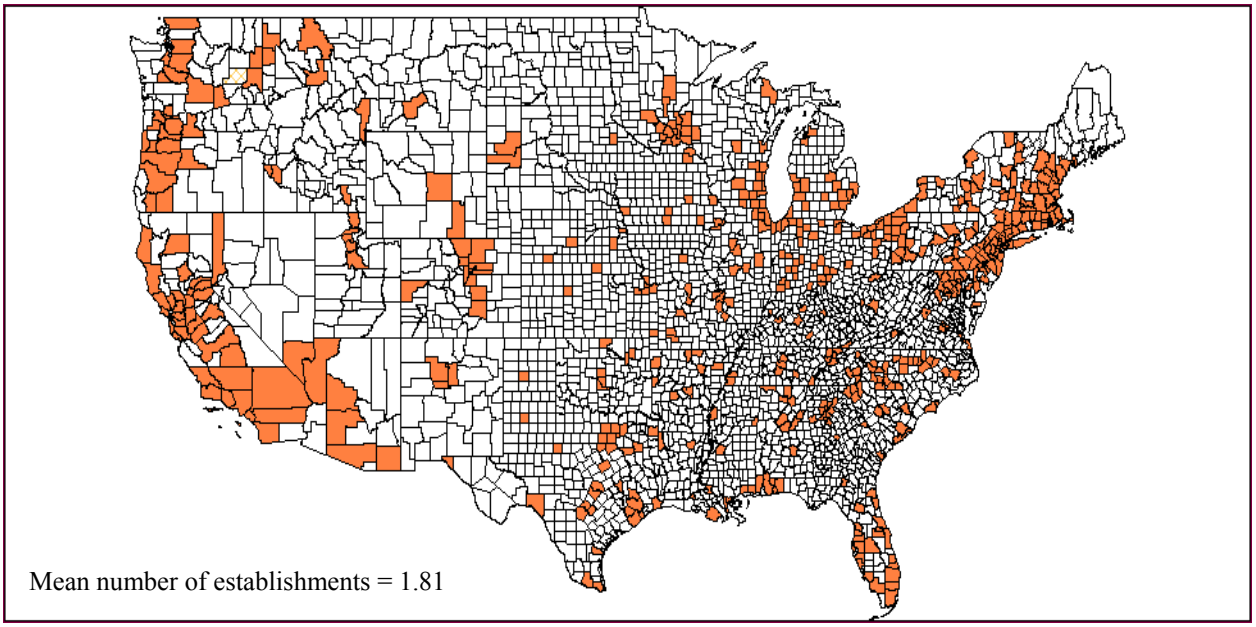


Figure 5.8. Geographical Distribution of Medical Devices and Equipment Subsector.

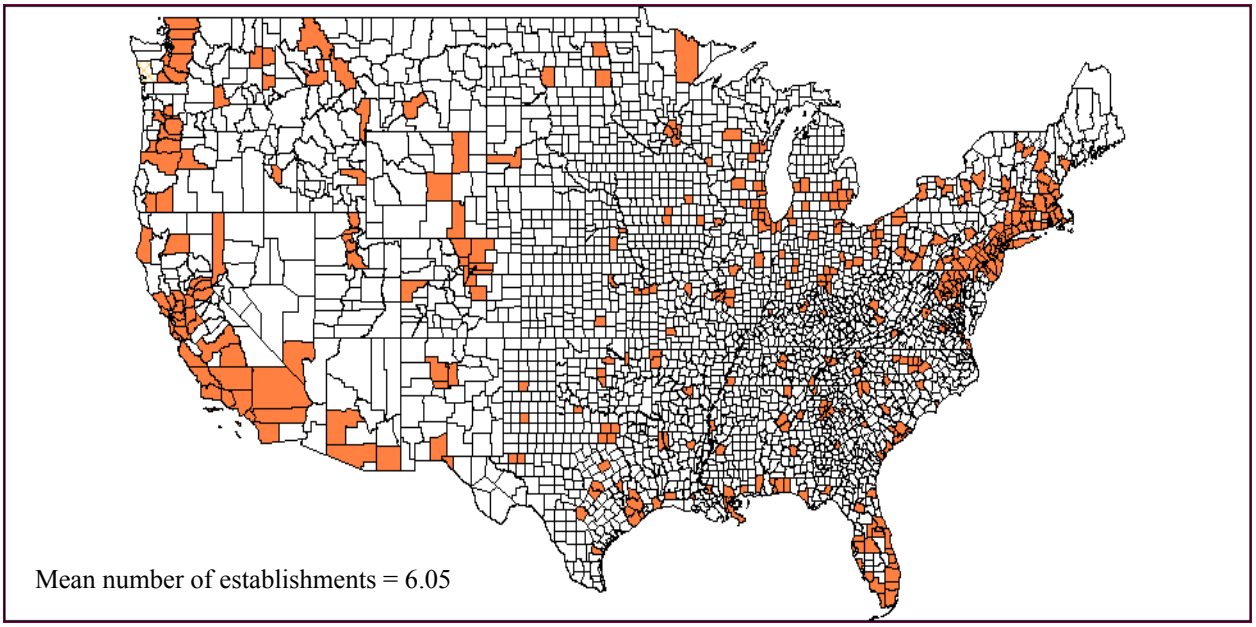


Figure 5.9. Geographical Distribution of Research and Testing Subsector.

third category includes low-cost periphery regions, such as Puerto Rico, the Southern states of the U.S., and other scattered isolated rural areas. Biotech firms in these regions

undertake the production of drug products that have achieved commercial scale and other intermediate products (Gray and Parker, 1998).

