To the Graduate Council:

I am submitting herewith a thesis written by Lance Andrew Stewart entitled "Identifying Rural Comparative Advantage: Ethanol Plant Location Determinants and Tennessee Value Chains." I have examined the final paper copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Agricultural Economics.

_____
Dayton M. Lambert, Major Professor

We have read this thesis and
recommend its acceptance:

Daniel De La Torre Ugarte
_____

Burton English
_____

Kim Jensen
_____

Acceptance for the Council:

Carolyn R. Hodges
_____

Vice Provost and Dean of the
Graduate School

(Original signatures are on file with official student records.)

IDENTIFYING RURAL COMPARATIVE ADVANTAGE: ETHANOL PLANT LOCATION DETERMINANTS AND TENNESSEE VALUE CHAINS

A Thesis
Presented for the
Master of Science Degree
The University of Tennessee, Knoxville

Lance Andrew Stewart
December 2008

Abstract

As rural areas struggle to adjust to the changing U.S. economy with increasing unemployment, falling wages, and constrained capital markets, economic developers look for strategies to promote economic expansion. Development strategies identifying and evaluating county comparative advantage may offer the promise of economic growth in rural areas. This thesis develops two models whereby county comparative advantage can be empirically identified and evaluated. The study first examines ethanol plant location determinants at the county level, in the contiguous forty-eight United States, the second identified industry clusters within Tennessee at the county level and estimated the extent to which these clusters contributed to growth in labor productivity.

In the first study, the location of grain-based ethanol plants is determined by infrastructure, product and input markets, fiscal attributes of local communities, and state and federal incentives. Bivariate probit regression along with spatial clustering methods are used to analyze investment activity of ethanol plants at the county level for the contiguous 48 United States from 2000-2007. The ability of a county to supply feedstock, and the absence of previously established ethanol plants, dominated the site selection decision between 2000 and 2007. Other factors, such as access to railroads or navigable rivers, product markets, low worker wages, producer credit and excise tax incentives, and methyl tertiary-butyl ether bans gave some counties comparative advantage with respect to attracting grain-based ethanol plant investment.

The second study identified industry clusters or economic linkages between purchasers and suppliers, at the county and regional level for 447 economic sectors in Tennessee. Information about value-added activities or innovative potential is possible by

determining the sector composition of the value chains defining an industry cluster. The cluster analysis was extended to estimate the extent to which specific value chains contributed to economic growth between 2001 and 2006 across Tennessee's 95 counties using an econometric model. County and regional comparative advantage was determined by testing whether the presence of a particular value chain in a given county increased labor productivity during this period.

TABLE OF CONTENTS

I: Introduction 1
References 7

II: Location Determinants of Ethanol Production Facilities and Planned Sites: 2000-2007 8
Introduction 9
Methods and Procedures 11
Empirical Model and Estimation Procedures 16
Empirical Model and Location Factors 17
Market Determinants (M) 19
Labor Quality and Availability (L) 22
Infrastructure (I) 23
Local Fiscal Determinants and State Incentives (S, F) 24
Regional control variables 25
Estimation Procedures 25
Spatial Heteroskedastic Autocorrelation Consistent (HAC) Covariance Estimation 25
Ex Post Spatial Cluster Analysis 26
Loess of Selection Probabilities 26
Data 27
Results 28
Correlated Location Decisions 28
First and Second Stage Independence 29
Spatially Heterogeneous Effects of Location Determinants 29
Market Determinants 29
Labor Determinants 31
Infrastructure Determinants 31
Local Fiscal Determinants and Incentives 32
Regional Variables 32
Spatial Clustering Analysis 33
Location Model Application 34
Loess Regression with Location Probabilities over the Urban-Rural Continuum 35
Summary and Conclusions 35
References 39
Appendix 43

III: Identification of Tennessee Value Chains 51
Introduction 52
Methods and Procedures 54
Conceptual Overview of Industry Clusters and Value Chains 54
Nonparametric Identification of Value Chains with Sector Transaction Data 56
Empirical Model Measuring the Influence of Clusters on Labor Productivity Growth 61

Estimation Methods 63
Spatial Clustering of County by Value Chain Coefficients 64
Data 65
Results 66
Step One: Industry Cluster Identification with 2001 Sector Transaction Data 67
Step Two: Where Did These Clusters Contribute to Growth in Labor Productivity? 70
Probit Selection Equation and Control Variables 70
Labor Productivity Growth Equation and Control Variables 71
Step Three: Analysis of County by Industry Cluster Interactions on Labor Productivity Growth 72
Summary and Conclusions 74
References 77
Appendix 80

IV: Summary and Conclusions 101
Vita 105

LIST OF TABLES

Part 2: Location Determinants of Ethanol Production Facilities and Planned Sites: 2000-2007

1. Descriptive Statistics of Location Determinants....44
2.1. Bivariate Probit Estimates, Main Effects....45
2.2. Bivariate Probit Estimates, Interaction of Location Determinants with the Rurality Index....46
3. Marginal Effects....47

Part 3: Identification of Tennessee Value Chains

4. Singleton, Local Serving, and Government Industrial Enterprise Sectors Excluded from the Value Chain Analysis, Following Feser and Isserman (2007)....81
5. Descriptive Statistics of Control Variables....82
6. 2001 Agricultural Cluster....83
7. 2001 Forestry Cluster....84
8. 2001 New Building Construction Cluster....85
9. 2001 Textile Cluster....86
10. 2001 Electronics Manufacturing Cluster....87
11. Agricultural Cluster Regression....88
12. Forestry Cluster Regression....89
13. New Building Cluster Regression....90
14. Textiles Cluster Regression....91
15. Electronics Manufacturing Cluster Regression....92

LIST OF FIGURES

Part 2: Location Determinants of Ethanol Production Facilities and Planned Sites: 2000-2007

1. Spatial Distribution of Active Ethanol Plants....48
2. Spatial Distribution of Announced Ethanol Plants....49
3. Loess Curves for Clustered Location Probabilities....50

Part 3: Identification of Tennessee Value Chains

4. Schematic of the Industry Cluster and Value Chain Concept Used in this Research....93
5. Candidate Clusters Identified with Ward's Clustering Algorithm....94
6.1. Spatial Distribution of County by Industry Cluster Interactions, Agriculture and Forestry....95
6.2. Spatial Distribution of County by Industry Cluster Interactions, New Building Construction and Textile....96
6.3. Spatial Distribution of County by Industry Cluster Interactions, Electronics Manufacturing....97
7.1. Local Indices of Spatial Association (LISA) for Industry Clusters, Agriculture and Forestry....98
7.2. Local Indices of Spatial Association (LISA) for Industry Clusters, New Building Construction and Textiles....99
7.3. Local Indices of Spatial Association (LISA) for Industry Clusters, Electronics Manufacturing....100

I: Introduction

Rural areas encounter difficult challenges adapting to changing economic bases, industrial restructuring, technological innovation, and competition from abroad (Barkley 1995; Johnson 2001). Since the 1990's, rural places have struggled to compete with urban areas that provide access to skilled labor, agglomeration economies, and product and input markets. Hard-pressed to adjust to these challenges, residents of rural communities often encounter declining employment opportunities, which results in lower wages and slower growth of wage earnings compared to urban counties (Johnson 2001). As the production technology of the basic sector became more capital intensive, rural employment in the basic sectors declined as relative wages increased. Demand for skilled labor by some manufacturers could not be met, and comparative advantage of many rural areas waned (Johnson 2001). Problems are compounded because young adults migrate out of rural areas without looking for jobs, and families are not drawn to these communities because employment opportunities are limited (McGranahan and Sullivan 2005).

The rural development paradigm of the 1970's and 1980's has been reinvented. While the development strategies implied by economic base theory (i.e., recruitment of manufacturing in general) may still be an option for some rural communities, such an approach makes sense for fewer communities today than it did twenty or thirty years ago. Therefore, it is generally understood today that alternative strategies promoting development of the non-basic sector could be useful in certain communities.

Rural economic development strategies typically focus on the identification and enhancement of community comparative advantage (Waldorf 2007). The Ricardian concept of comparative advantage is useful for understanding how rural and urban

communities benefit from the exchange of goods and services, given that the costs of producing a particular good or service may be relatively lower in rural locations compared to urban places (Deardorff 2005). For some economic activities, rural communities are positioned to specialize and export certain products and/or services for which they have a comparative advantage. The ability of a rural area to produce a good or provide a service with lower opportunity costs, given limited resources, bestows comparative advantage to rural areas (Barkley 2002). Other development strategies have centered on economic base theory which has proven to be a widely accepted model for analyzing local development problems and economic potential (Barkley 2002). Economic base theory divides economic activity in a region into "basic" or "export" activities and "non-basic" or "local" activities. Basic activities are identified as activities that serve markets outside of the community which result in outside money being brought into the local community. The forward and backward linkages that comprise the basic sector connect the local community to outside product and supply markets (Barkley 2002). But, some rural communities have struggled to promote growth in basic sector economies due to the new global economy. However, new development strategies, such as the identification of niche markets, the promotion of natural amenities for recreational markets, and increasing the potential of innovative value-added activities, may offer the promise of development for certain rural areas (Kraybill and Johnson 1989).

This research seeks to develop a model whereby county comparative advantage can be empirically identified and evaluated. This is accomplished with two methods. Local comparative advantage is identified, (1) by examining grain-based ethanol plant location determinants at the county level in the contiguous forty-eight United States,

2000-2007 and (2) by identifying industry value chains within Tennessee at the county level and how they influenced economic growth, 2001-2006.

The second chapter develops a limited dependent variable regression model to estimate which local demographic, economic, and physical resources influence the geographic distribution of grain-based ethanol plant investment decisions from years 2000 to 2007. The promise of alternative fuel sources from corn, switchgrass, and other cellulosic materials has rekindled the notion of basic sector development in some rural areas because of the potential comparative advantage due to their access to feedstock materials (English et al. 2006). Therefore, ethanol production facilities are perceived as one strategy to foster investment in rural America (Novack and Henderson 2007). Rural farming areas are typically rich in energy feedstock. This has encouraged the use of cropland for energy crop production, collection of agricultural residues, and biomass feedstock for conversion to biofuels. Locating ethanol plants in communities with strong agricultural traditions could potentially increase farm income, create jobs, and stimulate development of other small businesses (Shapouri and Gallagher 2005). If the counties more likely to attract investment can be identified and the community attributes that influence the site-selection decision can be compared, insight can be gained as to where local communities could focus scarce resources if recruitment of ethanol production facilities is pursued as a development strategy. However, a one-size-fits-all policy is unlikely to work for all counties.

The third chapter develops a two-part econometric model identifying local industry clusters/value chains within the state of Tennessee. Creating a value-added development strategy based on available human and material resources may present a

viable alternative to manufacturing recruitment for some rural communities (Kraybill and Johnson 1989). Value added activities can potentially increase the value of raw materials in a region, and/or add value to a product by altering its current characteristics to more preferred characteristics in the marketplace (Kraybill and Johnson 1989). Economic development strategies promoting activities that increase the value of a region's available resources (e.g., raw materials, human capital, and amenities) may be an important component of rural development policy. Such strategies could potentially create jobs by attracting, retaining, and enhancing value-added clusters through initiatives that strengthen forward and backward linkages (Breault 2000). In order to create value-added opportunities, setting the stage for the development of agglomeration economies and fostering rural partnerships may provide a means of enhancing regional advantage for rural communities (Waldorf 2007).

The separate sections of this research are linked by the overarching concept of comparative advantage as related to the potential for economic development in rural areas. The second chapter of the thesis is titled "Location Determinants of Ethanol Production Facilities and Planned Sites: 2000-2007". The third chapter of the thesis is titled "Identification of Tennessee Value Chains". Both parts are organized in five sections: (1) Introduction, (2) Methods and Procedures, (3) Data, (4) Results, and (5) Summary and Conclusions.

References

References

Barkley, D.L. 1995. "The Economics of Change in Rural America." *American Journal of Agricultural Economics* 77(5):1252-1258.

Barkley, D.L. 2002. "Employment Generation Strategies for Small Towns: An Overview of Alternatives." *REDRL Research Report*, 09-2001-02.

Breault, R. 2000. "The Evolution of Structured Clusters." *Photonics Tech Briefs*.

Deardorff, A.V. 2005. "How Robust is Comparative Advantage?" *Review of International Economics* 13(5):1004-1016.

English, B.C., D. De La Torre Ugarte, M. Walsh, C. Hellwinkel, and J. Menard. 2006. "Economic Competitiveness of Bioenergy Production and Effects on Agriculture of the Southern Region." *Journal of Agricultural and Applied Economics* 38(2):389-402.

Johnson, T.G. 2001. "The Rural Economy in a New Century." *International Regional Science Review* 24(1):21-37.

Kraybill, D.S. and T.G. Johnson. 1989. "Value-Added Activities as a Rural Development Stategy." *Southern Journal of Agricultural Economics* 21(1):27-36.

McGranahan, D., and P. Sullivan. 2005. "Farm Programs, Natural Amenities, and Rural Development." *Amber Waves*, February 3(1):28-35.

Novack, N., and J. Henderson. 2007. "Can Ethanol Power the Rural Economy*?" The Main Street Economist* 2(1).

Shapouri, H., and P. Gallagher. 2005. *USDA's 2002 Ethanol Cost-of-Production Survey*. Washington DC: U.S. Department of Agriculture, ESCS for. Agr. Econ. Rep. 841, July.

Waldorf, B.S. 2007. "The Effects of Rurality and Industrial Specialization on Income Growth U.S. Counties 2000 to 2003." Working Paper, Dept. of Agr. Econ., Purdue University.

II: Location Determinants of Ethanol Production Facilities and Planned Sites: 2000-2007

Introduction

From 2001 to 2006, ethanol production increased to more than 445 million gallons per year, with an annual growth rate of 20% (Kenkel and Hoclomb 2006). Nearly 20% of the12.73 billion bushels of corn produced in the U.S. went to produce ethanol in 2006 (USDA 2007; Novack and Henderson 2007). Some estimate that one third of the United States' corn crop will be used to produce ethanol by 2012 (Doering 2006). Factors contributing to the mounting interest in ethanol production include biofuel producers' interest in the value-added activities associated with ethanol production, increasing petroleum prices, increasing fossil fuel prices and uncertainty in many oil producing countries, and the growing perception of farmers that the biofuel feedstock market is less competitive compared to traditional commodity markets (Kenkel and Hoclomb 2006). Government incentives also play a pivotal role in biofuel production and plant location. Until recently, ethanol production was not profitable without blender subsidies (Parcell and Westhoff 2006) and profit margins still would fall short for many ethanol producers if government incentives were not in place (Doering 2006). Policies directly or indirectly increase ethanol demand, such as recent bans on methyl tertiary butyl ether (MTBE) in some states, the Clean Air Act Amendments of the 1990s, and the 2005 Energy Policy Act (Novack and Henderson 2007). Conversion of grain crops to ethanol is considered by some as a way to improve environmental quality, reduce the nation's energy concerns, create jobs, and increase income in rural communities (e.g., Novack and Henderson 2007; Parcell and Westoff 2006; Kenkel and Holcomb 2006; English et al. 2006; Low and Isserman 2008). These factors also contribute to the competitive advantage of ethanol production relative to other crop uses (Kleinschmit 2007).

However, a community must have comparative advantage with respect to local resources and other physical and natural amenities that provide investors the competitive advantage they seek before considering ethanol plant recruitment strategies. Investors aim to minimize costs at all levels of production and choosing an optimal location is essential for the long-term profitability of an ethanol plant (Mielenz, Koepping, and Parson 1996). Important site-selection factors include access to feedstock, marketing opportunities for coproducts, access to demand markets, community-related issues (e.g., local policies, subsidies, or regulations), and transportation infrastructure (Mielenz, Koepping, and Parson 1996). However, not all rural communities are endowed with the combination of attributes sought by potential investors, even if they are located in the Corn Belt or other grain producing areas. Therefore, understanding a range of factors that potentially confer comparative advantage with respect to potential ethanol production sites may inform local policy makers about the prospects of recruiting biofuel firms, given community resources constraints. Many local resources are fixed in the short or medium run, but other factors can be adjusted. By identifying community resources that provide potential investors competitive advantage, and then adjusting those factors, a community may increase their comparative advantage relative to their neighbors with respect to attracting potential ethanol plant investment.

This paper examines the influence local product and input market factors, transport and utility infrastructure, labor, state policy and local demographic characteristics within the contiguous forty-eight United States have on the site selection decision of ethanol plants, 2000-2007. Appreciating the geographic heterogeneity of location determinants and their relationship with investor location decisions provides a

model for ranking local communities competing for similar investments. Bivariate probit regression measures which factors influenced the likelihood an ethanol facility locates in a county during this period. A bivariate probit model is used to jointly estimate the location probabilities of established and proposed plants. It is hypothesized that the effects of location determinants are spatially heterogeneous, which would suggest that comparative advantage with respect to attracting ethanol plant investment will vary across the urban-rural landscape. Spatial clustering methods isolate regions more likely to attract investment from the expanding ethanol industry. Finally, loess regression is used to examine how site selection probabilities vary over a rural-urban continuum.

The Methods and Procedures section reviews the relevant literature summarizing important ethanol plant location determinants, the conceptual model analyzing the ethanol plant location decision, and the empirical model and estimation methods. The Data section identifies the data sources used to measure location determinants and site selection. Discussion of results follows, and the final section concludes.