Munroe, Craft, and Hutton (2002) conducted a survey of biotech companies in three California counties (Alameda, Contra Costa, and Solano). The results indicated that proximity to leading research centers (i.e. ready supply of skilled labor, access to ongoing research activities, new technology, etc.) as primary determinants for their current location. Access to venture capital, a well-trained workforce, space for expansion and access to new technology were considered to be the most critical requirements for the growth and prosperity of respondents' business. The results also indicated that state and local economic development agencies that provide financial assistance packages (i.e. subsidies, tax advantages, loan guarantees, etc.), biotechnology incubators/research parks (with appropriate zoning, infrastructure and public transportation), promote public awareness and training programs for the workforce. The results also indicated some of the respondent's willingness to locate in regions associated with lower costs (housing, space, wages, etc.), less congestion/ commuting, and good incentives such as subsidies and tax credits (Munroe, Craft, and Hutton, 2002).

Goetz and Rupasingha (2002) analyzed the site-specific determinants of U.S. high-tech industry, which includes firms that are involved in biotech activities, such as drug and pharmaceutical manufacturing firms and R&D services. Their results indicated that the availability of an existing high-tech firm, number of college graduates, local property taxes, population (urbanization economies), total county income, highway access, and county amenity scale have a positive and significant impact on the location of high-tech firms. Conversely, a county's unemployment and unionization rate, per capita pollution, and the percentage of black population were found to have a negative and significant impact on the high-tech firm's location. The present article differs from

previous literature in that we examine the differences in factors affecting the location of four biotech subsectors simultaneously.

5.3 Data

The present study analyzed several categories of variables that are considered to affect the location of four biotech subsectors, such as agglomeration factors, infrastructure factors, and local economic and socioeconomic factors. County-level data were obtained from the 2003 county business patterns (U.S. Census Bureau), Economic Research Service, National Agricultural Statistics Service, and U.S. Dept. of Labor. The study involve four dependent variables: 1. number of establishments belonging to the AF&C subsector, 2. number of establishments belonging to the D&P subsector, 3. number of establishments belonging to the MD&E subsector, and 4. number of establishments belonging to the R&T subsector.

A factor that is considered to be a prerequisite for attracting a biotech firm is proximity to research institutions. Several studies have analyzed the role of research institutes in the development and commercialization of biotechnology (Powell and Brantley, 1992; Darby and Zucker, 1996; Zucker, Darby, and Brewer, 1998; Prevezer, 1998; Zucker, Darby, and Armstrong, 2003; Dahlander and McKelvey, 2003; Xia and Buccola, 2005). Industry funded university research increased from $630 million in 1985 to $2.1 billion in 2004 (National Science Foundation, 2006), indicating an increasing affiliation between university and industry in technology advancement. Some of the primary reasons for this collaboration are: access to complementary research activity and human capital, increasing commercial opportunities, stringency of patent law and federal policies, and the relative decline of public research funding (Santoro and Alok, 1999; Yang and Buccola, 2003). This study includes county-level number of colleges, universities, and

professional schools (*colleges*) as a proxy for the proximity to research institutions and assumes it will have a positive effect on the location of four biotech subsectors. County-level data for this variable was collected from county business patterns (U.S. Census Bureau, 1998 and 2003) with 611310 NAICS classification, which includes number of colleges, universities, and professional schools. Ongoing research intensity in life sciences at research institutions of a particular state is also considered to be a critical factor in the location decision of a biotech firm (Zucker, Darby, and Brewer, 1998; Munroe, Craft, and Hutton, 2002; Xia and Buccola, 2005). State level university life science R&D expenditures (*R&D*) is included as a proxy for ongoing research intensity. This variable is assumed to have a positive effect on the site-selection of all biotech subsectors. State level data for R&D expenditures was collected from National Science Foundation.

Businesses that provide venture capital are considered to be an important source of capital, especially, for new and small firms (Powell, Koput, Bowie, and Smith-Doerr, 2002). For a small biotech firm, availability of venture capital in a particular region is as important as the strong research capacity of that region. During 2004, venture capital accounted for approximately 23.5% of the total financing for the biotech industry. (BIO, 2005; BioWorld, 2005). Most of the biotech firms are small and operate at a loss, spending large amounts of money on research and development for several years before earning a profit (Cortright and Mayer, 2002). For example, only 1 in 5,000 potential new medicines reach the pharmacy shelf, and that is after 12 to 15 years of R&D with an average expenditure of $500 million (California Trade and Commerce Agency, 2001). As a result, most of the small biotech firms depend on venture capital funds, on research contracts and equity investment from large biotech firms, and on sales of their company

stock in public markets (Cortright and Mayer, 2002). Therefore, the availability of local venture capital firms (*Venture Capital*) is hypothesized to have a positive effect on the location of all four biotech subsectors (data for county-level number of venture capital establishments was collected from county business patterns with 523910 NAICS classification).

Agriculture is an important component of the biotech industry. Some studies have analyzed issues related exclusively to the spatial distribution of agricultural biotech firm location and its relationship with research institutions (Kalaitzandonakes and Bjornson, 1997; Graff, 1997; Begemann 1997; Brennan, Pray, and Courtmanche, 1999; Sporleder, Moss, and Nickles, 2002; Yang and Buccola, 2003; Sporleder and Moss, 2004; Xia and Buccola, 2005). Around 13 percent of firms in biotechnology are primarily involved in agriculture (Dibner 1995; Graff 1997). According to Ernst and Young (2000), in 1999, agricultural biotech firms employed 21,900 workers, generated $2.3 billion in revenues and $1.4 billion in personal income for employees and owners. Since biotech firms that belong to the AF&C subsector seek applications directed toward agricultural production, it is hypothesized that, in order to gain positive economies of scale (low transaction costs), agricultural biotechnology prefer to locate in regions with significant agricultural production (*Farmland*) (county level data for farmland was collected from U.S. Census of Agriculture). Similarly, since the D&P and MD&E subsectors seek applications directed toward human health, we hypothesize that a county with more hospitals will have a positive effect on the location of these subsectors. County-level data for this variable was collected from county business patterns (U.S. Census Bureau, 1998 and 2003) with 622110 NAICS classification, which includes number of general medical and surgical hospitals.

In terms of conventional location theory, local property taxes (*Property Tax*) may discourage new investment by increasing the costs of production. However, in case of high-tech firms (such as the biotech industry), high property taxes are consider to be proxies for greater availability or higher quality of local public goods (Goetz and Rupasingha), which in turn reflects high standards of living of the local community. Therefore, we assume that property taxes are positively correlated with the location of all four biotech subsectors (county-level data for property tax collections was obtained from the U.S. Census Bureau). Similarly, counties with high unemployment (*unemployment*), which reflect low standards of living of the local community, is considered to have a negative effect on the location-decision of biotech firms. County-level data for unemployment and poverty rate are collected from Economic Research Service USDA and U.S. Bureau of Census, respectively. In order to measure the impact of urbanization economies on the location of biotech firms, we group U.S. counties into Metro (counties belonging to rural-urban continuum codes of 1-3) and Non-metro (counties belonging to rural-urban continuum codes of 4-9). The study assigns a value of 1 for non-metro counties and 0 for metro counties. Since, most of the existing biotech firms are located in the metropolitan cities, we hypothesizes that the variable (*Metro-Nonmetro*) will have a negative impact on the location-decision of biotech firms.

The impact of labor quality on the location decision of biotech industries is measured by county-level average wage per job (*Wage*) and percentage of persons with a college degree (*Education*) (Zucker, Darby, and Brewer, 1998). Both variables are considered to have a positive relationship with the site-selection of all four biotech subsectors. County-level data for *wage* and *education* are obtained from the Bureau of Economic Analysis and Economic Research Service USDA, respectively. Counties with

high median household incomes (*Income*), which represents a high standard of living, are considered to have a favorable impact on the biotech firms' site-selection (county-level data for median household income was obtained from Economic Research Service USDA). Similarly, median housing value (*Housing Value*) is used as a proxy for the quality of housing in a given county. It is expected to have a positive impact on the location-decision. County-level data for median housing value was obtained from U.S. Census Bureau. Table 5.2 presents the descriptive statistics of all the variables included in the model.

Table 5.2. Summary Statistics of Variables.

Variable	Mean	Std. Dev.
Agricultural Feedstock and Chemicals Establishments (number)	3.07	5.44
Drugs and Pharmaceuticals Establishments (number)	0.58	2.84
Medical Devices and Equipment Establishments (number)	1.81	7.95
Research and Testing Establishments (number)	6.06	28.40
Nonmetro-Metro (dummy)	0.65	0.48
Owner-Occupied Housing Units: Median Value (in $1000)	80.93	41.94
Unemployment Rate (percent)	5.97	1.96
Median Household Income (in $1000)	36.73	9.28
Total Number of Venture Capital Firms (number)	1.89	11.45
Number of Colleges, Universities and Professional Schools (number)	1.08	4.86
Number of General Medical and Surgical Hospitals (number)	1.75	3.70
Average Wage per Job (in $1000)	27.05	5.79
Percent of Persons with a College Degree (percent)	16.51	7.80
Property Tax (in $1000)	22.74	23.41
University Life Science R&D Expenditures (in $1000)	563.61	573.18
Land in Farm Acres (in 1000 acres)	301.14	385.21

5.4 Econometric Model

In order to analyze specific factors affecting the location of each of the four biotech subsectors, the present study utilizes seemingly unrelated regression (SUR). In general, there could be some non-observed characteristics of locations that are impacting the four biotech subsectors. Therefore, the error terms across the four equations may be correlated. A separate estimation of each of the four equations will ignore this correlation

and may result in inefficient parameter estimates. The SUR model, which estimates the four equations simultaneously, follows:

(5.1) $$Y_j = X_j^{'}\beta_j + \mu_{j,}$$

$$j = 1 \text{ to } 4,$$

where Y_j is the number of establishments, X_j is the set of explanatory variables, β_j is the vector of parameters to be estimated for equation j, μ_j indicates the error term which includes unobservable cross-section specific effects. Breusch-Pagan test (BP) could be employed to examine the correlation between residuals from the four equations. Failing to reject the BP test of independence, will allow us to estimate each equation separately. In addition to the BP test, the study tests for aggregation bias, which test hypothesis about the coefficients across equations. It test the hypothesis that the coefficients of a given variable across equations are similar indicating that no aggregation bias is present and all the four equations can be aggregated into one equation. For example, to test the hypothesis that coefficients of variable x_2 are similar across four equations, that is $\beta_{12} = \beta_{22} = \beta_{32} = \beta_{42}$, we may write in the form of linear hypotheses (Zellner 1962)

(5.2) $$\begin{aligned} \beta_{12} - \beta_{22} &= 0 \\ \beta_{12} - \beta_{32} &= 0 \\ \beta_{12} - \beta_{42} &= 0 \end{aligned}$$

5.5 Results

Tables 5.3 and 5.4 present the parameter estimates and error correlations for each of the four subsectors, respectively. The Breusch-Pagan test for the presence of correlation in the error terms across equations resulted in a test statistic value of 3874.06 and was found to be significant at the 1% level. Hence, the null hypothesis that error correlations are zero and each equation is independent is strongly rejected. Similarly, test for aggregation

Table 5.3. SUR Estimates of Factors Affecting the Location of U.S. Biotech Subsectors

Variables	Agriculture Feedstock & Chemicals	Drugs & Pharmaceuticals	Medical devices & Equipment	Research & Testing
Median Housing Value	-1.38E-05***	3.11E-06**	1.34E-05***	4.08E-05***
	(-4.51)	(2.18)	(4.16)	(3.20)
Unemployment	0.1570***	0.0112	0.0539	0.2197
	(3.54)	(0.54)	(1.15)	(1.18)
Median Household Income	0.0001***	3.04E-05***	0.0001***	0.0001*
	(6.29)	(4.55)	(5.50)	(1.92)
Venture Capital	0.0755***	0.0043	-0.0045	0.4485***
	(7.71)	(0.93)	(-0.43)	(10.91)
Colleges	0.5402***	0.3343***	1.0151***	3.3282***
	(22.82)	(23.63)	(33.06)	(33.48)
Hospitals		0.1034***	0.4274***	
		(6.66)	(13.58)	
Average Wage Per job	-3.25E-05*	8.04E-08	-3.57E-06	0.0003***
	(-1.83)	(0.01)	(-0.19)	(4.09)
Education	-0.0090	-0.0005	-0.0091	0.2537***
	(-0.62)	(-0.08)	(-0.59)	(4.14)
Property Tax	3.49E-05***	-1.59E-06	-1.16E-05***	-0.0001***
	(9.18)	(-0.88)	(-2.87)	(-3.25)
Farm Land	1.57E-06***			
	(8.32)			
Metro-Nonmetro	-0.6904***	0.0490	0.0232	0.4306
	(-3.76)	(0.57)	(0.12)	(0.56)
Life Science R&D	0.0006***	0.0002***	0.0006***	0.0025***
	(4.52)	(2.93)	(4.34)	(4.56)
Constant	-0.8945***	-1.4933***	-4.3303***	-20.2129***
	(-1.39)	(-5.03)	(-6.47)	(-7.62)
R-Square	0.47	0.58	0.72	0.66

Note: *,**,*** Statistical significance at the 10, 5, and 1% levels, respectively.
Values in the parenthesis indicate standard errors.

bias rejected the null hypothesis of no aggregation bias based on significant critical chi-square statistics for individual variables, thus indicating that the four equations SUR is a better fit for the data under consideration. Each equation in the SUR model was found to be statistically significant as indicated by strong chi-square values, which are found to be significant at the 1% level. The goodness of fit measure for each equation in the SUR model was indicated by R-square values varying from 0.46 to 0.72. The MD&E subsector had the highest R-square value of 0.72.

The median housing value had a negative and significant impact on the location of biotech firms related to the AF&C subsector, whereas it was found to have a positive and significant effect on the location-decision of the other three subsectors. This result is in accordance with the fact that most biotech firms are located in major metropolitan cities where housing costs are considered to be high. Conversely, biotech firms related to agriculture are considered to be located in counties with large farmland where housing costs are considered to be low. The unemployment rate variable was found to be insignificant in the location-decision of all biotech subsectors, except the AF&C subsector, where it was found to be positive and significant at the 1% level. Urbanization economies variable (*Nonmetro-Metro*) was found to be insignificant in the location of all four subsectors, except the AF&C subsector. The reason for this can be that this variable is correlated with other explanatory variables, which are picking up its effects. The variable was found to be negative and significant at the 1% level in the AF&C subsector equation.

Table 5.4. Error Correlations Across Equations.