Methods and Procedures

Conceptual Framework

The economic opportunities associated with biofuel have been brought to attention by agricultural economists since the beginning of the decade. As entrepreneurs and investors seek potential sites for ethanol plants in the wake of new plant construction and openings, research investigating the factors determining community comparative advantage with respect to attracting outside investment has flourished (Low and Isserman 2008). The location of a new ethanol plant may create new jobs, increase the local tax

base, and generate community wealth through the backward and forward linkages agriculture has with the wider regional economy (Novack and Henderson 2007; Low and Isseerman 2008). Profitable ethanol plants are usually located near abundant feedstock supplies, reliable transportation systems, close to an adequate water supply and energy sources, and where land is available for plant expansion (Rose, Detch, and Morgan 2005). Many potential sites are located in areas associated with declining populations and high farm program payments (Parcell and Westhoff 2006). But no matter the geographic location, the long-run survival of an ethanol plant depends on minimizing production costs (Dhuyvetter, Kastens, and Boland 2005). Low-cost production is achieved by minimizing feedstock procurement, natural gas, and labor costs. Feedstock procurement costs are reduced with reliable transportation infrastructure, as well as access to ample and reliable feedstock sources. Ethanol producers rely on efficient transportation and coproduct handling, in addition to the resources required to produce biofuels (Baker and Zahniser 2006). Additionally, to increase profitability, ethanol producers can market coproducts such as Distillers' Dried Grains with Solubles (DDGS), a potentially valuable livestock feed supplement. Therefore, locating near dairy or livestock operations may reduce DDGS transport costs (Baker and Zahniser 2006). A second major cost incurred by ethanol producers is due to the natural gas used in the fermentation process. On average, grain-based ethanol plants use 34,800 Btu of thermal energy per gallon of ethanol (Shapouri and Gallagher 2002). Thus, access to natural gas pipelines and their distribution centers is an important location determinant for ethanol plants (Shapouri and Gallagher 2002). Finally, state-wide and federal policies also influence site selection. Most states with significant ethanol production typically have some form of ethanol

subsidy, incentive, or initiative (Parcell and Westhoff, 2006). Several studies identify the geographical attributes attractive to ethanol producers with respect to locating production facilities (e.g. Baker and Zahniser 2005; Dhuyvetter, Kastens, and Boland 2005; Low and Isserman 2008). However, research identifying the location determinants at the national level remains limited therefore this study looks at the site selection problem with secondary data to identify the county attributes attractive for ethanol plant location in the contiguous 48 United States.

The same factors influencing food manufacturing plant investment determine ethanol plant location choices; access to product and input markets, agglomeration economies, and infrastructure (Henderson and McNamara 2000; Goetz 1997; Lambert, McNamara, and Garrett 2006). The purchase of a single input dominates the total cost structure of supply-oriented food manufacturing firms (Connor and Schiek 1997). The ethanol industry falls into the supply-oriented firm type because feedstock procurement costs dominate ethanol production costs (Shapouri and Gallagher 2005). Supply-oriented food processors locate near agricultural inputs to minimize procurement costs (Henderson and McNamara 2000). Ethanol investors prefer to locate in places with these characteristics because they minimize production costs.

Some firm location studies analyze the site selection decision as a two-stage process (Schmenner, Huber, and Cook 1987; Henderson and McNamara 1997, 2000; Lambert, McNamara, and Garrett 2006). Ethanol investors gauge potential sites based on regional, state, and local attributes. In the first stage, a firm selects a region based on broad company objectives such as access to inputs, market penetration, or other criteria in

the firms' objective function. In the second stage, a firm seeks a least-cost site in the selected region for their investment. The second stage of the location decision is

(1) $Z_i = g(\mathbf{M}_i, \mathbf{L}_i, \mathbf{I}_i, \mathbf{S}_i, \mathbf{F}_i)$,

where Z is the site choice in location *i*, g(•) is a cost minimizing site-selection function, and **M**, **L**, **I**, **S**, and **F** are vectors of community attributes representing input and product markets (**M**), labor attributes (**L**), infrastructure (**I**), state incentives (**S**), and local fiscal characteristics (**F**) influencing production costs, respectively.

The conversion of biomass to ethanol uses Leontief production technology (Mas-Collel, Whinston, and Green 1995) because of the fixed technology coefficients transforming grain or cellulosic materials to alcohol. Firms are hypothesized to select a site based on the aforementioned attributes, subject to an indirect cost function consistent with Leontief production technology, with an additional cost-reducing term accounting for potential access to coproduct markets. The indirect cost function is:

(2) $$C_m^i\left(q_m^i, \mathbf{w}^i; \boldsymbol{\alpha}, \boldsymbol{\theta}, \mathbf{p}\right) = q_m^i\left(\sum_k w_k^i \alpha_k^{-1} - \sum_l \theta^l p_{(i,j)}^l\right),$$

where C_m^i is the cost of ethanol production in location *i* incurred by firm *m*; q_m^i is firm *m*'s production capacity in location *i*; w_k^i are costs of the *k*th input at location *i* (including feedstock transportation costs); **α** and **θ** are fixed technical coefficients converting inputs to ethanol and *l* coproducts respectively; and $p_{(i,j)}^l$ is the market price for coproduct *l* discounted for transport costs from county *i* to market in county *j*. The optimal level of input *k*, given plant *m*'s output capacity, is $x_k = q_m^i \alpha_k^{-1}$ regardless of location. It is assumed that investors always choose a least cost location, but the expected costs determined by a potential ethanol investor choosing site *i* compared to other sites is not

completely observed. However, the location events of active and proposed ethanol plants and location factors are observable.

Site selection in county *i* depends on the difference between the expected cost of locating and operating in that county,

(3) $$E[C_m^i(q_m^i, \mathbf{w}^i)] = \mathbf{w}'^i\mathbf{x} - q_m^i \sum_l \theta^l p_{(i,j)}^l + \delta_i ,$$

compared to expected costs of locating and operating in county *j*,

(4) $$E[C_m^j(q_m^j, \mathbf{w}^j)] = \mathbf{w}'^j\,\mathbf{x} - q_m^j \sum_l \theta^l p_{(j,i)}^l + \delta_j ,$$

where δ is a random disturbance term associated with uncertainty about input availability, product and input transportation costs, infrastructure reliability, labor quality, product market potential, and other site-specific or regional attributes influencing costs that cannot be perfectly known by investors. When a firm has perfect information about a potential location, $E[\delta] = 0$.

In practice, the distribution of costs expected by investors and the marginal benefit of selecting location *i* over other sites is not observable by the researcher, but the locations of active and proposed ethanol plants and local factors are measurable. A firm locates in a given county when the expected costs of producing ethanol are lower compared to other counties. Combining the equations representing the expected location costs gives the reduced form equation defining a latent choice variable (*Z**):

(5) $$Z_i^* = E\left[C_m^j(q_m^j, \mathbf{w}^j) - C_m^i(q_m^i, \mathbf{w}^i)\right] = \mathbf{x}'(\mathbf{w}^j - \mathbf{w}^i) + (\delta_j - \delta_i) = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i ,$$

with $Z_i = 1$ if $Z_i^* > 0$; $Z_i = 0$ if $Z_i^* < 0$, with Z_i^* the marginal cost savings of locating in county *i*. Therefore, the location probability associated with any given county is $\Pr[Z_i = 1] = \Pr[\varepsilon_i > -\mathbf{x}_i'\boldsymbol{\beta}] = 1 - F(-\mathbf{x}_i'\boldsymbol{\beta})$, where *F* is a cumulative probability distribution

function. A common strategy to untangle this decision structure is to specify F as the cumulative density function of the normal or logistic distribution. This analysis uses the normal distribution to model county site selection probabilities.

Empirical Model and Estimation Procedures

This research jointly models plant location announcements (e.g., 'proposed' plants) and plants operating (e.g., 'active' or 'established' plants) from 2000-2007. An empirical model is developed from equation (5) to estimate the relationship between established and announced ethanol facilities in a given region. It is hypothesized that the location decisions of established plants influenced the location decisions of location announcements during this time period. This potential interaction is modeled with a bivariate probit regression.

Ethanol plant location announcements are hypothesized to be negatively correlated with active plants due to competition for limited grain feedstock supply. The location probability of ethanol plants established (EST) after 2000 is,

$$(6) \qquad {}^{EST}Z_i^* = \Pr[EST_i = 1] = \Phi({}^{EST}\mathbf{x}_i'\boldsymbol{\beta}_{EST}),$$

and location probability of new ethanol plant announcements after 2000 (ANN) is,

$$(7) \qquad {}^{ANN}Z_i^* = \Pr[ANN_i = 1] = \Phi({}^{ANN}\mathbf{x}_i'\boldsymbol{\beta}_{ANN}),$$

where Φ is the normal cumulative density function. The bivariate location model is,

$$(8) \qquad Z_i^* = \Pr[ANN_i = 1, EST_i = 1] = \Phi_{BVN}({}^{EST}\mathbf{x}_i'\boldsymbol{\beta}_{EST}, {}^{ANN}\mathbf{x}_i'\boldsymbol{\beta}_{ANN}, \rho)$$

where Φ_{BVN} is the bivariate normal cumulative distribution. The general specification for this two-equation system is;

(9) $${}^{EST}Z_i^* = \mathbf{x}'_{EST_i}\boldsymbol{\beta}_{EST} + {}^{EST}\varepsilon_i,\ {}^{EST}Z_i = 1 \text{ if } {}^{EST}Z_i^* > 0,\ 0 \text{ otherwise,}$$

(10) $${}^{ANN}Z_i^* = \mathbf{x}'_{ANN_i}\boldsymbol{\beta}_{ANN} + {}^{ANN}\varepsilon_i,\ {}^{ANN}Z_i = 1 \text{ if } {}^{ANN}Z_i^* > 0,\ 0 \text{ otherwise,}$$

with the cross-equation covariance matrix,

(11) $$\text{cov}({}^{ANN}\varepsilon_i, {}^{EST}\varepsilon_i) = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

The bivariate probit regression estimates are $\boldsymbol{\beta}_{EST}$, $\boldsymbol{\beta}_{ANN}$, and ρ; ${}^{EST}Z_i^*$ is a latent variable for established ethanol firms in county i; ${}^{ANN}Z_i^*$ is a latent variable for ethanol plant announcements in county i; $\mathbf{x}_i$ are local attributes; and ρ is the correlation between the disturbance terms, ${}^{ANN}\varepsilon_i$ and ${}^{EST}\varepsilon_i$.

Equations (8) and (9) are estimated as single equation probit regressions if ρ is not different from zero. Negative correlation between the disturbance terms may suggest feedstock resources are limited. That is, given an established ethanol facility, a proposed facility will tend to choose a site farther away from an established plant to access feedstock sources not consumed by the established operation.

Empirical Model and Location Factors

County-level factors are regressed against binary variables indicating where ethanol plants became operational and ethanol plant location announcements occurred during 2000-2007 with the following model;

(12) $$\Pr[Z_i^{ANN} = 1, Z_i^{EST} = 1] = \Phi_{BVN}(\text{RI}_i, \text{RI}_i*\mathbf{M}_i, \text{RI}_i*\mathbf{L}_i, \text{RI}_i*\mathbf{I}_i, \text{RI}_i*\mathbf{F}_i, \text{RI}_i*\mathbf{S}_i, \mathbf{M}_i, \mathbf{L}_i, \mathbf{I}_i, \mathbf{F}_i, \mathbf{S}_i, \text{Regional fixed effects}; \rho),$$

Z_i is a binary variable indicating if there was at least one active ethanol plant or ethanol plant announcement in a given county between 2000 and 2007, Φ_{BVN} is the standard bivariate normal cumulative density function, $\mathbf{M}_i$, $\mathbf{L}_i$, $\mathbf{I}_i$, $\mathbf{F}_i$, and $\mathbf{S}_i$ are the location determinants in a given county, and RI is an index measuring how rural a county is (Waldorf 2006) constructed with 2000 census data. Waldorf's rurality index is calculated as a function of population, population density, the percent of the population designated as rural or urban according to the US Census, and the distance of a county to a metropolitan statistical area according to the Office of Management and Budget (OMB) urban core-non-core county classification system (OMB 2007). The rurality index is a continuous variable, bounded between [0, 1]. Counties with a rurality index of 1 are remote, low population density counties far away from any urban core (e.g., "completely rural"). The converse is true for counties with a rurality index of 0 (e.g., "completely urban"). The interaction of this measure with location determinants tests the hypothesis that the geographical effects of the determinants vary geographically with respect to the likelihood of site selection, given the position of a county in the rural-urban hierarchy. The marginal effects of a location determinant are therefore a continuous function of the rurality index and vary geographically. The unconditional marginal effect of the *k*th continuous local factor is (Greene 2003),

$$(13) \qquad \frac{\partial \Pr[Z_i = 1]}{\partial x_{ik}} = (\hat{\beta}_k + RI_i \hat{\beta}^{RI}) \phi(\mathbf{x}_i' \hat{\boldsymbol{\beta}}),$$

where ϕ is the normal probability density function, and "^" is an estimate of the weight attributed to a local factor. For discrete (binary) variables (*d*), the marginal effect is (Wooldridge 2006),

$$(14) \qquad \frac{\partial \Pr[Z_i = 1]}{\partial x_{ik}} = \Phi(\mathbf{x}'_{ik-1}\hat{\boldsymbol{\beta}}_{k-1} + RI_i\hat{\beta}^{RI} + x_{ik}RI_i\hat{\beta}_k^{RI}) - \Phi(\mathbf{x}'_i\hat{\boldsymbol{\beta}} + RI_i\hat{\beta}^{RI}).$$

Details of the variables making up the location determinants in **M**, **L**, **I**, **F**, and **S** and regional control variables are discussed below, along with their expected relationships with active and announced ethanol plant locations. Also, descriptive statistics of variables are included in Table 1.

Market Determinants (M)

The market potential of a given area depends on the ability to match demand with the supply of competing goods. Larger product markets are penetrated by exploiting lower transportation costs, thereby increasing the competitive position of a site. Ethanol plants locate where the transportation costs for primary inputs (e.g., corn and other cellulosic materials as feedstock, natural gas for sterilizing mash and drying distiller grains, and electricity for regular operations) and coproduct distribution costs are minimized (Dhuvyetter, Kastens, and Boland 2005). Net feedstock costs account for the largest share of ethanol production costs. About 55% of the per-unit costs of ethanol production are due to feedstock acquisition (Shapouri and Gallagher 2005).

As the ethanol industry grows and becomes more competitive, profit margins will decrease and the marketing of coproducts will become more important (Dhuvyetter, Kastens, and Boland 2005). Distiller's grains (DG) are a highly valued supplement for livestock diets. Each gallon of corn ethanol produced yields roughly 6.3 pounds of dried distiller's grain with solubles (DDGS) (Roe, Jolly, and Wisner 2006). Locating near livestock operations and selling coproducts can potentially offset feedstock procurement costs. Natural gas and electricity costs also command a large share of the variable

operating expenses. Ethanol plants use natural gas in the distillation process, which accounts for the second highest variable operating expense for ethanol production (Shapouri and Gallagher 2005).Choosing sites with historically low natural gas prices as well as sufficient and proximate supplies may allow ethanol producers to hedge against volatile fossil fuel prices by keeping procurement costs low. Location quotients[1] were constructed to measure the influence of county natural gas distribution centers (NATGAS) on the location decision.

Three variables measure the effects of product markets on the location decision of grain-based ethanol plants. Assuming that ethanol will primarily be used as a fuel additive, the per-county number of retail gasoline businesses, plus the sum of the retail gas stations in surrounding counties would account for market potential and demand access. Retail gas establishments are used only as a demand proxy (GAS), due to the assumption that areas with higher concentrations of gas stations typically have higher concentrations of consumers. Ideally, the number of blending facilities in a county would be used to measure access to product markets, but the information was not available. Therefore, this clearly imperfect measure only roughly approximates demand potential. The 2000 total head of cattle per county plus surrounding counties (CATTLE) measures potential access to DDGS markets. Distiller's grain is marketed in wet and dry forms, and may need to be stored or dried before it is shipped to demand centers. A location quotient

[1] Location quotients are a measure of specialization in a given sector. Communities highly specialized in a given sector are more likely to export that particular service or good. $LQ_c^i = \left(\frac{e_c^i}{e_c^t}\right) \Big/ \left(\frac{e_n^i}{e_n^t}\right)$ where e_c^i is the level of economic activity for county *c* in industry *i* and e_c^t is total economic activity and *n* represents the national data (Shaffer, Deller, and Marcouiller, 2004).

of farm product warehousing operations (STORE) measures the influence storage facilities have on ethanol plant location decisions.

Three variables were used to measure the impact of access to input markets on ethanol plant location decisions. Ethanol production relies heavily on the agricultural sector, given demand for feedstock as well as potential DDGS markets. Farm proprietor income was divided by the nonfarm proprietor income of a given county to measure the relative importance of farming in a county, based on the assumption that counties with relatively more income generated by farming will proxy regions that have comparative advantage with respect to farming (FARMPROP). It is expected that ethanol producers are more interested in the total bushels available for feedstock use rather than the crop yield of a given county. Due to the limited ability of a single county to supply all of a large ethanol producer's corn needs, ethanol producers will likely import corn from surrounding counties by truck, rail, or barge. The average of total bushels of corn produced from 1990-2000 in a county was added to the sum of the average of total bushels of corn produced in neighboring counties to measure access potential to corn as feedstock for ethanol (CORN). Structural and strategic barriers to entry into product markets due to preexisting ethanol plants may be an important factor in location decisions. As more firms locate in a given region, competition for that area's limited feedstock resources increases. There were 31 plants producing ethanol before 2000. These firms were included as a measure of entry barriers (ESTAB) for plant location announcements. It is hypothesized that counties with existing active ethanol plants will be less attractive as potential sites for new entrants because of competition for feedstock.

Labor Quality and Availability (L)

Higher quality workers are typically more productive, which leads to increased productivity yielding a higher level of output at the same or lower costs, thereby increasing profits (McNamara, Kriesel, and Deaton 1988). It is hypothesized that a high-skilled labor force will be positively correlated with ethanol plant site selection. The 2000 percent of individuals over the age of twenty-five with a high school diploma was used to measure labor quality effects on ethanol plant location (HS00).

The cost of labor directly affects production costs and manufacturing plant profits. Locations with lower labor costs have lower operating costs, increasing the attractiveness of the area for manufacturing industries (Schmenner, Huber, and Cook 1987; Smith, Deaton, and Kelch 1978; McNamara, Kriesel, and Deaton 1988). It is hypothesized that counties with higher labor costs will be negatively correlated with ethanol plant location and site selection. The 2000 annual manufacturing wage per worker in each county was used to measure the effect of labor cost on the location decision (WAGE).