	Agriculture Feedstock & Chemicals	Drugs & Pharmaceuticals	Medical devices & Equipment	Research & Testing
Agriculture Feedstock & Chemicals	1.00	0.17	0.19	0.14
Drugs & Pharmaceuticals	0.17	1.00	0.61	0.57
Medical devices & Equipment	0.19	0.61	1.00	0.68
Research & Testing	0.14	0.57	0.68	1.00

Breusch-Pagan test of independence; chi(2) = 3874.06, Pr = 0.0000

A county's median household income, which reflects local standard of living, was found to have a positive and significant impact on the location of all four subsectors. This result is consistent with previous literature, which indicates biotech firms' preference for locating in regions with a high standard of living and well developed infrastructure.

Availability of local venture capital was found to be positive and significant in the location decision of the AF&C and R&T subsectors, indicating their dependency on local financial sources. This result is consistent with earlier findings of Powell, Koput, Bowie, and Smith-Doerr (2002) that venture capital firms are an important source of capital for new and small biotech firms. The variable that accounts for state level life science research and development expenditures was found to be positive and significant in the location decision of all four subsectors. This result is in accordance with the current spatial concentration of biotech firms in the West and Northeast, which include a greater number of states spending large amounts of money on research and development. The property tax variable was found to have a negative and significant impact on the location of firms related to the R&T and MD&E subsectors, indicating their preference to locate in counties with low property tax. Since the cost associated with research and development of an innovative new drug is more than $800 million (DiMasi, Hansen, and Grabowski 2003; Congressional Budget Office 2006), these firms prefer to locate in regions that provide R&D tax credits and funding sources for their facilities. However, property tax was found to have an unexpected negative and significant impact on the location of firms related to the AF&C subsector.

Number of colleges in a given county was found to be one of the most significant variables affecting the location of all four subsectors. The variable coefficients range between 0.33 (D&P) to 3.33 (R&T). In the case of R&T subsector, the coefficient for college variable indicates that as number of colleges in a given county increase by one, the number of firms that belongs to the R&T subsector, increase by 3.33. This result is broadly consistent with earlier findings in Audretsch and Feldman (1996), Zucker, Darby,

and Brewer (1998), and Furman et al. (2006) that knowledge spillovers from university research to innovative firms are at work in the biotech related R&D industry.

The amount of farmland in a given county was found to be positive and significant in the location of AF&C subsector. Agricultural biotech firms' aim at improving agricultural production, as a result of which, they prefer to locate in counties with large agricultural production. This will allow firms to utilize local farmland to perform the required tests to develop a genetically modified crop variety that can perform well under local climatic conditions. Similarly, the D&P and MD&E subsectors prefer to locate in close proximity of research hospitals to perform the required tests for developing a new drug. This result is indicated by the positive and significant coefficients of number of hospitals variable for both the subsectors.

A county's average hourly wage and education variable, which are considered to be proxies for availability of high-skilled labor, are found to be positive and significant at the 1% level for the R&T subsector. However, the average hourly wage variable was found to have a negative and significant impact on the location of AF&C subsector. Since the AF&C subsector requires low-skilled labor relative to the R&T subsector, the former is considered to pay less, compared to the later.

5.6 Conclusion

Biotech firms are mainly research and development (R&D) oriented and operate in collaboration with research-oriented universities, biomedical research centers, and other diversified companies that aid in production and distribution of biotech products. Since biotech products are related to drugs and pharmaceuticals, agriculture, and environment, there is no single NAICS or SIC code for the biotech industry. The study utilizes Battelle and SSTI's (2004) bioscience classification, which include four subsectors that are

involved in biotech activities. The four major subsectors are: 1.agricultutral feedstock and chemicals, 2. drugs and pharmaceuticals, 3. medical devices and equipment, and 4. research and testing.

Over the past two decades the U.S. biotech industry has experienced significant growth, resulting in an increase in size and number of establishments. Currently, several state and local economic development agencies are designing and implementing strategies to attract new biotech firms, resulting in stiff competition among and within states. As a result of this increasing competition, the U.S. biotech industry is experiencing some changes in its geographical distribution. However, only some new state/regions are likely to attract biotech firms, as most biotech firms are tending to cluster along existing biotech regions. Several studies have analyzed the location aspects of the biotech industry, however, our understanding of differences in factors affecting the location of each of the four biotech subsectors, is anecdotal. This study employs a seemingly unrelated model that simultaneously analyzes factors affecting site-selection of each of the four subsectors.

Proximity to research institutes to collaborate with skilled labor and obtain new technologies, was found to have the most significant impact on the location of R&T subsector. Positive and significant estimates of availability of local venture capital firms and state-level research and development expenditures indicated the importance of public and private financial sources in the location-decision of R&T subsector. This conclusion is also indicated by a negative and significant estimate of property tax variable. Firms related to the R&T subsector also prefer to locate in a county with access to skilled labor and a high standard of living; this is indicated by the positive and significant estimates of the median housing value, median household income, average wage, and education variables.

The D&P and MD&E subsectors prefer to locate in close proximity of research institutes and hospitals to access skilled-labor and develop and test the new drugs. Similar to the R&T subsector, the two subsectors prefer to locate in counties with high standard of living, which is indicated by the positive and significant coefficients associated with the median housing value and median household income variables. These findings may hinder rural areas hopes of attracting firms related to these three subsectors. However, rural areas that are adjacent to major metropolitan cities are still capable of attracting manufacturing firms that produce intermediate products and drugs that have achieved commercial scale (Gray and Parker 1998) . In order to attract these firms, the rural areas should develop their infrastructure, especially, highway access, housing facilities, and business incentives.

Firms belonging to the AF&C subsector prefer to locate in counties with low median housing values and average hourly wage, and high unemployment rate. They also prefer to locate in counties with large farmland, so that they can improve the quality of local crops by developing locally suitable genetically modified crop varieties. However, they avoid counties that are highly rural with a poorly developed infrastructure. These results indicate that counties with large farmland can attract agriculture related biotech firms provided they develop the necessary infrastructure suitable for them. Thus, the state and local economic development agencies should design strategies based on the type of firm they want to attract.

Future research is directed toward analysis factors affecting the location of each of the four subsectors by employing spatial econometric models, which account for spatial concepts (spatial dependence and heterogeneity) associated with data collected from points in space (Anselin 1988) . Furthermore, including county-level variables related to the state and local economic development incentives, R&D expenditures, and

environmental constraints may further enlighten our understanding of the biotech industry location.

5.7 References

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

Audretsch, D. B. and M. P. Feldman. (1996). "Knowledge Spillovers and The Geography Innovation and Production." American Economic Review, 86: 630-640.

Battelle Technology Partnership Practice and State Science and Technology Institute. 2004. " Laboratories of Innovation: State Bioscience Initiatives 2004." Biotechnology Industry Organization, Washington D.C.

Begemann BD. 1997. "Competitive strategies of biotechnology firms: Implication for US Agriculture." Journal of Agricultural and Applied Economics 29:117-22.

Brennan, M.F., Pray, C.E., & Courtmanche, A. 1999. "Impact of industry concentration on innovation in the U.S. plant biotech industry." Paper presented at the Transitions in Agbiotech: Economics of Strategy and Policy conference, Washington, DC.

Biotechnology Industry Organization. 2005-2006. "Guide to Biotechnology." Biotechnology Industry Facts. BIO, Washington D.C.

BioWorld Publishing Group. 2005. BioWorld Today.

California Trade and Commerce Agency, Office of Economic Research. 2001. Biotechnology. May 2001.

Congressional Budget Office. 2006. Research and Development in the Pharmaceutical Industry. A CBO Study, Congress of the United States. October 2006.

Cortright, J., & Mayer, H. (2002). "Signs of Life: The Growth of Biotechnology Centers in the U.S." Portland: The Brookings Institution Center on Urban and Metropolitan Policy.

Council for Biotechnology Information. 2004. "Biotech Acres: Global Biotech Plantings Show Double-Digit Growth for Ninth Straight Year." <www.whybiotech.com>.

Dahlander, L. and McKelvey, M. (2003): Revisting Frequency and Spatial Distribution:Innovation Collaboration for Biotech Firms. Paper presented at DRUID's Summer 2003 Conference, Helsingore, June 12-14.

Darby M.R. and L.G. Zucker, 1996. "Star Scientists, Institutions, and the Entry of Japanese Biotechnology Enterprises," NBER Working Papers 5795, National Bureau of Economic Research, Inc.

Dibner, Mark D. 1995. Biotechnology Guide U.S.A., 3rd ed., Research Triangle Park, NC: Institute for Biotechnology Information (IBI).

DiMasi, J. A., R. W. Hansen, and H. Grabowski. 2003. The price of innovation: new estimates of drug development costs. Journal Of Health Economics 22:151-185.

Ernst & Young. 2000. "The Economic Contributions of the Biotechnology Industry to the U.S. Economy." Prepared for Biotechnology Industry Organization. Ernst & Young Economics Consulting and Quantitative Analysis.

Ernst & Young. 2004. "Resurgence: The Americas Perspective–Global Biotechnology Report 2004 Ernst & Young LLP.

Ernst & Young. 2005. "Beyond Borders: The Global Biotechnology Report." Ernst & Young LLP.

Furman, Jeffrey L. and Cockburn, Iain M., "Public & Private Spillovers, Location and the Productivity of Pharmaceutical Research" (September 2006). NBER Working Paper No. W12509

Geospatial and Statistical Data Center. 'Crimes reported 1994-2002'. University of Virginia Library. http://fisher.lib.virginia.edu/collections/stats/crime/crimes94.html

Goetz, S. J. and A. Rupasingha. 2002. "High-Tech Industry Clustering: Implications for Rural Areas." American Journal of Agricultural Economics 84(5).

Goetz, S. and S. Morgan. 1995. "State-Level Locational Determinants of Biotechnology Firms". Economic Development Quarterly, (9):174-184.

Graff D. G. 1997. "The Agricultural Biotechnology Industry in Overview." Agricultural and Resource Economics, University of California, Berkeley.

Gray M. and Parker E. 1998. "Industrial Change and Regional Development: The Case of the US Biotechnology and Pharmaceutical Industries." ESRC Centre for Business Research, University of Cambridge, Working Paper No. 95.

Grudkova, V. (2001). The Technology Economy: Why do Tech Companies Go Where They Go? EDA National Forum, Washington, DC (May 30, 2001).

Kalaitzandonakes, N. and Bjornson, B. (1997). "Vertical and Horizontal Coordination in the Agro-biotechnology Industry: Evidence and Implications" Journal of Agricultural and Applied Economics 29: 129-139.