Productivity within the manufacturing industry depends on available labor. A deep labor pool requires less recruiting and would adequately meet the needs for a larger number of diverse firms. A diversified work force increases the likelihood of acquiring workers with the necessary skills to fill positions at all levels of production. A Herfindahl index was used to measure the effects of a diversified workforce on the location decision of potential ethanol plant investors (Davis and Schluter 2005). As the index approaches one, more individuals are employed by a single sector. The measure was calculated as $HERFEMP_i = \Sigma_k s_{ki}^2$, where s_{ki} are the 2000 shares of workers employed in the

agricultural, forestry and mining, wholesale, retail, service, finance, insurance, real estate, and manufacturing sectors in a county.

Infrastructure (I)

Infrastructure consists of the physical and natural features in an economy that support community and commercial needs by creating access to regional, national, and international markets. Smith, Deaton, and Kelch (1978), Woodward (1992), and Rainey and McNamara (1999) looked at how county level infrastructure effected manufacturing location decisions and all found that infrastructure was a positive determinant of plant location decisions. Ethanol production requires reliable transportation systems to acquire the necessary inputs and to distribute ethanol as well as distiller's grains. Ethanol is typically shipped in tanks by train, truck, or barge and then mixed directly with gasoline in the tanker trucks that deliver fuel to gas stations (Baker and Zahniser 2006). Transportation networks include federal and state roads, railroads, and waterways capable of transporting barges. The total county road network miles, including state highways and the federal interstate system, was normalized by the total square miles of the county to measure the road network potential of a county (ROAD). A similar measure was constructed for the class 1 and 2 railroad networks for each county (RAIL). It is expected that these transportation density measures will be positively correlated with active ethanol plants and announcements. County adjacency to a major river (RIVER) was used to measure the influence of river transportation opportunities on ethanol plant location decisions. Location quotients measure the influence of establishments specializing in truck transport (TRLQ).

Local Fiscal Determinants and State Incentives (S, F)

By the end of 2000, nine states (CA, CO, CT, IA, ME, MI, MN, NE, and NY) had completely banned methyl tertiary-butyl ether (MTBE) (DOE 2001). Given that MTBE is no longer used as a fuel additive in these states, ethanol emerged as a likely substitute. The adverse environmental effects associated with MTBE induced demand for a replacement in which the "eco-friendly status" of ethanol made it a prime candidate as a comparable gas-additive. By July 2001, legislation had passed an excise tax supporting ethanol producers in eight states (AK, CT, HI, ID, IL, IA, MN, and SD) (DOE 2001). The federal excise tax for ethanol producers was designed to make ethanol more competitive as a gasoline additive. For this reason, it is expected that states with a federal excise tax were more likely to attract investment (TAX). In addition, by July 2001, ten states (KS, MN, MI, MT, NE, ND, OK, SD, WI, and WY) had authorized ethanol producer credit incentives that credited the sale of corn for ethanol production (PRODCR) (DOE 2001). A producer credit is a supply-side policy that should have a greater effect in more remote grain producing counties. It is expected that ethanol producer credit incentives have a positive impact on attracting potential ethanol production facilities.

Fiscal policy includes the government expenditure patterns and tax policies of counties and states. While higher state spending can be a benefit in some instances, states with high corporate taxes are less attractive with respect to site selection in the manufacturing industry (Goetz 1997). County-level per capita property taxes were normalized by total county expenditures per capita in 2000 to measure the fiscal effects on the site location decision (FISC).

Regional control variables

The Economic Resource Service (ERS) farm resource regions control for unobserved factors associated with the first-stage location decision of firms; Heartland, Northern Crescent, Fruitful Rim, Northern Great Plains, Prairie Gateway, Basin and Range, Mississippi Portal, Southern Seaboard (ERS 2000). These regions are characterized by the dominant agricultural commodities produced in an area, along with soil, climatic, and demographic attributes.

Estimation Procedures

A Wald statistic (W) tests the equality of RI slope and intercept coefficients, with the null hypothesis that location determinants were geographically stationary; H_0: all RI interaction terms = 0. A Wald test also tests the independence between the first and second stage decisions of ethanol plant location announcements and active plant locations. Lastly, the null hypothesis that the first and second stage location decisions were independent is tested using the Wald statistic, with the null hypothesis that the regional effects were not correlated with the location decisions.

Spatial Heteroskedastic Autocorrelation Consistent (HAC) Covariance Estimation

A spatial heteroskedastic autocorrelation consistent (spatial HAC, Kelejian and Prucha 2007) estimator is used in this research to attend to potential spatial error dependence. Spatial error dependence compromises the efficiency of standard errors. The Spatial HAC covariance is based on the Newey-West estimator (Newey and West 1987) commonly used to attend to temporal autocorrelation and heteroskedasticity in time series applications. In the case of the spatial HAC estimator, covariance between locations is

modeled using a kernel density function that downweights the influence of shocks in counties farther away from a target location. The kernel density function must be real, continuous, and integrate to unity. This application uses the Parzen kernel, which meets these criterions.

Ex Post Spatial Cluster Analysis

Regions exhibiting greater likelihood of attracting ethanol plant investment relative to other regions are identified using the selection probabilities generated by the bivariate probit location model. Spatial clustering techniques are applied using a Local Indicator of Spatial Association (LISA) to identify groups of counties forming high-probability location clusters (Anselin 1995). A 5% level of significance discriminates areas forming low and high probability clusters. The spatial weight matrix used in the procedure is an inverse distance matrix based on county nearest neighbors. Distances between counties are measured from the centroid coordinate of each county (Anselin 2003).

Loess of Selection Probabilities

The rurality index is regressed against the estimated probabilities corresponding with high location probability counties (i.e., counties comprising the positive LISA clusters) to examine the relationship between active ethanol plants and new ethanol plant announcements using loess regression (Greene 2003). The curve generated by the regression indicates where counties are most competitive along the urban – rural continuum.

Data

The plant location and announcement information was collected from the Renewable Fuels Association (2008). The total number of active ethanol plants as of January 3, 2008 was 141, with 70 ethanol plant location announcements. The 2000 cutoff point was chosen for two reasons. First, all plant location announcements documented by the RFA occurred during or after 2000. Announcements are defined as plants that report zero production because the announcements were recent and physical construction had yet to be completed. Whether these plants actually begin production is not a concern since it is the local factors associated with the county which initially attracted interest. Second, 78% (110) of the active ethanol plants had begun production in or after 2000, resulting from the recent and renewed emphasis on expanding renewable fuel supplies in the United States.

Location determinants measured in 2000 (or prior to 2000) are used in the regression analysis to avoid potential simultaneity problems. To evaluate the feedstock input and coproduct output determinants on the site selection decision, crop and livestock production data for the year 2000 were collected from the National Agricultural Statistics Service (2000). Demographic variables were extracted from the 2000 Census files (2000), and information about state policy incentive initiatives was obtained from the United States Department of Energy (2001). Inter-county distances, interstate and state highway miles, county physical attributes, navigable river access, and per county miles of class I and II rail lines are available from ESRI (2006). Information on trucking and natural gas distribution establishments was extracted from the 2000 U.S. Census County Business Pattern files (2000). The RI is Waldorf's (2006) 2000 rurality index constructed

using 2000 census data and the Office of Management and Budget (OMB) urban core non-core county classification system (OMB 2007). Regional control variables were identified using the Economic Research Service resource regions (ERS 2000). After eliminating counties with incomplete information, there were 3,064 usable observations in the final data set.

Results

The pseudo-R^2 measure was relatively low for the bivariate location model (R^2 = 0.26), but overall, location determinants and their interaction with the rurality index were significantly correlated with active ethanol plant locations and new plant announcements (joint Wald test [W], H_0: $\boldsymbol{\beta} = \mathbf{0}$, W = 12,990, degrees of freedom [df] = 90) (Table 2.2). The correlation between the disturbances was significant and strongly negative (discussed below).

Correlated Location Decisions

The correlation between disturbances of the established and announced ethanol plants was strongly negative and significant (ρ = -0.98, W = 4.82, df = 1). The negative correlation between the unobserved factors associated with the location decisions of active and announced ethanol plants suggests some systematic repulsion between active and planned facilities with respect to determining location outcomes. While it is difficult to pinpoint which unobserved factors contribute to this effect, mounting pressure on feedstock availability in certain regions most likely play some role. In a competitive market, new investment would select sites as far away as possible from established plants to access what grain-based feedstock resources remains.

First and Second Stage Independence

The coefficients of the regional variables should not be significant if the assumption that the first and second stage location decisions are independent is tenable. The hypothesis was tested with a joint Wald test. The null hypothesis that the farm resource regions did not influence the second stage decision of ethanol plant location announcements and active plan locations was rejected at the 5% level. Therefore, the regional variables remained in the location model.

Spatially Heterogeneous Effects of Location Determinants

The null hypothesis that location factors were homogenous across the urban-rural continuum was rejected at the 5% level. The marginal effects of the location determinants are discussed looking at their effects in groups of counties falling into the rurality index quantiles of 0 – 0.2, 0.4 – 0.6, 0.6 – 0.8, and 0.8 – 1.0; 0 – 0.2 representing more urban counties, and 0.8 – 1.0 representing remote, rural counties. Discussion of the results focuses on the marginal effects of the location determinants, with emphasis on their interaction with the rurality index (Table 3).

Market Determinants

Product markets had varying effects on location decisions and active plant sites, depending on the rurality of counties (Table 2). The relative importance of farming in a county (FARMPROP) was positively associated with active and announced plant locations. However, the effect decreases moving away from metropolitan areas. The marginal effects in Table 3 indicate that in counties where income from farming is relatively more important than income from other sources just beyond urban core areas

are more attractive to ethanol producers than relatively remote counties. Areas with abundant farm income located in extremely rural areas are generally likely to have access to some types of infrastructure needed for ethanol production. Likewise, the total average corn production in a county is a positive determinant for attracting active ethanol plants and new plant investment. But remote areas appear be negatively correlated with announced and established ethanol plants location (Tables 2, 3), again suggesting the importance of infrastructure determinants in more remote locations. Farm storage operations are an important location determinant for established ethanol plants in more rural areas, perhaps because urban areas do not typically specialize in warehousing agricultural products. Alternatively, established ethanol plants may already command available storage facilities, limiting availability for new plants. The number of cattle in the 0.4 – 0.6 quantile of the rurality index is a significant determinant for attracting ethanol plants. Due to the increasingly competitive nature of the ethanol industry, plants entering the industry may have strong incentives to locate near potential DG markets to lower input procurement costs. Plant announcements from 2000-2007 were negatively associated with plants active prior to 2000, suggesting firms entering the market avoid locations with established ethanol plants already competing for feedstock resources. The number of retail gasoline stations outside urban areas (within the 0.2 – 0.4 quantile) was positively related with ethanol plants. Conversely, announced facilities were positively correlated with retail gas stations in more rural areas. Given that the number of retail gasoline stations in a given area proxies demand potential, it appears that established ethanol plants from 2000-2007 may have already saturated locations near urban markets.

Labor Determinants

Urban counties in the 0.2 – 0.4 quantile range of the rurality index were less likely to attract ethanol plant investment from 2000 – 2007 when wages were relatively higher (Table 3). Labor quality, as measured by the percent of persons with a high school diploma, appears to be an important consideration for plant location. Rural areas become more attractive to potential ethanol plant investors when higher-skilled labor is potentially greater. Labor pool diversity in more rural areas is an important factor attracting new plant investment. The probability an ethanol plant located in a given county decreases as the proxy for a diverse labor force approached 1 (e.g., a more homogenous labor force, suggesting a less diversified economy).

Infrastructure Determinants

Road density in counties along the middle of the urban-rural spectrum is an important location determinant for active ethanol plants (Table 3). However, road density is not correlated with proposed plant sites. New plants flooding the ethanol market at the turn of the century probably occupied prime locations first, including well developed primary and secondary road networks, which in turn may have sent latecomers in search of sites with access to secondary transportation sources. Counties with well developed rail transport systems may have a comparative advantage over other counties with respect to attracting potential ethanol plant investment. Counties in the 0.2 – 0.4 rurality index quantile range with access to river transportation were positively correlated with plant location announcements, but more remote rural areas may not have the infrastructure to support such activities. Trucking establishment density in more urban areas appears to be

an important location determinant. Counties with relatively more trucking facilities may be able to support the transportation demand of ethanol plants in more densely populated counties. Access to natural gas distribution centers did not appear to be a factor with respect to ethanol plant location in either urban or rural counties from 2000-2007.

Local Fiscal Determinants and Incentives

State excise taxes are positively correlated with announced ethanol plant locations (Table 3). In addition, producer credit incentives proved to be a positive determinant for established ethanol plants in counties in the 0.4 – 0.6 quantile of the rurality index. County level per-capita property taxes became an increasingly negative factor for more rural areas with respect to attracting ethanol plants from 2000 to 2007. The ban on MTBE was positively correlated with active ethanol plant sites in counties within the 0.2 – 0.4 range of the rurality index. However, rural areas with bans on MTBE appeared to have a comparative advantage with respect to attracting new ethanol plant announcements.

Regional Variables

The Heartland and Northern Crescent Regions were more likely to attract ethanol plant investment during 2000-2007 because of the importance of corn production and feedstock availability. The Fruitful Rim, Northern Great Plains, and Prairie Gateway appear to be attractive regions for ethanol production. Counties in the Basin Range were more likely to have an established ethanol plant, but were less attractive to announced ethanol plants. The Mississippi Portal is negatively correlated with established ethanol plants. However, as noted previously, navigable river access is an important determinant for announced ethanol plants, which may be why the Mississippi Portal is positively

correlated with announced ethanol plants. The Southern Seaboard region is negatively associated with announced plants, but did not appear to be an important factor for established plants' first-stage location decision.

Spatial Clustering Analysis

The spatial distribution of the estimated site selection probabilities for active grain ethanol plants and proposed plants are presented in Figures 1 and 2, respectively. The differences in the two spatial distributions are worth noting. The estimated site selection probabilities for active ethanol plants are concentrated in the Corn Belt. The spatial distribution of the estimated site selection probabilities suggests that some counties in Iowa, Minnesota, Illinois, Nebraska, western Wisconsin, western Indiana, northern Missouri, east North Dakota, parts of South Dakota, the Texas Panhandle, portions of Oklahoma and Kansas, east Wyoming, northern Colorado, and a small section of California exhibit qualities that were attractive for established ethanol plants. Compared to Figure 2, the spatial distribution of the location probabilities of ethanol plant announcements appears to be more dispersed and fragmented. Areas in Ohio, Missouri, a larger portion of Indiana, a small portion of Pennsylvania and Oregon, Idaho, larger areas in California and Colorado, and a small area in Arizona now appear to be attractive to new ethanol plant investment. Areas likely to attract ethanol plants are less evident in Figure 1. This may suggest that as profits continue to be realized and the feedstock market becomes progressively saturated, prime locations are occupied, forcing new ethanol plants to consider alternative sites in marginal locations.

Location Model Application

Based on the results of the ethanol location model, ideal sites likely to attract future grain-based ethanol investment will most likely seek out counties with qualities such as a history of substantial corn production, relatively more income generated by farming, a lack of previously established ethanol plants, low worker wages, establishments specializing in trucking activities, as well as counties adjacent to navigable rivers. Rural areas could prove competitive if they exhibit qualities such as an educated and/or diversified labor force, dense railroad structures, located within a state with an excise tax incentive, located within a state with bans on MTBE, significant cattle production, and/or a high concentration of gasoline stations. Counties able to adjust some of these factors could also potentially gain a comparative advantage with respect to attracting ethanol plant investment.

In addition, the bottom panel of Figure 2 identifies the 25 counties within the high probability clusters that had the highest probabilities with respect to attracting ethanol plant announcements. The potential production capacity of those sites is also displayed which was calculated by dividing the corn availability in a county plus corn production in neighboring counties by a corn-to-ethanol conversion rate, 2.8 in this case. Admittedly, this is assuming a new ethanol plant will consume all available corn produced in and around a county and will be withheld from other uses. However, looking at the production potential of these areas provides a useful comparison of possible plant capacity in each area.

Loess Regression with Location Probabilities over the Urban-Rural Continuum

The rurality index was regressed against the probabilities isolated from the LISA high probability regions determined with the announced and active plant location models using loess regressions to gauge the extent to which location probabilities varied for active and announced plants over the urban – rural continuum (Figure 3). Very urban or very rural areas exhibit relatively lower probabilities with respect to attracting ethanol plant investment. The loess curve for announced ethanol plants is much flatter, showing little variation in location probabilities across the urban-rural continuum. This may suggest that as the ethanol industry matures, newer plants must search for ethanol support elements along the entire spectrum of urban and rural in second-best sites. Tracing the "frontier" of these location probabilities is revealing (Figure 3). Compared to the probability distribution of active plants over the rural-urban continuum, the likelihood of a new plant announcement occurring between the 0.35 – 0.75 rurality index range plateaus well below the active plant peak. The pattern further demonstrates that newer entrants may be struggling with respect to finding competitive locations because established plants already occupy many of the best sites.

Summary and Conclusions

This analysis used bivariate probit regression and spatial clustering techniques to isolate which location determinants were important with respect to attracting ethanol plant investment in the contiguous 48 United States, during 2000-2007. While many of the factors hypothesized to be important ethanol plant location determinants were significant, some factors were not. The importance of ethanol plant location determinants

varied, depending on how remote the county was, and whether an ethanol plant was established, or announced. While some rural areas exhibited a comparative advantage with respect to attracting ethanol plants, it appears that many rural communities are less likely to attract ethanol plant investment. The primary drivers behind the decision to locate an ethanol plant in a given county are access to feedstock and the absence of operating ethanol plants. Access to coproduct markets and transport infrastructure is also important. Some infrastructure variables in rural areas, such as farm product warehousing operations and road density, were important determinants for established plants from 2000 to 2007 but were not important for new plant announcements. This may be due to established ethanol plants obtaining prime areas that meet infrastructure needs leaving new plants to find areas with other means of transportation and storage as they occupy increasingly marginal locations. Local fiscal policy and state incentives also influence the location decisions of potential ethanol investors. There appears to be some potential with respect to recruiting ethanol plant investment in relatively rural areas. However extremely rural areas may be challenged with respect to providing any comparative advantage to prospective plants.

These findings are a first-step towards understanding the interaction between ethanol plant location and local factors that provide comparative advantage to counties considering ethanol plant recruitment as a development strategy. While the results appear encouraging for some rural areas, a rural county's access to or availability of important location determinants should be put in perspective before limited funds are invested in aims to attract ethanol plant investment.