Lerner, J. and Merges, R. 1998. "The Control of Strategic Alliances: An Empirical Analysis of the Biotechnology Industry." Journal of Industrial Economics, Vol. 46. pp. 125–156.

Munroe T., G. Craft, and D. Hutton. 2002. "A Critical Analysis of the Local biotechnology Industry Cluster — Counties of Alameda, Contra Costa, & Solano in California." Volume III, Appendix. Munroe Consulting Inc., Craft Consulting and Hutton Associates. Research Monograph prepared for a Consortium of Bay Area Organizations, May 2002.

National Science Foundation. 2006. Industrial Funding of Academic R&D Continues to Decline in FY 2004. NSF 06-315.

Prevezer, M. 1998. Clustering in biotechnology in the USA, pp. 124–193 in Swan, G. M. P., Prevezer, M. and Stout, D. K. (eds), The Dynamics of Industrial Clustering. International Comparisons in Computing and Biotechnology, Oxford, Oxford University Press.

Powell, W. W., and P. Brantley. 1992. Competitive cooperation in biotechnology:Learning through networks? In Networks and Organizations. Structure, Form andAction, edited by R. Eccles. Boston: Harvard Business School Press.

Powell W. W., K.W. Koput, J. Bowie,. and L. Smith-Doerr. 2002. "The Spatial Clustering of Science and Capital. Journal of Regional Studies, 36 (3).

Sankula S. and E. Blumenthal. 2004. Impacts on US Agriculture of Biotechnology-Derived Crops Planted in 2003– An Update of Eleven Case Studies. National Center for Food and Agricultural Policy, October 2004 Report.

Santoro, Michael D., and Alok K. Chakrabarti. "Building Industry-University Research Centers: Some Strategic Considerations." International Journal of Management Reviews (September 1999): 225 - 245.

Sporleder, T.L., L.E. Moss, and L.A. Nickles. 2002. "Agricultural Biotechnology Start-Ups: Does Venture Capital Matter?" Prepared for presentation at the Third Annual Agricultural Biotechnology Conference, Ohio's Future in Functional Foods, Reynoldsburg, OH.

Sporleder, T.L. and L.E. Moss. 2004. "Knowledge Capital, Intangible Assets, and Leverage: Evidence from U.S. Agricultural Biotechnology Firms." Paper Presented at the 2004 Symposium of the International Food and Agribusiness Management Association Meetings, Switzerland, June 12-13, 2004.

Stewart R. 2004. States of Innovation Biotechnology Locations. Business Facilities The Location Advisor. http://www.businessfacilities.com/bf_04_11_cover.asp (last visited 12/12/06)

U.S. Census Bureau. 2003. County Business Pattern. Data Available at http://censtats.census.gov/cbpnaic/cbpnaic.shtml

U.S. Department of Agriculture. 2003. Economic Research Service. Data Available at http://www.ers.usda.gov/Data/

___. National Agricultural Statistics Service. Historical Data. Available at http://www.nass.usda.gov/Data_and_Statistics/index.asp

U.S. Department of Labor. Bureau of Labor Statistics. Available at http://www.bls.gov/home.htm

Xia, Y., and Buccola, S. 2005. University life science programs and agricultural biotechnology. American Journal of Agricultural Economics, 87(1), 229-243.

Yang H. and S. Buccola. 2003. "University-Industry Relationships and the Design of Biotechnology Research." Paper Presented at the American Agricultural Economics Association Annual Meeting, Montreal, Canada, July 2003.

Zucker, L. G., M. R. Darby, and M. B. Brewer. 1998. Intellectual human capital and the birth of U.S. biotechnology enterprises. The American Economic Review 88:290-306.

Zucker L. G. & M.R. Darby & J.S. Armstrong, 2003. "Commercializing knowledge: university science, knowledge capture and firm performance in biotechnology," Proceedings, Federal Reserve Bank of Dallas, issue Sep, pages 149-170.

CHAPTER 6

SUMMARY, CONCLUSIONS AND FUTURE RESEARCH

6.1 Summary and Conclusions

The biotech industry is one of the fastest growing industries in the United States. Biotechnology is applied in fields as diverse as agriculture, environment, and drugs and pharmaceuticals. The biotech industry includes firms involved in R&D services, testing laboratories, and manufacturing of biotech products. In addition, the industry collaborates with research universities and hospitals for developing new biotech products. Over the past two decades the U.S. biotech industry has experienced significant growth, resulting in an increase in size and number of establishments. Since most of the state and local economic development agencies and research economists believe industrial cluster analysis as a policy solution for all regional problems (vom Hofe and Chen 2006). Biotechnology cluster development is gaining importance as one of the most vital strategies for economic development to enhance regional growth (Grudkova, 2001). Several studies have analyzed the location aspects of the biotech industry, however, our understanding of the spatial influence on the regional distribution of biotech establishments, is anecdotal.

To analyze the effects of numerous firm-specific, location-specific, and inter- and intra-industry agglomeration factors on the location of the U.S. biotech industry, three studies were conducted. The first study utilized a Bayesian spatial tobit model and examined the overall and regional differences in factors affecting the location of the U.S. biotech industry. The second study examined the inter- and intra-industry spatial association of biotech related R&D and testing facilities across all contiguous U.S. counties employing a Spatial 2SLS model. Finally, the interdependence between the four

subsectors of the U.S. biotech industry was analyzed using a Seemingly Unrelated Regression model.

The first study confirmed the hypothesis of spatial agglomeration for the spatial structure of the biotech industry, indicating that biotech firms are positively correlated across counties, resulting in clustering of biotech production. Availability of venture capital firms, research institutions, and hospitals were found to have the most significant impact on the location of biotech firms. This indicates that the biotech firms prefers to locate in regions where they have a source for financing their business , access to research institutes to collaborate with skilled labor and obtain new technology, and access to hospitals for research, testing and marketing of new biotech products. Biotech companies also prefer to locate in counties with a well developed infrastructure and a high standard of living.

The regional models in this study also shredded some light on the regional differences in factors affecting the location of new biotech establishments. Prospective biotech firms willing to locate in the West prefer to establish in metro-counties with easy access to research institutes and skilled labor pool, and that are associated with low median housing values, property taxes, and crime rate. Since the existing biotech firms in the Northeast are mainly associated with manufacturing and marketing of biotech products, biotech firms preferences are different in the Northeast compared to other regions. Biotech firms that are willing to locate in the Northeast prefer counties with easy access to funding sources (venture capital firms and state incentives), hospitals for research, testing and marketing of new biotech products, and the ones that are associated with low median housing value and property tax.

Spatial clustering of biotech research and testing activities was confirmed in the second study. Proximity to manufacturing firms and research universities, and availability of venture capital firms were found to have the most significant impact on the location of

R&D and testing facilities. The significance of both spatial agglomeration economies and research institutes indicate that public as well as private spillovers are at work in the R&D and testing industry, resulting in their spatial clustering. The significant negative sign associated with property tax estimate indicates that in order to attract new R&D and testing firms and develop the existing ones, the state and local economic development agencies should provide certain tax credits and business incentives.

The third study, which analyzed factors affecting the location of four subsectors simultaneously, indicated that firms belonging to the agricultural feedstock and chemicals subsector prefer to locate in counties with low median housing values and average hourly wage, and high unemployment rate. They also prefer to locate in counties with large farmland, so that they can improve the quality of local crops by developing locally suitable genetically modified crop varieties. However, they avoid counties that are highly rural with a poorly developed infrastructure. These results indicate that counties with large farmland can attract agriculture related biotech firms provided they develop the necessary infrastructure suitable for them. Conversely, the drug and pharmaceuticals, and medical devices and equipment subsectors prefer to locate in close proximity of research institutes and hospitals to access skilled-labor and develop and test the new drugs. Similar to the research and testing subsector, the two subsectors prefer to locate in counties with high standard of living, which is indicated by the positive and significant coefficients associated with the median housing value and median household income variables. These findings may hinder rural areas hopes of attracting firms related to these three subsectors. However, rural areas that are adjacent to major metropolitan cities are still capable of attracting manufacturing firms that produce intermediate products and drugs that have achieved commercial scale (Gray and Parker 1998) . In order to attract these firms, the rural areas should develop their infrastructure, especially, highway access, housing facilities, and provide business incentives. Thus, the state and local economic

development agencies should design strategies based on the type of firm they want to attract.

One of the limitations of this research is that information related to the extent of biotech activities (direct or indirect) performed by each establishment (dependent variable) included in the model is unknown. The study assumes that all establishments included the model (dependent variable) perform biotech activities at the same level and they are all of same size.

6.2 Future Research

Future research is directed toward simultaneous analysis of factor affecting the location of each of the four subsectors by employing spatial econometric models, which account for spatial concepts (spatial dependence and heterogeneity) associated with data collected from points in space (Anselin 1988) . Furthermore, including county-level variables related to the state and local economic development incentives, R&D expenditures, and environmental constraints may further enlighten our understanding of the biotech industry location.

6. 3 References

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

Gray M. and Parker E. 1998. “Industrial Change and Regional Development: The Case of the US Biotechnology and Pharmaceutical Industries.” ESRC Centre for Business Research, University of Cambridge, Working Paper No. 95.

Grudkova, V. (2001). The Technology Economy: Why do Tech Companies Go Where They Go? EDA National Forum, Washington, DC (May 30, 2001).

vom Hofe, R. and Chen, K. 2006 Whither or not Industrial Cluster: Conclusions or Confusions. The Industrial Geographer 4: 2-28.

VITA

Pramod Sambidi attended high school at St. Joseph's high school, Ramanthapur, Hyderabad, India, and graduated in May of 1991. He obtained a bachelor of science majoring in agriculture from the College of Agriculture, Acharya N. G. Ranga Agricultural University, Hyderabad, India, in July 1999. He attended Louisiana State University, US, and earned a Master of Science degree in Agricultural Economics in 2003. In Summer of 2003, he enrolled in the Agricultural Economics doctoral program at Louisiana State University. He is currently a candidate for the degree of Doctor of Philosophy, which will be conferred in May of 2006. Upon completion of his doctoral studies, he will be employed as a Regional Economist at the Houston-Galveston Area Council in Houston.