While the model used in this research seeks to thoroughly address the determinants thought to dominate ethanol plant location decisions, there are certain limitations. For one, a majority of the factors used in the model are mainly input drivers and a more thorough exploration of demand proxies may be in order for future research efforts. Also, a gauge on ethanol plant capacity is lacking. Therefore the potential effects of plant scale on the location decision are not picked up. Other limitations include an absence of an appropriate variable for ethanol blending terminals, the lack of an appropriate variable to control for pasture-fed cattle as opposed to cattle most likely to be fed distillers' grains, as well as an appropriate variable to evaluate the importance of ethanol policies at the turn of the century. Future research efforts could potentially address some of these limitations.

Ethanol production is not a new technology, the recent flurry of activity in the ethanol market indicates the industry is still in its infancy. As profit margins and access to prime production locations disappear, there will be fewer entrants into the market. And as demand for fuel continues to grow and grain feedstock struggle to meet ethanol and food demands, the role of biomass and other feedstocks will become increasingly important. Future studies analyzing location determinants may shed some light on county comparative advantage and local capacity to attract new plant investment as conversion technologies advance, and alternative feedstock becomes available.

References

References

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

Anselin, L. 2003. *GeoDa 0.9 User's Guide*. Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign, Urbana, IL 61801, (accessed online 9/8/2008, http://sal.agecon.uiuc.edu).

Baker, A. and S. Zahniser. 2006. "Ethanol Reshapes the Corn Market." *Amber Waves*, April, 4(2):30-35.

Connor, J.M. and W. Schiek. 1997. *Food Processing: An Industrial Powerhouse in Transition*, 2nd ed. New York: Macmillan Co.

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

Dhuyvetter, K.C., T.L. Kastens, and M. Boland. "The U.S. Ethanol Industry: Where will it be Located in the Future." Agricultural Issues Center, University of California, November 2005.

Doering, O., 2006. "Ethanol and Energy Policy." Dept. Agr. Econ. ID-340, Purdue University.

Economic Research Service. 2000. "Farm Resource Regions," Economic Research Service-United States Department of Agriculture Agricultural Information Bulletin Number 760, 2 pp. (www.ers.usda.gov/publications/aib760/aib-760.pdf)

English, B.C., D. De La Torre Ugarte, M. Walsh, C. Hellwinkel, and J. Menard. 2006. "Economic Competitiveness of Bioenergy Production and Effects on Agriculture of the Southern Region." *Journal of Agricultural and Applied Economics* 38(2):389-402.

ESRI--GIS and Mapping Software. Internet Site: http://www.esri.com/ (Accessed February 9, 2007).

Goetz, S. 1997. "State and County-Level Determinants of Food Manufacturing Establishment Growth: 1987-1993." *American Journal of Agricultural Economics* 79(3):838-850.

Greene, W.H. 2003. *Econometric Analysis*. 2nd ed. New York: Macmillan Co.

Henderson, J.R. and K.T. McNamara. 1997. "Community Attributes Influencing Local Food Processing Growth in the U.S. Corn Belt." *Canadian Journal of Agricultural Economics* 25(2):680-697.

Henderson, J.R. and K.T. McNamara. 2000. "The Location of Food Manufacturing Plant Investments in Corn Belt Counties." *Journal of Agricultural and Resource Economics* 25(2):680-697.

Kelejian, H.H. and I.R. Prucha. 2007. "HAC Estimation in a Spatial Framework." *Journal of Econometrics* 140(1):131-154.

Kenkel, P. and R. B. Holcomb. 2006. "Challenges to Producer Ownership of Ethanol and Biodiesel Production Facilities." *Journal of Agricultural and Applied Economics* 38(2):369-375.

Kleinschmit, J. 2007. "Biofueling Rural Development: Making the Case for Linking Biofuel Production to Rural Revitalization." Carsey Institute. Policy Brief No. 5, University of New Hampshire.

Lambert, D., K. McNamara, and M. Garrett. 2006. "Food Industry Investment Flows: Implications for Rural Development." *Review of Regional Studies* 36(2):140-162.

Low, S. and A. Isserman. 2008. "Ethanol: Implications for Rural Communities." Paper presented at AAEA Annual Meetings, Orlando FL, 27-29 July.

Mas-Collel, A, M. D. Whinston, and J. R. Green, 1995. *Microeconometric Theory*, New York, Oxford University Press.

McNamara, K.T., W.P. Kriesel, and B.J. Deaton. 1988. "Manufacturing Location: The Impact of Human Capital Stocks and Flows," *Review of Regional Studies* 18(1):42-48.

Mielenz, J.R., D. Koepping, and F. Parson. 1996. "Commercialization of Biomass Ethanol Technology." *Applied Biochemistry and Biotechnology* 57/58:667-676.

National Agricultural Statistics Service (NASS), United States Department of Agriculture www.usda.gov/nass, 2000.

Newey, W.K. and K.D. West. 1987."A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55(3):703-708.

Novack, N., and J. Henderson. 2007. "Can Ethanol Power the Rural Economy*?" The*

Main Street Economist 2(1).

Office of Management and Budget-Standards for Defining Metropolitan and Micropolitan Statistical Areas; Notice. Internet site: http://www.whitehouse.gov/omb/fedreg/metroareas122700.pdf (Accessed July 27, 2007).

Parcell, J.L., and P. Westhoff. 2006. "Economic Effects of Biofuel Production on States and Rural Communities." *Journal of Agricultural and Applied Economics* 38(2):377-387.

Rainey, D.V. and K.T. McNamara. 1999. "Taxes and Location Decision of Manufacturing Establishments," *Review of Agricultural Economics* 21(1):86-98.

Renewable Fuels Association (RFA). Internet Site: http://www.ethanolrfa.org (Accessed January 3, 2008.

Roe, J.D., R. W. Jolly, and R. N. Wisner. 2006. "Another Plant?!...The Rapid Expansion in the Ethanol Industry and its Effect all the Way Down to the Farm Gate." Paper presented at AAEA annual meeting, Long Beach CA, 23-26 July.

Rose, M., E. R. Detch, and J. G. Morgan. 2005. "Ethanol and Biodiesel: Questions and Answers." Comptroller of the Treasury, Office of Research. Authorization No. 307328. Internet site: http://www.comptroller1.state.tn.us/repository/RE/final_alt_fuels.pdf (Accessed March 9, 2007).

Schmenner, R.W., J.C. Huber, and R.L. Cook.1987. "Geographic Differences and the Location of New Manufacturing Facilities." *Journal of Urban Economics* 21(1):83-104.

Shaffer, R., S. Deller, and D.W. Marcouiller. *Community Economics: Linking Theory and Practice*, 2nd ed. Oxford: Blackwell Publishing, 2004.

Shapouri, H., and P. Gallagher. 2005. *USDA's 2002 Ethanol Cost-of-Production Survey*. Washington DC: U.S. Department of Agriculture, ESCS for. Agr. Econ. Rep. 841, July.

Smith, E.D., B.J. Deaton, and D.R. Kelch. 1978. "Location Determinants of Manufacturing Industry in Rural Areas," *Southern Journal of Agricultural Economics* 10(1):23-32.

U.S. Department of Agriculture. 2007. *USDA Agricultural Projections to 2016*. Long-term Projections Report OCE-2007-1. Washington DC, February.

U.S. Department of Commerce, Bureau of the Census. 2000. *United States Census*

Bureau. Washington DC, June.

U.S. Department of Energy, Energy Information Administration. 2001. *Biofuels for the State, Helping the Economy and the Environment.* Washington DC, September.

Waldorf, B. 2006. "A Continuous Multi-dimensional Measure of Rurality: Moving Beyond Threshold Measures," Paper selected for the Annual Meetings of the Association of Agricultural Economics, Long Beach, CA, July. Available at: http://agecon.lib.umn.edu/cgi-bin/pdf_view.pl?paperid=21522&ftype=.pdf

Woodward, D.P. 1992. "Locational Determinants of Japanese Manufacturing Start-ups in the United States," *Southern Economics Journal* 58(3):690-708.

Wooldridge, J.M. 2006. *Introductory Econometrics*, 3rd. ed. Thomson South-Western.

Appendix

Table 1. Descriptive Statistics of Location Determinants

Variable		Mean	σ
Location Announcements (2000-2007)		0.020	0.140
Active Ethanol Plants (2000-2007)		0.035	0.183
FARMPROP	Farm proprietor income/nonfarm proprietor income (2000)	0.190	0.557
CATTLE	Cattle, plus surrounding counties (100,000s head)	2.007	2.045
CORN	Average total corn production plus surrounding counties (1990-2000) (100,000s bushels)	171.266	217.756
STORE	Farm product warehousing operations (Location Quotient) (2000)	2.117	13.963
NATGAS	Natural gas distribution centers (Location Quotient) (2000)	3.308	8.114
GAS	Gas stations, plus surrounding counties (2000)	6.894	3.429
ESTAB	Existing ethanol plant before 2000 (1 = yes)	0.010	0.106
HERFEMP	Employment concentration index, 2000 (between [0,1])	0.121	0.052
WAGE	Average wage per worker ($), 2000	12.307	2.761
HS00	% with highschool diploma, 2000	77.321	8.732
TRUCKLQ	Trucking companies (Location Quotient), 2000	2.076	1.897
ROAD	Road density (road miles/county area)	0.457	0.272
RAIL	Rail density (rail road miles/county area)	0.307	0.402
RIVER	River adjacency	0.326	0.469
FISC	Per capita income taxes/county expenditures, 2000	0.337	0.229
TAX	State excise tax incentive (2001) (1 = yes)	0.133	0.339
PRODCR	Ethanol producer credit program (2001) (1 = yes)	0.233	0.423
MTBE	Methyl tertiary-butyl ether ban, 2000 (1 = yes)	0.185	0.388
RI2000	Waldorf's (2006) 2000 rurality index (between [0,1])	0.501	0.177
HLAND	Heartland (1 = yes)	0.178	0.382
NOCRES	Northern Crescent (1 = yes)	0.138	0.345
FRUIT	Fruitful Rim (1 = yes)	0.091	0.288
NOGRTPL	Northern Great Plains (1 = yes)	0.058	0.235
PRGATE	Prairie Gateway (1 = yes)	0.128	0.334
BRANGE	Basin and Range (1 = yes)	0.064	0.245
MISSPORT	Mississippi Portal (1 = yes)	0.054	0.226
SOSEA	Southern Seaboard (1 = yes)	0.155	0.362

Table 2.1. Bivariate Probit Estimates, Main Effects

	Announced		Established	
Variable 1/	Estimate	T value 2/	Estimate	T value
CONSTANT	0.740	0.403	2.814	1.669
FARMPROP	2.388	3.135	2.186	3.228
CATTLE	-0.073	-1.132	0.137	2.303
CORN	0.001	2.163	0.001	2.971
STORE	-0.012	-1.261	-0.044	-1.954
NATGAS	0.028	1.374	-0.010	-0.326
GAS	-0.126	-2.438	0.126	2.753
ESTAB	-7.538	-3.046	0.431	0.588
HERFEMP	8.357	2.893	-2.174	-0.684
WAGE	-0.170	-3.132	0.035	1.015
HS00	-0.017	-0.806	-0.075	-3.833
TRUCKLQ	0.392	2.818	0.244	1.944
ROAD	-0.303	-0.488	-0.703	-1.385
RAIL	-1.008	-2.254	-0.589	-1.625
RIVER	0.857	3.681	-0.101	-0.468
FISC	-1.348	-1.907	0.982	1.644
TAX	-0.331	-0.931	0.161	0.570
PRODCR	-0.040	-0.119	-0.291	-1.032
MTBE	-0.083	-0.293	0.947	3.845
RI 2000	-12.565	-3.569	-13.924	-4.345
HLAND	0.633	4.207	0.895	5.512
NOCRES	0.314	1.911	0.670	3.927
FRUIT	0.443	2.580	0.316	1.702
NOGRTPL	0.707	3.067	1.145	5.652
PRGATE	0.788	4.420	0.704	4.056
BRANGE	-7.920	-6.620	0.528	2.465
MISSPORT	0.607	3.220	-6.023	-7.254
SOSEA	-5.899	-18.292	-0.076	-0.398

Note:
1/ Variable labels are in Table 1.
2/ T values 1.645, 1.961, and 2.577 significant at the 10%, 5%, and 1% level respectively.

Table 2.2. Bivariate Probit Estimates, Interaction of Location Determinants with the Rurality Index

	Announced ethanol plants		Established ethanol plants	
Variable 1/	Estimate	T value 2/	Estimate	T value
RI x FARMPROP	-3.109	-3.035	-2.677	-2.943
RI x CATTLE	0.226	1.991	-0.124	-1.190
RI x CORN	-0.004	-3.385	-0.002	-2.277
RI x STORE	0.025	1.343	0.073	2.080
RI x NATGAS	-0.028	-0.979	-0.010	-0.219
RI x GAS	0.149	2.155	-0.130	-2.126
RI x ESTAB	1.894	0.383	-1.564	-1.003
RI x HERFEMP	-9.107	-1.855	-6.747	-1.164
RI x WAGE	0.266	2.601	-0.119	-1.515
RI x HS00	0.106	2.558	0.178	4.644
RI x TRUCKLQ	-0.476	-2.182	-0.283	-1.522
RI x ROAD	-0.315	-0.212	2.695	2.245
RI x RAIL	4.918	4.230	2.224	2.243
RI x RIVER	-1.764	-3.578	-0.228	-0.530
RI x FISC	0.983	0.766	-2.802	-2.467
RI x TAX	1.266	1.967	0.149	0.297
RI x PRODCR	-0.166	-0.268	1.287	2.387
RI x MTBE	0.880	1.678	-0.899	-1.978
N	3,064			
Log likelihood	-563			
ρ	-0.989	4.82 3/		
Wald test: $\boldsymbol{\beta}_{ANN} = \boldsymbol{\beta}_{EST} = \mathbf{0}$		12,990 4/		
Pseudo R^2	0.258			

Note:
1/ Variable labels are in Table 1.
2/ T values 1.645, 1.961, and 2.577 significant at the 10%, 5%, and 1% level respectively.
3/ χ^2 variate with 1 degree of freedom. Critical value at the 5% level is 3.84.
4/ χ^2 variate with 90 degrees of freedom. Critical value at the 5% level is 113.15.

Table 3. Marginal Effects 1/

Plant announcements	Rurality index /2				
Variable	0.0 - 0.2	0.2 - 0.4	0.4 - 0.6	0.6 - 0.8	0.8 - 1.0
FARMPROP	0.002284	0.028299	0.006526	0.000270	-0.000104
CATTLE	-0.000056	-0.000111	0.000309	0.000108	0.000033
CORN	0.000001	0.000000	-0.000006	-0.000002	-0.000001
GAS	-0.000122	-0.001574	-0.000401	-0.000027	0.000002
HERF00	0.008190	0.109393	0.029786	0.002531	0.000041
WAGE	-0.000158	-0.001754	-0.000290	0.000021	0.000018
HS00	-0.000007	0.000291	0.000284	0.000074	0.000020
TRUCK	0.000379	0.004848	0.001207	0.000075	-0.000009
RAIL	-0.000568	0.009089	0.011362	0.003109	0.000864
RIVER	0.000748	0.006370	-0.000198	-0.000483	-0.000185
TAX	-0.000225	0.000944	0.002364	0.000709	0.000204
MTBE	0.000005	0.003510	0.002791	0.000680	0.000179
Active plants	Rurality index				
Variable	0.0 - 0.2	0.2 - 0.4	0.4 - 0.6	0.6 - 0.8	0.8 - 1.0
FARMPROP	0.009255	0.019703	0.007415	0.000376	-0.000039
CORN	0.000005	0.000010	0.000003	0.000000	0.000000
STORE	-0.000178	-0.000317	-0.000067	0.000008	0.000004
GAS	0.000544	0.001236	0.000532	0.000042	0.000002
HS00	-0.000275	-0.000306	0.000123	0.000060	0.000015
ROAD	-0.002093	0.001498	0.005634	0.001426	0.000299
RAIL	-0.001768	0.001116	0.004576	0.001166	0.000245
FISC	0.003386	0.002016	-0.003664	-0.001180	-0.000267
PRODCR	-0.000784	0.001351	0.003080	0.000735	0.000150
MTBE	0.004135	0.009649	0.004353	0.000383	0.000024

Notes:
1/ Marginal effects displayed here are significant at the 10% level, based on the bivariate probit results reported in Table 2.
2/ Rurality index quantiles.

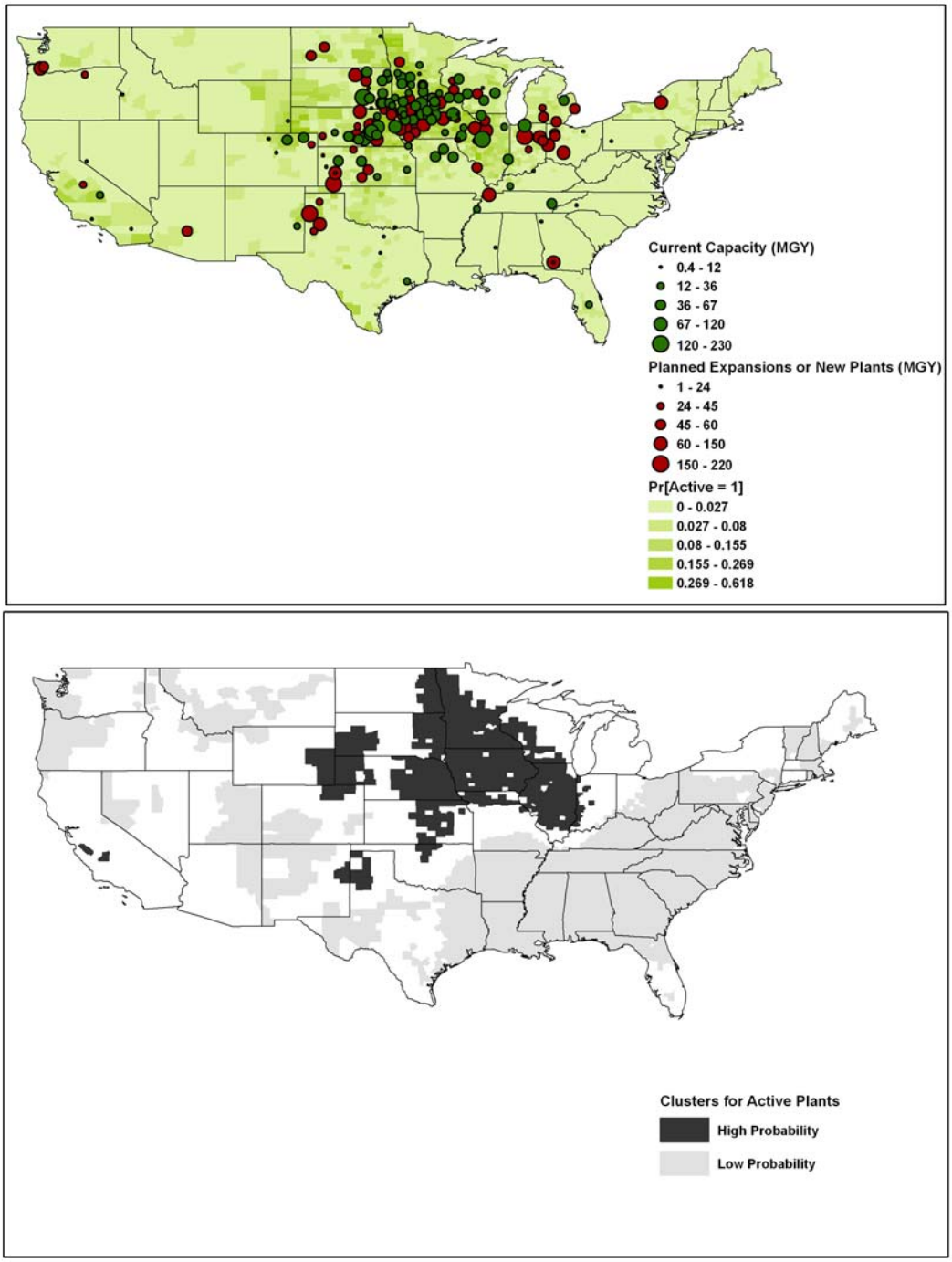


Figure 1. Top panel: spatial distribution of estimated location probabilities for active ethanol plants. Bottom panel: plant location probability clusters. High probability (HH) clusters are counties where the probability of attracting a potential investor is surrounded by other counties with high probabilities. Low probability (LL) clusters are counties where the probability of attracting a potential investor is surrounded by other counties with low probabilities.

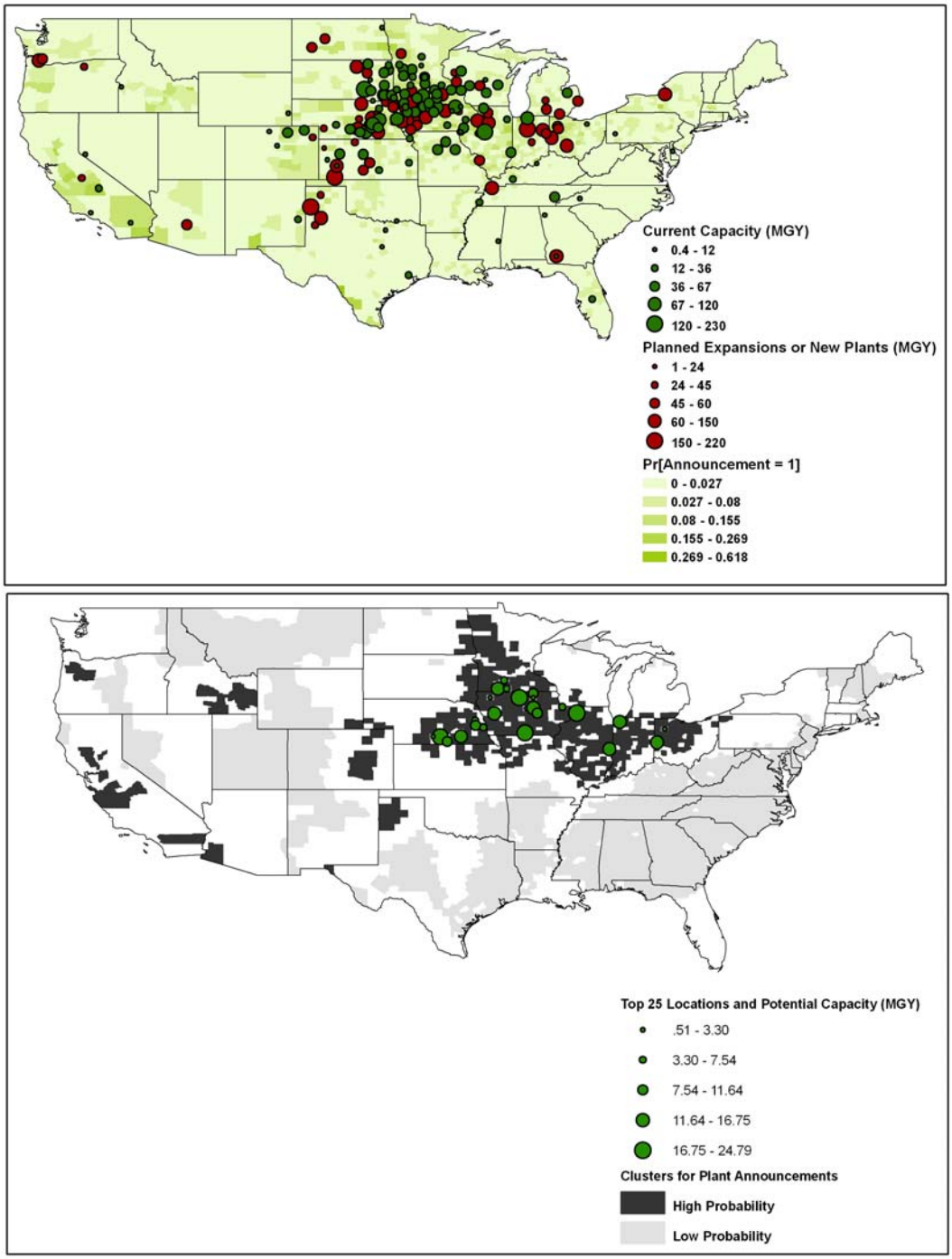


Figure 2. Top panel: spatial distribution of estimated location probabilities of ethanol plant location announcements. Bottom panel: plant location probability clusters. High probability (HH) clusters are counties where the probability of attracting a potential investor is surrounded by other counties with high probabilities. Low probability (LL) clusters are counties where the probability of attracting a potential investor is surrounded by other counties with low probabilities.

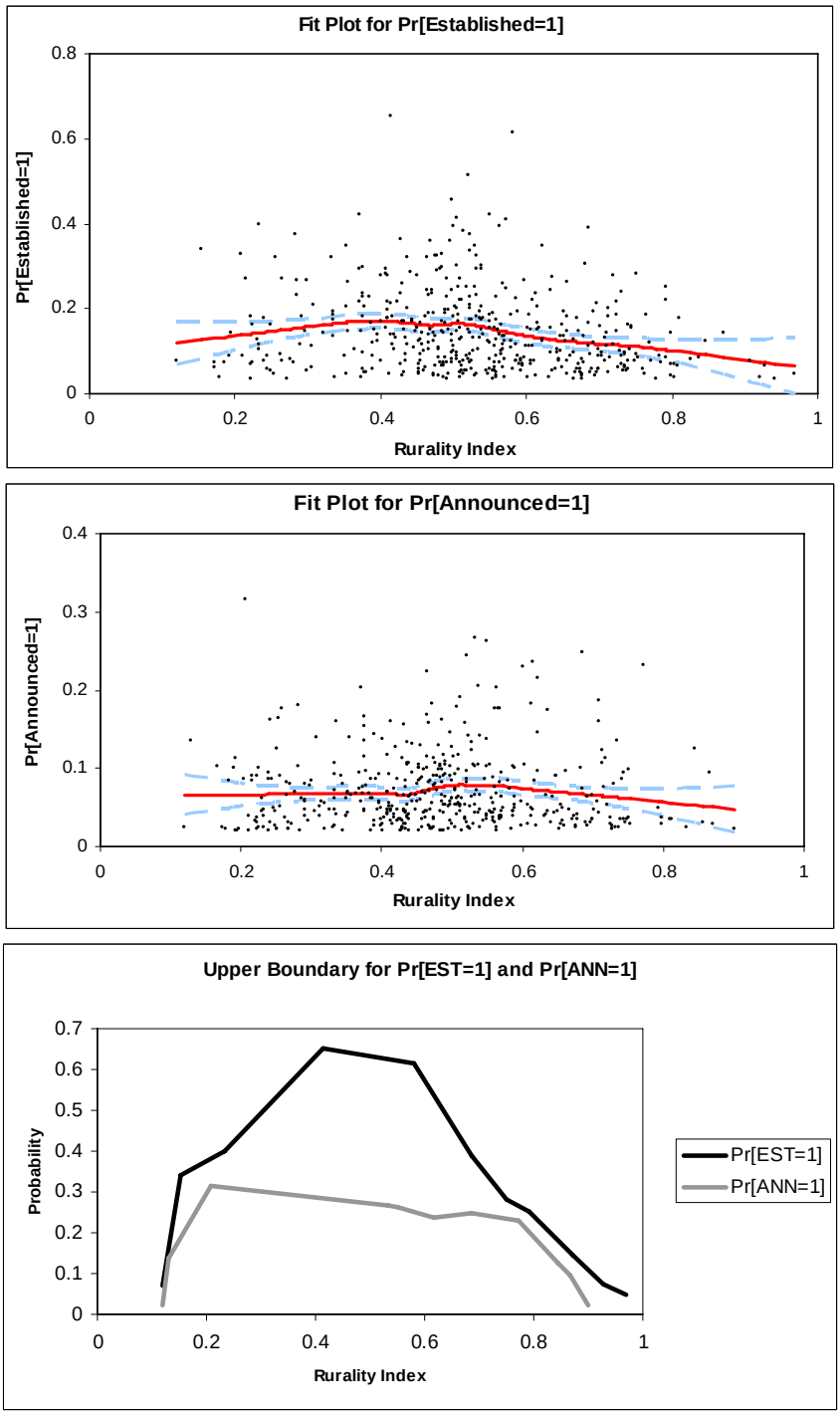


Figure 3. Upper panel; loess curve for clustered location probabilities associated with active plant locations and the rurality index. Middle panel; loess curve for clustered location probabilities associated with plant location announcement and the rurality index. Lower panel; upper boundaries of the probabilities associated with the active and announced plant locations, the plant location announcements, and the rurality index. Key: Solid line denotes the fit and the dotted line denotes the upper and lower confidence levels, $\alpha = 0.05$.

III: Identification of Tennessee Value Chains

Introduction

The U.S. economy weathered significant shocks in the past decade driven by technological innovations, an expanding global economy, industry restructuring, and competition from abroad. These changes may be attributed to several major macroeconomic events, one being the "dot-com" bubble of the 1990's (Kraay and Ventura 2005). Fueled by growth in computing power and the widespread use of the internet, U.S. financial markets expanded in the late 1990's and peaked in 2000. Due to investor optimism, technology stocks soared and money flowed out of real estate and into financial markets (Kraay and Ventura 2005). However, the turn of the century benchmarked the burst of the dot-com bubble, mounting energy costs, and increased attention towards global security; all of which ended with a sharp downward market adjustment. Financial markets went into a recession after poor earnings, stagnating inventories, and uncertainty (Kraay and Ventura 2005). The U.S. economy adjusted to staggering markets by lowering interest rates to record levels, which in turn, ignited the housing market boom that lasted well into 2006. By late 2003, there were some signs of economic recovery, but the economy had undergone a profound transformation.

During this period, manufacturing jobs continued to leave rural areas as firms outsourced labor abroad to minimize production and distribution costs. Employment in the manufacturing sector yielded to growth in retail, service, and construction jobs as rural communities struggled to adjust to the changing economy (Johnson 2001). Communities unable to adapt to these changes continued to experience population decline as employment opportunities decreased and economic growth stagnated

(McGranahan and Sullivan 2005). However, some rural communities appeared to weather market shocks by fostering entrepreneurship, marketing their natural amenities and cultural attractions, recruiting innovative value-added activities, or maintaining and expanding existing industry clusters (McGranahan and Wojan 2007).

Regional economists increasingly emphasize the identification of industry clusters given their contribution to rural economic growth and regional development (e.g., Barkley and Henry 1997; Porter 1998; Stimson, Stough, and Roberts 2006; St. John and Pouder 2006). This research applies the industry clusters concept and the value chains that define them to examine the transformation of the Tennessee economy amidst recent changes in the local, regional, and global economy. From 1997 to 2007, Tennessee lagged behind other states in terms of job and population growth, which led to a reduction in the labor force (CBER 2008). Identifying industry sectors that flourished in Tennessee during the 2001 to 2006 time period despite a tumultuous and uncertain economy provides insight to potential economic development strategies for some counties with regards to business retention and industry recruitment. The 2001-2006 period spans the brief recession in 2001 and the height of robust economic expansion in 2006. By identifying industry clusters that thrived during this time period, along with the value chains supporting them, it is possible to determine which counties were positioned to withstand these difficult economic times in terms of continued job growth, business retention or expansion, and growth in labor productivity in certain sector. For example, some industry clusters may have positively contributed to county or regional economic growth during this short business cycle. But in other instances, the presence of a given industry cluster may have had no effect, or even a negative impact, on local growth.

Locations where a given industry cluster persisted or expanded during 2001-2006, and contributed to economic growth, may suggest that a county has comparative advantage with respect to a given value chain in particular, or an industry cluster in general.

This research develops an econometric model to identify the sectors comprising an industry cluster in Tennessee counties, and the clusters subsequent impact on county economic growth. A three-step procedure accomplishes this goal. In the first step, a nonparametric cluster algorithm (Feser 2005; Feser and Isserman 2007; Feser, Sweeney, and Renski 2005) is applied to sector data for Tennessee's 95 counties to isolate value chains defining and supporting industry clusters. Sector level information comes from the Tennessee Impact Analysis for Planning (IMPLAN) data base (2001 and 2006) which includes information about employment, income, and value added information for 509 government and private industry sectors. In the second step, the extent to which a given value chain identified in 2001 contributed to labor productivity growth from 2001 to 2006 for each sector in every county is estimated using regression analysis. In the third step, the spatial distribution of the effect specific industry clusters had on labor productivity growth is analyzed for Tennessee's 95 counties.

Methods and Procedures

Conceptual Overview of Industry Clusters and Value Chains

The industry cluster concept is useful for understanding industry and small-business interconnections. Industry clusters are typically regarded as a mix of industries linked together both geographically and functionally that draw upon common suppliers making up an industry cluster (Figure 4). A cluster may be recognized by firms in related

industries seeking comparative advantage by locating near to natural resources, large markets, or labor pools (Gibbs and Bernat 1997). The source of competitive advantage for industry clusters is in local qualities, such as shared knowledge, relationships, and motivation, which are more difficult for distant competitors to obtain (St. John and Pouder 2006). Agglomerations of competing and collaborating industries in a region are arranged into horizontal and vertical relationships involving similar resource and/or labor needs, and rely on a shared formulation of specialized economic industries (Fujita and Thisse 2002). These cluster relationships allow industries to compete while promoting avenues of cooperation (Porter 1998).

Industry clusters affect competition by stimulating innovation, research, and development which supports future productivity growth and stimulates the formation of new businesses. These events, further stimulating another round of interaction, advance core industry sectors and strengthen the cluster itself (Porter 1998). Businesses within the clusters benefit by gaining greater access to suppliers and customized support service while building a framework for companies to work together meeting common needs and promoting common interests (CCED 2008). In many aspects, cluster members enjoy the benefits of scale economies without sacrificing autonomy (Porter 1998).

Because they are built around core export-oriented firms, industry clusters bring new wealth into a region and help drive regional economic growth (Stimson, Stough, and Roberts 2006; Barkley and Henry 1997; St. John and Pouder 2006). Business and retention programs endorsing the forward linkages (via supply) and backward linkages (via demand) of industry clusters will likely increase local multiplier effects (Shaffer, Deller, and Marcouiller 2004). These spatial concentrations of activity increase local

income by increasing factor productivity, decreasing production costs, and providing access to specialized goods and services (Johnson 2001). In other words, output will be higher for a given dollar of input due to agglomeration economies created by industry clusters (Gibbs and Bernat 1997). Gibbs and Bernat's (1997) analysis of rural industry clusters suggest that rural worker earnings are about 13% higher than those of comparable workers outside clusters. Through regional cooperation, groups of counties may sometime consider cluster recruitment as a mutual economic development strategy because of the network externalities resulting from business transactions and communications stemming from cluster formation across a region (CCED 2008).

An industry cluster's economic development potential provides an incentive to identify value chains in rural communities. Identifying value chains gives researchers the ability to determine if a particular industry cluster contributes to economic growth. Identifying value chains also provides policy makers, local planners, potential entrepreneurs, and industry leaders with information to make complex decisions regarding new business prospects, market potential, and supply acquisition. Distinguishing which value chains contribute to economic growth is also helpful for local planners with respect to allocating scarce resources to business retention or recruitment projects, or regional cooperation efforts where several local governments target similar development objectives.

Nonparametric Identification of Value Chains with Sector Transaction Data

The literature cites numerous methods for identifying and analyzing industry clusters. Qualitative approaches use techniques such as interviews, focus groups, and

surveys to learn about supply chain structures (Stimson, Stough, and Roberts 2006). Quantitative approaches typically analyze industrial sector data to measure industry size and change, as measured by employment, wage earnings, the number of establishments, and related dynamics (Stimson, Stough, and Roberts 2006). Quantitative methods include the analysis of input-output linkages and location quotients (Stimson, Stough, and Roberts 2006; Feser and Bergman 2000; Barkley and Henry 1997; Feser and Isserman 2007). Recent attention has also applied exploratory spatial data analysis techniques to identify and analyze the geographic distribution of regional clusters (Gibbs and Bernat 1997).

This study uses a nonparametric clustering procedure suggested by Feser and Bergman (2000), Feser, Sweeney, and Renski (2005), Feser (2005), and Feser and Isserman (2007) to identify Tennessee industry clusters in 2001. Feser and Bergman (2000) suggest a step-wise algorithm to identify industry clusters. Their approach uses input-output purchase and sales transaction data to determine significant inter-sector linkages that define mutually exclusive industry clusters. Recent studies Feser (2005) and by Feser and Isserman (2007) extended this approach by applying Ward's (1963) hierarchal clustering algorithm to construct a series of industry clusters based on a non-parametric matching procedure. Feser and Isserman's (2007) procedure identifies a reduced number of value chains that are essentially groups of industries with highly similar, and therefore, linked sectors. Sales and purchases of intermediate and final products between sectors determine linkages between sectors. An input–output transaction matrix of industry purchasing and sales patterns calculate pairwise correlations to compare industry linkages along four dimensions, thereby identifying

sectors whose economic linkages with each other are stronger than linkages with sectors outside of the group.

Define sets $S_{i(j)}$ and $B_{i(j)}$, where $S_{i(j)}$ is the set of supplier industries to industry $i(j)$ and $B_{i(j)}$ is the set of purchasing industries (buyers) from industry $i(j)$. Given S and B, we can define:

$$(1) \qquad I_{ij}^{SS} = S_i \cap S_j, \qquad U_{ij}^{SS} = S_i \cup S_j$$

$$I_{ij}^{BB} = B_i \cap B_j, \qquad U_{ij}^{BB} = B_i \cup B_j$$

$$I_{ij}^{SB} = S_i \cap B_j, \qquad U_{ij}^{SB} = S_i \cup B_j$$

$$I_{ij}^{BS} = B_i \cap S_j, \qquad U_{ij}^{BS} = B_i \cup S_j$$

where I and U are sets of buyer and/or supplier intersection and unions, respectively. From these relationships, the following measures are constructed;

$$(2) \qquad R_{ij}^{SS} = \frac{I_{ij}^{SS}}{U_{ij}^{SS}}, \quad R_{ij}^{BB} = \frac{I_{ij}^{BB}}{U_{ij}^{BB}}, \quad R_{ij}^{SB} = \frac{I_{ij}^{SB}}{U_{ij}^{SB}}, \quad \text{and } R_{ij}^{BS} = \frac{I_{ij}^{BS}}{U_{ij}^{BS}}$$

These ratios measure the strength of the linkage shared between industries i and j along four dimensions. R^{SS} is the number of supplier industries that industries i and j have in common over the total number of supplier industries to i and j. Larger values of R^{SS} indicate a stronger value chain linkage between i and j because of joint sourcing from the same suppliers. Likewise, R^{BB} is the share of common buyer industries. R^{SB} and R^{BS} measure second level relationships between sector pairs, and increase as one industry's suppliers become another industry's buyers. The maximum value of these measures is chosen to construct an n by n intersectoral linkage matrix, with (i,j) elements $R_{ij}^{MAX} = MAX\left(R_{ij}^{SS}, R_{ij}^{BB}, R_{ij}^{BS}, R_{ij}^{SB}\right)$.

Feser (2005) and Feser and Isserman (2007) suggest several adjustments for constructing $\mathbf{R}^{MAX}$, which we follow here. Some sectors are eliminated from the value chain analysis which includes sectors not present in any county, local consumer and personal services, retail trade, primary and secondary schools, and government enterprise sectors (Table 4). Also, following Feser and Isserman (2007), a linkage threshold of 0.02 was set for purchases and 0.01 for sales, which essentially requires that sector *j* must represent at least 2% of sector *i*'s total intermediate input purchases to be considered one of *i*'s key suppliers. Likewise, *j* must account for at least 1% of *i*'s intermediate sales to be considered one of *i*'s key buyers (Feser 2005). Once this step is completed, sectors such as wholesale trade, information, legal services, advertising, finance, and insurance are defined as "enabling industries". Enabling industries are assigned a weight of 0.01, which reduces their influence in the calculation of the R_{ij} measures from equation (2). This allows distinct or unique linkages between sectors to define value chains rather than the joint consumption of similar mixes of producer services, while still entirely including the linkages with producer services. The weighting scheme is admittedly ad hoc, but it moderates linkages between industries *i* and general enabling industry by the selected weight which serves to emphasize more concretely the linkage among more specialized industries (Feser 2005). In this application suggested by Feser and Isserman (2007), a sensitivity analysis was performed to determine the extent to which the weights influenced value chain identification, and which combinations produced the cluster with statistically strong connections (discussed below).

The next step of the analysis includes identifying sectors with few linkages with other sectors. These sectors are either insular (i.e. "with" themselves; sectors with no

purchases or sales to other sectors) or have connections with primarily local serving industries. In network analysis, these sectors are called "singletons", and were eliminated from the linkage matrix prior to applying Ward's cluster analysis (Ward 1963).

Ward's cluster algorithm produces the value chains in the final phase of the first step of this analysis. The reduced $\mathbf{R}^{MAX}$ matrix is converted to a "dissimilarity" (or "distance" matrix, **D**) matrix, and is used in Ward's clustering method. We used the simple distance measure $d_{ij} = \left(1 - R_{ij}^{MAX}\right)$ as the basis for the dissimilarity matrix. The larger d_{ij} is, the less closely related sales and purchases are between a sector, and the less likely sectors i and j will be included in the same cluster. One thing to note is that every industry in a county is linked at some level to every other industry, with the strength of that linkage ranging from $R_{ij}^{MAX} = 0$ (no joint buyers or suppliers) to $R_{ij}^{MAX} = 1$ (identical buyer and supplier linkages). Industries making up the clusters simply have stronger linkages than other identified groups, but it is still the case that any given industry may have reasonably tight linkages with other sectors defining specific value chains.

Feser (2005) suggests an *ad hoc Z* test as a cutoff point to determine cluster membership: $Z_{ij} = \dfrac{\bar{r}_{ij} - mean(\bar{r}_{ij})}{s.d.(\bar{r}_{ij})}$, with $\bar{r}_{ij}$ the average value of the maximum linkage between sector i and the set of primary sectors in core cluster j. The critical Z value was set to 1.96 (e.g., a 95% confidence interval). Lastly, it is important to note that cluster membership is not mutually exclusive. That is, a sector may belong to a value chain in more than one cluster. In this analysis, Tennessee input-output data from IMPLAN (year 2001) was used to benchmark value chains in the base year (2001) of the regression model.

There are several methods useful for determining how many clusters aptly describe a particular data set. The pseudo Hotelling's T^2 value is used to determine the optimal number of industry clusters following analysis of the linkage matrix with Ward's algorithm (Milligan and Cooper 1985).

Empirical Model Measuring the Influence of Clusters on Labor Productivity Growth

In the second step of the analysis, once industry clusters are identified based on 2001 transaction tables, we test whether a given cluster influenced labor productivity growth in a county from 2001-2006. The empirical model assumes a Cobb-Douglas production function that measures the relationship between local county level value chains, socio-demographic controls, input factors, and economic growth. This particular functional form has a long history in the economic growth literature because of its flexibility in terms of theoretical model development, interpretability, and application (e.g., Solow, 1954; Mankiw, Romer, and Neil, 1992; Benhabib and Spiegel, 1994).The function allows one to measure how capital per worker determines the amount of output per worker, and therefore how much output the economy produces for any given level of capital stock (Mankiw 2000). The Cobb-Douglas production function is,

(3) $$Y = tA^{\alpha_1} M^{\alpha_2} L^{\alpha_3} Z^{\alpha_4},$$

where Y is industry output, total factor productivity is measured by t, A represents total land area, M is a measure of capital, L is labor availability, Z are additional factors of production, and the $\boldsymbol{\alpha}$'s are elasticities of substitution. Dividing both sides of the production function by labor, and assuming that production is subject to stochastic shocks (u), yields the concentrated expression,

(4) $y = \bar{t} a^{\alpha_1} m^{\alpha_2} z^{\alpha_4} u$, with E(u) = 1.

Additional discrete or continuous factors thought to affect production (c) enter the equation as,

(5) $y = \bar{t} a^{\alpha_1} m^{\alpha_2} z^{\alpha_4} u \exp(c)$.

Effects of a given value chain on labor productivity growth at the county level are estimated by including county fixed effects in the exponential term in (5) (c_i = 1,…, 95). In addition, because the goal is to isolate the effects of a single value chain on economic growth (as measured by labor productivity), labor productivity stemming from the s = 1,…, 509 sectors is estimated simultaneously across all counties;

(6) $\Delta y_{is} = \bar{t} a_i^{\alpha_1} m_{i,2001}^{\alpha_2} z_{i,2001}^{\alpha_3} u_{is} \exp(c_i \delta + \gamma g_{2001}^{s} + \eta c_i g_{2001}^{s})$,

with g_{2001}^{s} a binary variable indicating that the sth sector in county i is a member of industry cluster g; $c_i g_{2001}^{s}$ an interaction term between value chain s and county i; and Δy_{is} denotes the change in labor productivity growth ($(Y_{is}^{2006} / L_{is}^{2006}) / (Y_{is}^{2001} / L_{is}^{2001})$). Anywhere a sector appears in a county, and that sector is a member of the industry cluster under evaluation, g_{2001}^{s} = 1, otherwise, g_{2001}^{s} = 0. The parameter η measures the extent to which industry cluster g contributed to average labor productivity growth in county i. Together, the terms bracketed inside the exponential are fixed effects and are essentially a composite error term (Wooldridge 2006). The right hand side variables in equation (6) are lagged, denoted by the 2001 subscripts. An orthogonal restriction is applied to the county fixed effects (c_i) such that $\sum_i \delta_i = 0$ (Neter, et al. 1998). The restriction shifts the distribution of the fixed effects away from a reference group where inference of the fixed effect coefficients is with respect to the average growth of a particular county (i.e., the

usual "compare all groups to one location" specification) to one where county growth is compared to the overall average growth in labor productivity for the state.

Estimation is facilitated by taking the natural logarithm of (6), which yields the log-log linear form of the production function,

$$(7) \qquad \ln \Delta y_{ik,2001}^{2006} = t^* + \mathbf{x}'_{ik,2001}\boldsymbol{\alpha} + \gamma g_{2001}^{s} + \delta_i c_i + \eta_i g_{2001}^{s} c_i + \bar{u}_{ik} \ ,$$

where **X** is a matrix containing the natural logs of the continuous exogenous production factors, $\bar{u}$ is an independent and identically distributed shock with $E(\bar{u}_{ik}) = 0$ and $Var(\bar{u}_{ik}) = \sigma_{ik}^2$.

Estimation Methods

Ideally, Ordinary Least Squares (OLS) would be used to estimate equation 7, and standard error would be estimated with a heteroskedastic – robust covariance estimator. The regression includes all sectors (n = 48260), but some sectors may have never existed in a county in 2001 or 2006 (e.g., value added dollars and employment was zero). Alternatively, a given sector in a county may have existed in 2001, but exited the county in 2006. In addition, a sector may have not been present in 2001, but entered in 2006. Given the log-log linear specification, this makes the calculation of labor productivity impossible between 2001 and 2006 for some sectors. However, simply eliminating these sectors with zero output during one or both of the periods and focusing only on the sectors that existed in 2001 and 2006 (n = 15221) may bias OLS estimates. If the pattern of missing sectors is random, then selection bias is not an issue. However, it is entirely possible that such a pattern in the data set is not random due to imperfect competition, market power, or other unobserved local factors systematically influencing firm entry and

exit. To attend to this potential selection bias, we estimate the system using Heckman's (1979) sample selection model. A significant correlation coefficient between the disturbance terms of the selection and outcome equations is evidence that sample selection bias exists, the observed pattern of missing sectors is not random, and the properties of the OLS estimator are compromised. Standard errors were calculated with White's (1980) heteroskedastic robust covariance matrix.

Spatial Clustering of County by Value Chain Coefficients

The use of spatial statistical techniques is useful for analyzing spatially clustered data, dispersed data, outliers, and spatial dependency. Local Indices of Spatial Association (LISA) are estimated to measure the extent to which locations share similar growth in labor productivity over a geographic region, conditional on the presence of a value chain (Anselin 1995). LISAs are a local adaptation of Moran's I (1950), which focuses on the local variations within patterns of spatial dependency. LISAs measure the degree to which a target value is similar to the values displayed by adjacent locations and also measures the similarity between each target value and those of its spatially defined neighbors (Anselin 1995). In the *ex post* analysis, the coefficients of the value chain by county dummy variable (the η_i's) are analyzed using this procedure to identify regional linkages between value chains and value chain membership. Such information may be useful for local policy makers seeking to coordinate regional development planning efforts with other counties.

Data

Tennessee IMPLAN (2001) gross absorption coefficients (GACs) were used to construct the input-output transaction matrix, which served as the base for the industry clustering procedure. The GACs contain information for 509 intersectoral purchases and sales. Sector employment information was also gathered from the 2001 and 2006 IMPLAN files, as well as employee compensation, and the value added for each sector. Eliminating local retail and personal serving sectors, singletons, and government industries reduced the transactions matrix (**R**) to a 447 by 447 dissimilarity matrix (**D**) evaluated by Ward's cluster algorithm.

The tax assessments for the state of Tennessee came from the 2001 Tax Aggregate Report of Tennessee (State Board of Equalization 2001). While clearly an imperfect measure, this variable was used to proxy local capital endowment. Additional control variables included county area (square miles), the percent of manufacturing, agriculture, business service establishments in 2000 (all from the U.S. Census County Business Pattern files), median household income, and the percent of workers commuting to another county (in 2000). Land area, total sector wages, value added, and total county property tax assessment in county *i* was normalized by the number of workers in each sector *s*. All variables are in natural logarithms except the variables entering the model in percent form or as fixed effects. Descriptive statistics of control variables are in Table 5. After eliminating observations with missing data, the full sample size used in the regression analysis was 48,260, with 15,221 observations where value-added and employment in a given sector was observed in 2001 and 2006.

Results

Discussion of the results focuses on five industry clusters; agricultural, forestry, new building construction, textiles, and electronics manufacturing, due to the strength of the connections between member sectors as well as the influence these clusters have on the Tennessee economy. Agricultural and forestry industries are expected to be influenced by the growing biofuels industry. As ethanol production and corn scarcity increases, alternative biofuels sources are emerging as potential substitutes for feedstock. U.S. corn supplies are struggling to keep pace with ethanol demand and future predictions indicate that ethanol will likely be supported manufactured by wood (i.e. poplar, a fast growing tree that can be grown as a row crop on fallow farmland) and/or switchgrass, a native perennial grass as feedstock (De La Torre Ugarte et al. 2006). These dedicated energy crops are expected to have a considerable impact on agricultural and forestry industries, as some traditional food and fiber farmlands convert to growing feedstock for fuel. In 2001 (2006), the Tennessee's Agriculture Sector employed 304,880 (311,711) persons, and produced 27,763 (35,018) millions of dollars in value added output (Table 6). Value added output from the forestry sector increased 33% from 5,548 millions of dollars in 2001, but employment in this sector slightly decreased from 21,237 employees in 2001 to 20,987 employees in 2006 (Table 7). The new building construction cluster reflects the impact of the housing boom that occurred at the beginning of the century. Value added output in the new building construction industry increased significantly by 43% from 24,286 millions of dollars in 2001. Furthermore, in 2006 the new building construction industry added 7,255 employees to its labor force, increasing the number of employees from 274,640 in 2001 to 281,895 in 2006 (Table 8).

The textile industry is expected to be influenced by the exodus of manufacturing jobs to least-cost production sites. Value added output for the textiles sector fell from 3,893 millions of dollars in 2001 to 3,524 millions of dollars in 2006. Employment in the textiles sector dropped significantly, employing 32,574 workers in 2001 and only employing 21,323 workers in 2006 (Table 9). The electronics manufacturing sector may be insulated from the exit trends of other manufacturing sectors. This may be, in part, due to the fact that in production, the most significant economic forces are the rising importance of information, communication, robotics, and other types of technology (Johnson 2001). Value added output in the electronics manufacturing industry increased from 5,554.11 millions of dollars in 2001 to 9,436.34 millions of dollars in 2006. However, employment fell from 41, 799 employees in 2001 to 32,832 in 2006 (Table 10).

In general, the *Z* scores in the agricultural, forestry, new building construction, textiles, and electronics manufacturing clusters indicate strong relationships within the representative clusters, but the linkage values between the groups are relatively low compared to Feser (2005) and Feser and Isserman's (2007) study at the national level. The relatively lower values are expected since the focus is on the state and county level, and some regional sector linkages are lost. Discussion of the results focuses on each of the three steps used to identify value chains and how they impacted economic growth.

Step One: Industry Cluster Identification with 2001 Sector Transaction Data

According to the pseudo T^2 statistic, Figure 5 suggests three candidate clusters at 19, 43, and 46. Clusters were too aggregated using 19 as a membership value, making it difficult to discern any particular relational patterns between sectors. The second and

third candidates – 43 and 46 numbers of clusters – produced memberships that were easily discernable with relatively strong linkages and significant *Z* scores. The lower, more conservative value (43 clusters) was used in the analysis.[2]

There were 9 primary sectors and 18 supporting sectors identifying what were termed the agricultural cluster (Table 6). The primary sectors are consistent with what one might typically consider "agricultural activities" (e.g., cotton farming, grain farming, and tobacco farming). The supporting industries include sectors closely affiliated with farming (e.g., veterinary services, agriculture and forestry support services, and animal production) and farm by–products (waste management, logging and refractory mining [from conversion of forest to farmland, or farmland to other uses], various transport services [rail and ground transportation], hunting and trapping [a positive amenity of agriculture]). Potentially less obvious relations include insurance carriers (possibly for crop and livestock insurance) and timber and forest products. Some surprising members include hospitals (possibly due to the risks associated with agriculture, or perhaps identifying a linkage with the veterinary sector), air transportation (perhaps catering to some perishable agricultural products, seeds, or animal semen delivered from/to far away locations), and pipeline transportation (perhaps associated with natural gas or petro-chemical distribution).

The forestry cluster (FS) (Table 7) contained fewer members, with 8 primary members and 4 secondary sectors. Not surprisingly, forest nurseries, forest products, and timber sectors stand out as the forestry cluster core. Logging and sawmill industries represented the next strongest membership linkages. Secondary sectors included pulp

[2] Feser and Isserman (2007) identified 45 clusters in their national analysis.

mills, paper mills, wood window frames and door manufacturing, and reconstituted wood product manufacturing.

The new building construction cluster (Table 8) contained 14 primary sectors and 3 secondary sectors. The cluster core is somewhat broadly defined by sectors such as new residential structures, commercial and institutional buildings, and architectural and engineering services. Sectors such as water, sewer, and pipeline construction, road construction, and telecommunications also make up a significant portion of the cluster core, possibly representing the infrastructure needs that are likely needed to support the construction of new building structures. Secondary sectors are made up of agriculture and forestry support activities and environmental consulting which are likely due to the conversion of farm land and forest to residential and commercial areas. Pipeline transportation is the third support sector which is likely an additional infrastructure support need.

The textiles cluster (Table 9) is made up of 14 primary sectors with no identified secondary sectors. The sectors that define the textiles cluster consist mainly of clothing mills, carpet mills, curtain mills, and other various fabric mills that one might typically associate with textile products. Other members closely affiliated with textile activities are fabric coating mills and upholstered household furniture manufacturing.

The electronics manufacturing cluster (Table 10) contained 14 primary members and 9 secondary members. The core of the electronics manufacturing cluster is made up of a variety of different electronic component production activities (e.g., communications equipment, electronic computers, photographic equipment, and navigation instruments). The secondary sectors that represented weaker membership linkages consist of

semiconductor manufacturing, relay manufacturing, and miscellaneous electrical equipment manufacturing which may provide the raw components needed for the manufacturing of products in the core of the cluster.

Step Two: Where Did These Clusters Contribute to Growth in Labor Productivity?

The effect agricultural, forestry, new building construction, textiles, and electronics manufacturing clusters had on labor productivity growth at the county level was estimated with equation 7. We report the results from the Heckman sample selection model because the null hypothesis that the disturbance terms between the selection and growth equations was strongly rejected for all 5 models (Tables 11, 12, 13, 14, and 15). In terms of identifying the effect of industry clusters on labor productivity growth, the relevant estimates are the interaction terms of the county fixed effects and the industry cluster dummies, so the relationship between the control variables and labor productivity growth is discussed briefly.

Probit Selection Equation and Control Variables

About 32% of the observations (15, 221) were included in the growth equation. Land area per worker was negatively correlated with the likelihood of a sector existing in 2001 and 2006 for each industry cluster regression. Less densely populated areas were less likely to have a firm (or firms) locate there, given that population and employment are highly correlated. Higher worker wages increased the likelihood a sector in a given county was observed in 2001 and 2006. There is limited evidence that counties with higher property taxes were less likely to have any given sector operating there in 2001 or 2006. The more persons commuting into or out of a county in 2000, the more likely that

county retained a business in 2001 and 2006. This may seem contrary to conventional logic: the more persons commuting to another county to work, the smaller the potential labor force in the county that they originated from, unless there is significant commuter back flow. On the other hand, it may simply reflect a dearth of jobs in a given location. However, it seems likely that commuting workers attract certain sectors and support elements for suburban developments such as, retail and personal service industries or new housing construction.

Labor Productivity Growth Equations and Control Variables

Total land area/worker in the clusters analyzed was negatively correlated with labor productivity growth, suggesting that more remote places grew more slowly between 2001 and 2006 with respect to labor productivity. Property tax assessment per worker was positively associated with labor productivity growth, but growth was slower when wages were higher. The outcome measure is value added per worker, and essentially represents sector profits. Higher labor costs lower profit margins, assuming the same level of productivity. Median household income was positively associated with labor productivity growth, but the more business services available in a county, the slower was growth in labor productivity. The percent of workers commuting was negatively correlated with labor productivity growth. Workers leaving for jobs in other counties may reduce the available labor pool, making it difficult to find skilled labor for a variety of positions.

Step Three: Analysis of County by Industry Cluster Interactions on Labor Productivity Growth

On average, labor productivity growth in the agriculture cluster was 68% slower compared to counties absent in this industry cluster (Table 11, 68% = exp(-1.16) – 1, [Wooldridge 2006]). In contrast, the change in labor productivity was 78% higher in the forestry sector compared to the rest of the economy (Table 12). Likewise, labor productivity grew by 46% in counties present in the new building construction cluster (Table 13). For counties with members of the textile cluster, labor productivity fell by 66% (Table 14). The most significant growth in labor productivity of the five clusters examined was exhibited by the electronics manufacturing cluster which grew by 532% for counties present in this industry cluster (Table 15).

Turning attention to the interaction between the county fixed effects and the industry clusters, about 17% of the counties with one or more sectors belonging to the agricultural cluster experienced growth in labor productivity from 2001 – 2006 (Figure 6.1). Only two counties exhibited negative growth with respect to the agricultural cluster. The interaction term between the county fixed effects and the forestry cluster was positive in roughly 44% of Tennessee's counties, with only one county experiencing negative growth in labor productivity with respect to this sector (Figure 6.1). Only five Tennessee counties belonging to the new building industry cluster proved to be statistically significant, with two of those counties exhibiting positive growth and the other three counties exhibiting negative growth (Figure 6.2). As expected, the interaction term between county fixed effects and the textile industry cluster exhibited a larger proportion of negative growth than positive growth in Tennessee counties between 2001

and 2006. About 19% of Tennessee counties experienced negative growth in labor productivity within the textile sector, while only roughly 12% of counties experiencing positive growth (Figure 6.2). Surprisingly, roughly 27% of Tennessee counties experienced a decline in labor productivity with respect to the electronics manufacturing industry cluster while only three counties experienced growth in labor productivity (Figure 6.3).

Local indices of spatial association (LISA) were estimated using the interaction coefficients (η_i's) from the agriculture, forestry, new building construction, textile, and electronics manufacturing cluster regressions (Figure 7). Counties highlighted in red are those with relatively high cluster-by-county interaction coefficients, or counties where the industry cluster positively influenced labor productivity growth between 2001 and 2006. A regional agricultural cluster is evident in the northwestern corner of the state. This is not too surprising because this region produces most of Tennessee's soybean and corn (Figure 7.1). There appears to be a strong but small spatial cluster of positive county–by–forestry cluster interaction terms in the western most portion of middle Tennessee (Figure 7.1). Within the new building construction cluster, two counties in middle Tennessee (possibly reflecting suburban development outside of Nashville) and one county in east Tennessee positively influenced labor productivity growth between 2001 and 2006 (Figure 7.2). There are no strong spatial clusters of counties where the interaction between the textile industry cluster and the county fixed effect is relatively high, which is likely representative of the decline in the textiles industry and other manufacturing industries (Figure 7.2). Likewise, the electronic manufacturing industry

cluster showed no strong spatial clusters of counties with relatively high cluster-by-county interaction coefficients (Figure 7.3).

The spatial cluster patterns suggest that counties belonging to these respective clusters might benefit from regional coordination with respect to focusing limited resources on business recruitment or retention strategies targeting sectors belonging to clusters that experienced positive growth between 2001 and 2006. In other cases where an industry cluster positively influenced labor productivity growth during this time period, but did not belong to a highly productive "spatial cluster", counties may also consider focusing scarce resources to develop and enhance businesses who are members of these sectors, but with less emphasis on trying to establish channels of regional cooperation with counties.

Summary and Conclusions

This research proposed a three step approach for identifying industry clusters, determining how these clusters influence economic growth at the local and regional level, and the extent to which industry clusters are integrated over a wider geographic region. The first step applied a recently developed nonparametric approach to identify industry clusters based on input–output transactions. The goal was to identify industry clusters in Tennessee (in 2001), and how these clusters influenced labor productivity growth between 2001 and 2006. The second step used regression analysis to gauge the performance of counties hosting one or more sectors of an industry cluster. The third step examined the spatial distribution of counties whose economies expanded (or contracted), given presence of one or more firms belonging to a given industry cluster. County and

regional comparative advantage was determined by testing whether the presence of a particular value chain in a given county contributed to increased labor productivity during this period, controlling for other local determinants of economic performance.

Identification of industry clusters is an important component of regional economic development strategies. By determining the sector composition of industry clusters, more accurate information regarding recruitment of traditional value-added activities (such as manufacturing) or incubation of innovative endeavors (such as niche agricultural marketing or biotech firms) is possible when a community or region is able to identify which industries make up a cluster, the forward and backward linkages of purchases and sales between cluster members, and the degree to which industry costs are minimized through cluster membership.

While the methods used to identify and evaluate industry clusters and their influence on economic growth proved to be successful, certain limitations remain. First, regional "trading areas" could have been developed to more accurately assess the forward and backward linkages between regions rather than relying on county data. Other practical possibilities include applying this methodology to include a broader regional analysis. Tennessee's economy is regionally, nationally, and internationally link with other economies. Wider regional coverage may produce a more detailed county-level analysis and may account for some economic leakage at the county level. Lastly, while multiple benefits exists within cluster-based analysis, industry clusters are not necessarily the cure all for regional development, and a one-size-fits-all strategy is unlikely to work for all regions. Industry clusters must be studied and reevaluated to gain the full benefits of co-location.

References

References

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

Barkley, D. and M. Henry. 1997. "Rural Industrial Development: To Cluster or Not To Cluster?" *Review of Agricultural Economics* 19(2):308-325.

Benhabib, J. and M. Spiegel. 1994. "The role of human capital in economic development: evidence from aggregate cross-country data." *Journal of Monetary Economics* 34(2):143–174.

Center for Business and Economic Research (CBER). 2008. "An Economic Report to the Governor of the State of Tennessee." College of Business Administration, University of Tennessee.

Center for Community and Economic Development (CCED). 2008. "2007-2008 Building Communities Educational Series: Clusters and Regions-They Both Matter." University of Wisconsin Extension.

De La Torre Ugarte, D., B. English, K. Jensen, C. Hellwinkel, J. Menard, and B. Wilson. 2006. "Agricultural Impacts of Ethanol and Biodiesel Expansion." Study Report, Dept. of Agr. Econ., University of Tennessee.

Feser, E. 2005. "Benchmark Value Chain Industry Clusters for Applied Regional Research." *Regional Economics Application Laboratory*. University of Illinois at Urbana-Champagne.

Feser, E.J. and E.M. Bergman. 2000. "National Industry Cluster Templates: A Framework for Applied Regional Cluster Analysis." *Regional Studies* 34(1):1-19.

Feser, E. and A. Isserman. "The Rural Role in National Value Chains and Regional Clusters." Working paper, Spatial Policy Analysis Research Consortium, University of Illinois, 2007.

Feser, E., S. Sweeney, and H. Renski. 2005. "A Descriptive Analysis of Discrete U.S. Industrial Complexes." *Journal of Regional Science* 45(2):395-419.

Fujita, M. and J. Thisse. 2002. *The Economics of Agglomeration*. Cambridge: Cambridge University Press.

Gibbs, R. M., and G. A. Bernat, Jr. 1997. "Rural Industry Clusters Raise Local Earnings." *Rural Development Perspectives* 12(3):18–25.

Heckman, J.J. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica* 74(1):153-161.

Johnson, T.G. 2001. “The Rural Economy in a New Century.” *International Regional Science Review* 24(1):21-37.

Kraay, A. and J. Ventura. 2005. “The Dot-Com Bubble, the Bush Deficits and the U.S. Current Account.” Working Paper, National Bureau of Economic Research, Cambridge, Massachusetts.

Mankiw, N.G. 2000. *Macro Economics*, 4th ed. New York: Worth Publishers.

Mankiw, N.G., D. Romer, and D.N. Neil. 1992. “A Contribution to the Empirics of Economic Growth.” *The Quarterly Journal of Economics* 107(2):407-437.

McGranahan, D. and P. Sullivan. 2005. “Farm Programs, Natural Amenities, and Rural Development.” *Amber Waves*, February 3(1):28-35.

McGranahan, D. and T. Wojan. 2007. “Recasting the Creative Class to Examine Growth Processes in Rural and Urban Counties.” *Regional Studies* 41(2):197-216.

Milligan, G.W. and Cooper, M.C. 1985. “An Examination of Procedures for Determining the Number of Clusters in a Data Set,” *Psychometrika* 50:159 - 179.

Moran, P.A.P. 1950. “Notes on Continuous Stochastic Phenomena.” *Biometrika* 37:17-23.

Neter, J., M.H. Kutner, C.J. Nachtsheim, and W. Wasserman. 1998. *Applied Linear Statistic Models*, 4th ed. New York: McGraw Hill Book Co.

Porter, M. E. 1998. “Clusters and the New Economics of Competition.” *Harvard Business Review*, 77-90.

Shaffer, R., S. Deller, and D. Marcouiller. 2004. *Community economics: Linking theory to practice*. Oxford, UK: Basil Blackwell.

Solow, R. 1954. “A Contribution to the Theory of Economic Growth.” *The Quarterly Journal of Economics* 70(1):65–94.

State Board of Equalization. 2001. *2001 Tax Aggregate Report of Tennessee*. Tennessee.

Stimson, R.J., R.R. Stough, and B.H. Roberts. 2006. *Regional Economic Development*, 2nd ed. New York: Springer Berlin Heidelberg.

St. John, C. H. and R. W. Pouder. 2006. "Technology Clusters versus Industry Clusters: Resources, Networks, and Regional Advantages." *Growth and Change*, June 37(2):141-171.

U.S. Department of Commerce, Bureau of the Census. 2000. *United States Census Bureau*. Washington DC, June.

Ward, J.H. 1963. "Hierarchical Grouping to Maximize an Objective Function." *Journal of the American Statistical Association* 58(3):236-244.

White, H. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48:817-838.

Wooldridge, J.M. 2006. *Introductory Econometrics*, 3rd. ed. Thomson South-Western.

Appendix

Table 4. Singleton, Local Serving, and Government Industrial Enterprise Sectors Excluded from the Value Chain Analysis, Following Feser and Isserman (2007).

Industry Code	Description	Industry Code	Description
16	Fishing	470	Social assistance- except child day care services
21	Iron ore mining	476	Fitness and recreational sports centers
23	Gold- silver- and other metal ore mining	477	Bowling centers
49	Rice milling	481	Food services and drinking places
50	Malt mfg.	482	Car washes
56	Sugar mfg.	483	Automotive repair and maintenance- except car wash
78	Roasted nuts and peanut butter mfg.	486	Household goods repair and maintenance
81	Flavoring syrup and concentrate mfg.	487	Personal care services
90	Cigarette mfg.	488	Death care services
127	Flexible packaging foil mfg.	489	Drycleaning and laundry services
208	Alumina refining	490	Other personal services
214	Primary smelting and refining of copper	491	Religious organizations
303	Computer storage device mfg.	495	Federal electric utilities
310	Electron tube mfg.	496	Other Federal Government enterprises
331	Household laundry equipment mfg.	497	State and local government passenger transit
348	Motor home mfg.	498	State and local government electric utilities
354	Guided missile and space vehicle mfg.	499	Other State and local government enterprises
355	Propulsion units and parts for space vehicles	500	Noncomparable imports
356	Railroad rolling stock mfg.	501	Scrap
360	Military armored vehicles and tank parts mfg.	502	Used and secondhand goods
412	Nonstore retailers	503	State & Local Education
432	Automotive equipment rental and leasing	504	State & Local Non-Education
433	Video tape and disc rental	505	Federal Military
435	General and consumer goods rental except video tapes	506	Federal Non-Military
461	Elementary and secondary schools	507	Rest of the world adjustment to final uses
469	Child day care services	509	Owner-occupied dwellings

Table 5. Descriptive Statistics of Control Variables

Variable	Mean	Standard Error
Total land area/worker in sector k	-3.559	1.032
Total property tax assessment/worker in sector k	9.753	0.311
Employee compensation in sector k /worker in sector k	2.601	4.431
% manufacturing establishments in 2000	26.545	7.198
% agriculture and forestry establishments in 2000	2.558	1.548
% business services in 2000	7.045	3.287
Median household income in 2000	10.372	0.190
% commuting workers in 2000	39.400	15.506

Table 6. 2001 Agriculture Cluster

Description	IMPLAN Code	Type	Z-Score	Linkage
Agriculture and forestry support activities	18	Secondary	3.00992	0.17201
Logging	14	Secondary	1.99199	0.18758
Forest nurseries, forest products, and timber	15	Secondary	2.86202	0.13959
Hunting and trapping	17	Secondary	5.48808	0.12196
Waste management and remediation services	460	Secondary	2.06114	0.18758
Other accommodations	480	Secondary	2.21895	0.16627
Other ambulatory health care services	466	Secondary	2.44403	0.15406
Hospitals	467	Secondary	2.59739	0.17135
Veterinary services	449	Secondary	3.53534	0.20720
Animal production, except cattle and poultry	13	Secondary	3.68751	0.26751
Cattle ranching and farming	11	Secondary	3.95177	0.28572
Pipeline transportation	396	Secondary	2.18919	0.14524
Rail transportation	392	Secondary	2.11774	0.10362
Air transportation	391	Secondary	2.1719	0.10362
Transit and ground passenger transportation	395	Secondary	2.8029	0.20720
Insurance carriers	427	Secondary	2.04378	0.13404
Insurance agencies, brokerages, and related	428	Secondary	2.72779	0.04745
Sand, gravel, clay, and refractory mining	25	Secondary	2.07674	0.10389
Greenhouse and nursery production	6	Primary	3.96090	0.26050
Cotton farming	8	Primary	5.01410	0.36356
Grain farming	2	Primary	5.02075	0.34246
All other crop farming	10	Primary	5.08411	0.32293
Tree nut farming	4	Primary	5.12422	0.36337
Oilseed farming	1	Primary	5.14225	0.34059
Vegetable and melon farming	3	Primary	5.14993	0.31769
Tobacco farming	7	Primary	5.23144	0.36408
Fruit farming	5	Primary	5.27465	0.36408

Variable	*Sum total*	
output_2001 (Value added output, 2001)	27,763.16	(millions $)
output_2006 (Value added output, 2006)	35,018.54	(millions $)
imp_emp_2001 (Employment, 2001)	304,880	
imp_emp_2006 (Employment, 2006)	311,711	

Table 7. 2001 Forestry Cluster

Description	IMPLAN Code	Type	Z-Score	Linkage
Reconstituted wood product mfg.	114	Secondary	2.04365	0.21739
Pulp mills	124	Secondary	2.16229	0.21456
Paper and paperboard mills	125	Secondary	2.41774	0.23452
Wood windows and door mfg.	117	Secondary	2.11123	0.18697
Veneer and plywood mfg.	115	Primary	2.41896	0.22046
Cut stock, resawing lumber, and planing	118	Primary	2.51333	0.28056
Wood preservation	113	Primary	2.51405	0.28056
Wood container and pallet mfg.	120	Primary	2.51405	0.28056
Hunting and trapping	17	Primary	2.68654	0.06250
Logging	14	Primary	2.72569	0.22500
Sawmills	112	Primary	2.84700	0.28157
Forest nurseries, forest products, and timber	15	Primary	4.71125	0.21944

Variable	*Sum total*	
output_2001 (Value added output, 2001)	5,548.07	(millions $)
output_2006 (Value added output, 2006)	7,382.42	(millions $)
imp_emp_2001 (Employment, 2001)	21,237	
imp_emp_2006 (Employment, 2006)	20,987	

Table 8. 2001 New Building Construction Cluster

Description	IMPLAN Code	Type	Z-Score	Linkage
Agriculture and forestry support activities	18	Secondary	2.38684	0.14992
Pipeline transportation	396	Secondary	2.35608	0.15110
Environmental and other technical consulting	445	Secondary	3.70603	0.18166
Other maintenance and repair construction	45	Primary	2.12844	0.23183
Maintenance and repair of highways, streets,	44	Primary	2.16423	0.18920
Water, sewer, and pipeline construction	40	Primary	2.26508	0.17410
Highway, street, bridge, and tunnel construction	39	Primary	2.42852	0.20660
Manufacturing and industrial buildings	37	Primary	2.89901	0.18561
Commercial and institutional buildings	38	Primary	2.93545	0.26826
Telecommunications	422	Primary	3.06343	0.13042
New multifamily housing structures, nonfarm	34	Primary	3.33948	0.23573
Other new construction	41	Primary	3.34927	0.26252
New residential 1-unit structures, nonfarm	33	Primary	3.41953	0.26569
Maintenance and repair of nonresidential buildings	43	Primary	3.42496	0.26569
New farm housing units and additions	36	Primary	3.69456	0.23277
New residential additions and alterations	35	Primary	3.69626	0.25858
Architectural and engineering services	439	Primary	4.78881	0.15020

Variable	*Sum total*	
output_2001 (Value added output, 2001)	24,285.61	(millions $)
output_2006 (Value added output, 2006)	34,718.75	(millions $)
imp_emp_2001 (Employment, 2001)	274,640	
imp_emp_2006 (Employment, 2006)	281,895	

Table 9. 2001 Textiles Cluster

Description	IMPLAN Code	Type	Z-Score	Linkage
Narrow fabric mills and schiffli embroidery	94	Primary	2.06468	0.21796
Cut and sew apparel mfg.	107	Primary	2.44990	0.20018
Nonwoven fabric mills	95	Primary	2.51757	0.25934
Fabric coating mills	98	Primary	2.58697	0.23323
Upholstered household furniture mfg.	363	Primary	2.65137	0.14195
Other apparel knitting mills	106	Primary	2.95429	0.22986
Other hosiery and sock mills	105	Primary	3.06284	0.29503
Knit fabric mills	96	Primary	3.11771	0.29787
Other miscellaneous textile product mills	103	Primary	3.28875	0.29713
Broadwoven fabric mills	93	Primary	3.32728	0.29045
Curtain and linen mills	100	Primary	3.39298	0.22387
Carpet and rug mills	99	Primary	3.55801	0.29812
Textile bag and canvas mills	101	Primary	3.63819	0.21209
Textile and fabric finishing mills	97	Primary	4.27126	0.22688

Variable	*Sum total*	
output_2001 (Value added output, 2001)	3,892.78	(millions $)
output_2006 (Value added output, 2006)	3,523.50	(millions $)
imp_emp_2001 (Employment, 2001)	32,574	
imp_emp_2006 (Employment, 2006)	21,323	

Table 10. 2001 Electronics Manufacturing Cluster

Description	IMPLAN Code	Type	Z-Score	Linkage
Industrial process variable instruments	316	Secondary	2.78363	0.24972
Semiconductors and related device mfg.	311	Secondary	3.17871	0.15699
Miscellaneous electrical equipment mfg.	343	Secondary	1.9608	0.20375
Electromedical apparatus mfg.	313	Secondary	2.05737	0.25198
Relay and industrial control mfg.	336	Secondary	2.09629	0.21396
Printing machinery and equipment mfg.	266	Secondary	2.25353	0.22586
Automatic environmental control mfg.	315	Secondary	2.51823	0.29187
Audio and video equipment mfg.	309	Secondary	3.62061	0.23344
Business support services	455	Secondary	2.36307	0.01651
Search, detection, and navigation instruments	314	Primary	2.69279	0.28442
Office machinery mfg.	270	Primary	2.73636	0.21450
Irradiation apparatus mfg.	320	Primary	2.90108	0.19610
Computer terminal mfg.	304	Primary	2.92021	0.22619
Photographic and photocopying equipment mfg.	272	Primary	3.28487	0.24708
Aircraft mfg.	351	Primary	3.29394	0.21082
Other communications equipment mfg.	308	Primary	3.45960	0.28829
All other electronic component mfg.	312	Primary	3.86071	0.31131
Electronic computer mfg.	302	Primary	3.89569	0.25963
Other computer peripheral equipment mfg.	305	Primary	3.92029	0.27803
Electricity and signal testing instruments	318	Primary	3.95912	0.28442
Broadcast and wireless communications equipment	307	Primary	4.02521	0.31833
Telephone apparatus mfg.	306	Primary	4.08301	0.29275
Electronic equipment repair and maintenance	484	Primary	4.37496	0.26126

Variable	*Sum total*	
output_2001 (Value added output, 2001)	5,554.11	(millions $)
output_2006 (Value added output, 2006)	9,436.34	(millions $)
imp_emp_2001 (Employment, 2001)	41,799	
imp_emp_2006 (Employment, 2006)	32,832	

Table 11. Agriculture Cluster Regression

Variable	Coefficient	Marginal Effects
Linear Regression		
Total land area/worker in sector k	-0.273*	-0.549***
Total property tax assessment/worker in sector k	2.794***	2.535***
Employee compensation in sector k /worker in sector k	-1.811***	-0.943***
% manufacturing establishments in 2000	-0.022	-0.019
% agriculture and forestry establishments in 2000	-0.012	0.001
% business services in 2000	-0.157***	-0.148***
Median household income in 2000	1.742**	1.628*
% commuting workers in 2000	-0.057***	-0.046***
Agriculture cluster dummy variable	-1.160***	-1.160***
Probit Regression		
Total land area/worker in sector k	-0.105***	-0.037***
Total property tax assessment/worker in sector k	-0.099*	-0.034*
Employee compensation in sector k /worker in sector k	0.331***	0.115***
% manufacturing establishments in 2000	0.001	0.000
% agriculture and forestry establishments in 2000	0.005	0.002
% business services in 2000	0.003	0.001
Median household income in 2000	-0.044	-0.015
% commuting workers in 2000	0.004***	0.001***
Constant	-0.570	
ρ	-0.800***	
σ	4.480***	
λ	-3.560***	
Wald test, H0: $\rho = 0$	192 (1 degree of freedom)	
Censored observations	33,039	
Uncensored observations	15,221	
Log likelihood	-53,554	

Note: Significance at the 1%, 5%, and 10% levels denoted by ***, **, and * respectively.

Table 12. Forestry Cluster Regression

Variable	Coefficient	Marginal Effects
Linear Regression		
Total land area/worker in sector k	-0.200	-0.457***
Total property tax assessment/worker in sector k	4.221***	4.070***
Employee compensation in sector k /worker in sector k	-1.701***	-0.880***
% manufacturing establishments in 2000	-0.033**	-0.029*
% agriculture and forestry establishments in 2000	0.014	0.028
% business services in 2000	-0.196***	-0.183***
Median household income in 2000	2.401***	2.314***
% commuting workers in 2000	-0.098***	-0.090***
Forestry cluster dummy variable	0.582***	0.582***
Probit Regression		
Total land area/worker in sector k	-0.108***	-0.038***
Total property tax assessment/worker in sector k	-0.063	-0.022
Employee compensation in sector k /worker in sector k	0.344***	0.122***
% manufacturing establishments in 2000	0.002	0.001
% agriculture and forestry establishments in 2000	0.006	0.002
% business services in 2000	0.005	0.002
Median household income in 2000	-0.037	-0.013
% commuting workers in 2000	0.003***	0.001***
Constant	-1.003	
ρ	-0.724***	
σ	4.514***	
λ	-3.269***	
Wald test, H0: ρ =0	252 (1 degree of freedom)	
Censored observations	33,039	
Uncensored observations	15,221	
Log likelihood	-53,982	

Note: Significance at the 1%, 5%, and 10% levels denoted by ***, **, and * respectively.

Table 13. New Building Construction Cluster Regression

Variable	Coefficient	Marginal Effects
Linear Regression		
Total land area/worker in sector k	-0.487***	-0.699***
Total property tax assessment/worker in sector k	1.459	2.442**
Employee compensation in sector k /worker in sector k	-1.730***	-0.937***
% manufacturing establishments in 2000	-0.035	-0.038*
% agriculture and forestry establishments in 2000	0.096	-0.024
% business services in 2000	-0.186**	-0.241***
Median household income in 2000	2.110*	2.564**
% commuting workers in 2000	-0.021	-0.036*
New building construction cluster dummy variable	0.381***	0.414***
Probit Regression		
Total land area/worker in sector k	-0.106***	-0.038***
Total property tax assessment/worker in sector k	-0.071	-0.033*
Employee compensation in sector k /worker in sector k	0.342***	0.116***
% manufacturing establishments in 2000	0.002	0.000
% agriculture and forestry establishments in 2000	0.005	0.002
% business services in 2000	0.005	0.001
Median household income in 2000	-0.032	-0.014
% commuting workers in 2000	0.003***	0.001***
Constant	-0.957	
ρ	-0.742***	
σ	4.515***	
λ	-3.352***	
Wald test, $\rho = 0$	288 (1 degree of freedom)	
Censored observations	33,039	
Uncensored observations	15,221	
Log likelihood	-53,872	

Note: Significance at the 1%, 5%, and 10% levels denoted by ***, **, and * respectively.

Table 14. Textile Cluster Regression

Variable	Coefficient	Marginal Effects
Linear Regression		
Total land area/worker in sector k	-0.474***	-0.728***
Total property tax assessment/worker in sector k	1.753*	1.586
Employee compensation in sector k /worker in sector k	-1.709***	-0.888***
% manufacturing establishments in 2000	-0.036*	-0.032
% agriculture and forestry establishments in 2000	0.057	0.069
% business services in 2000	-0.192**	-0.180**
Median household income in 2000	2.110*	2.034*
% commuting workers in 2000	-0.026	-0.019
Textiles cluster dummy variable	-1.090	-1.090
Probit Regression		
Total land area/worker in sector k	-0.106***	-0.038***
Total property tax assessment/worker in sector k	-0.070	-0.025
Employee compensation in sector k /worker in sector k	0.343***	0.121***
% manufacturing establishments in 2000	0.002	0.001
% agriculture and forestry establishments in 2000	0.005	0.002
% business services in 2000	0.005	0.002
Median household income in 2000	-0.032	-0.011
% commuting workers in 2000	0.003***	0.001***
Constant	-0.968	
ρ	-0.777***	
σ	4.329***	
λ	-3.486***	
Wald test, $\rho = 0$	306 (1 degree of freedom)	
Censored observations	33,039	
Uncensored observations	15,221	
Log likelihood	-53,815	

Note: Significance at the 1%, 5%, and 10% levels denoted by ***, **, and * respectively.

Table 15. Electronics Manufacturing Cluster Regression

Variable	Coefficient	Marginal Effects
Linear Regression		
Total land area/worker in sector k	-0.206	-0.461***
Total property tax assessment/worker in sector k	4.258***	4.103***
Employee compensation in sector k /worker in sector k	-1.697***	-0.882***
% manufacturing establishments in 2000	-0.028*	-0.024
% agriculture and forestry establishments in 2000	0.023	0.038
% business services in 2000	-0.189***	-0.177***
Median household income in 2000	2.275***	2.192**
% commuting workers in 2000	-0.098***	-0.090***
Electronics cluster dummy variable	1.844***	1.844***
Probit Regression		
Total land area/worker in sector k	-0.107***	-0.038***
Total property tax assessment/worker in sector k	-0.066	-0.023
Employee compensation in sector k /worker in sector k	0.344***	0.122***
% manufacturing establishments in 2000	0.002	0.001
% agriculture and forestry establishments in 2000	0.006	0.002
% business services in 2000	0.005	0.002
Median household income in 2000	-0.035	-0.012
% commuting workers in 2000	0.003***	0.001***
Constant	-0.990	
ρ	-0.770***	
σ	4.343***	
λ	-3.461***	
Wald test, $\rho = 0$	256 (1 degree of freedom)	
Censored observations	33,039	
Uncensored observations	15,221	
Log likelihood	-53,908	

Note: Significance at the 1%, 5%, and 10% levels denoted by ***, **, and * respectively.

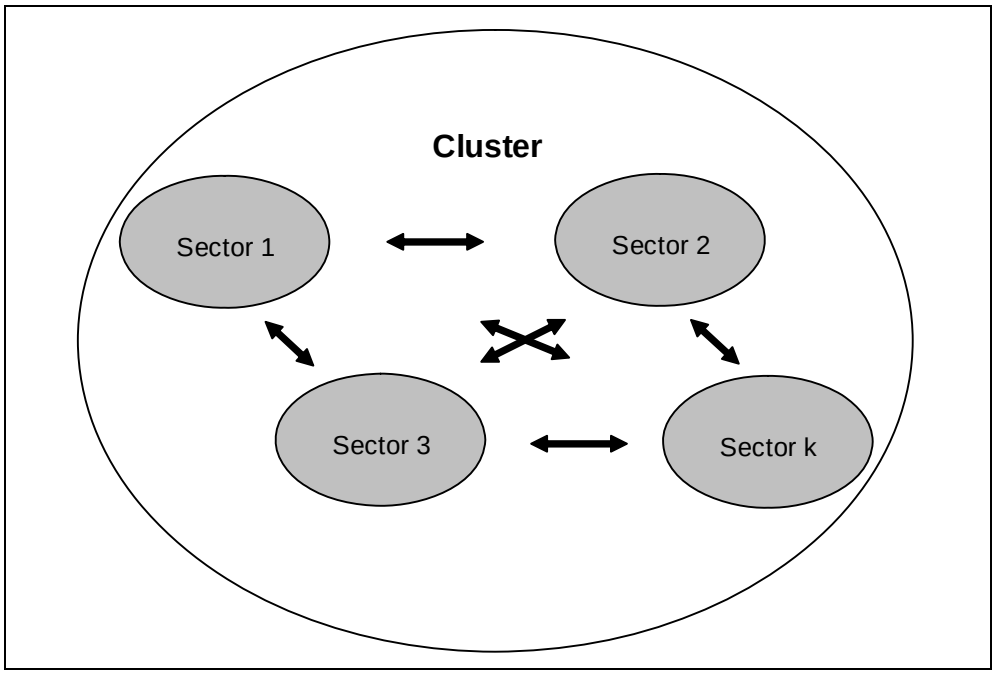


Figure 4. Schematic of the Industry Cluster and Value Chain Concept Used in this Research.

Arrows denote potential purchases and sales linkages between sectors, representing a cluster– specific value chain.

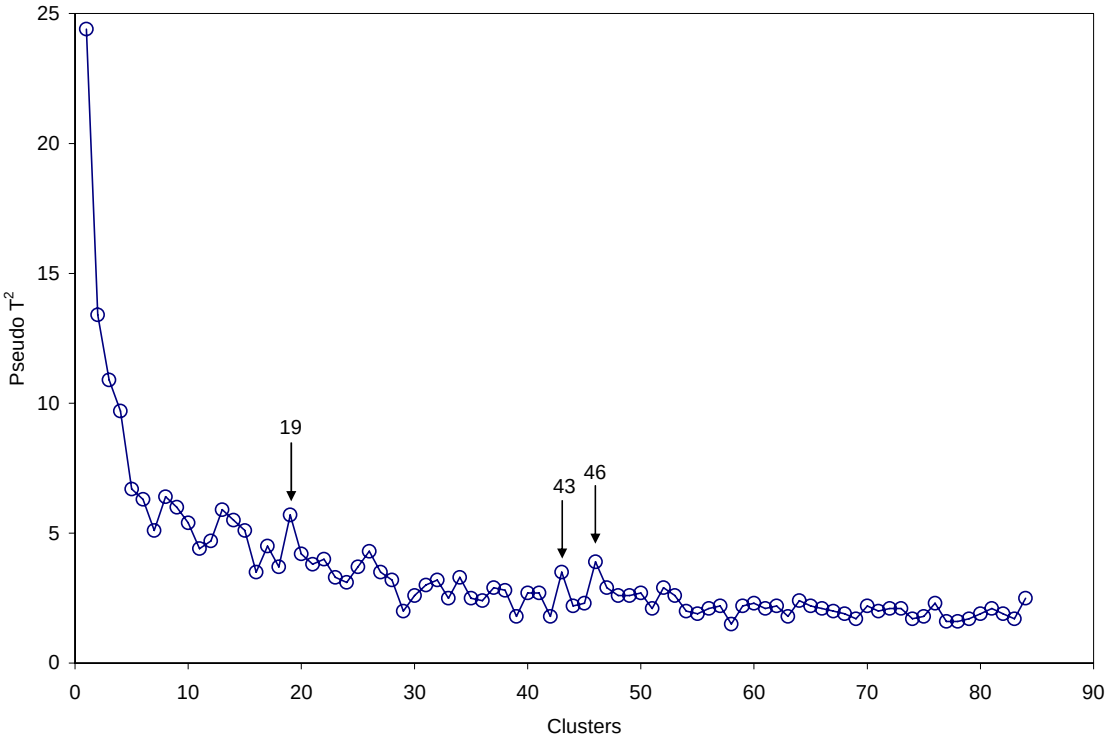


Figure 5. Candidate Clusters Identified with Ward's Clustering Algorithm.

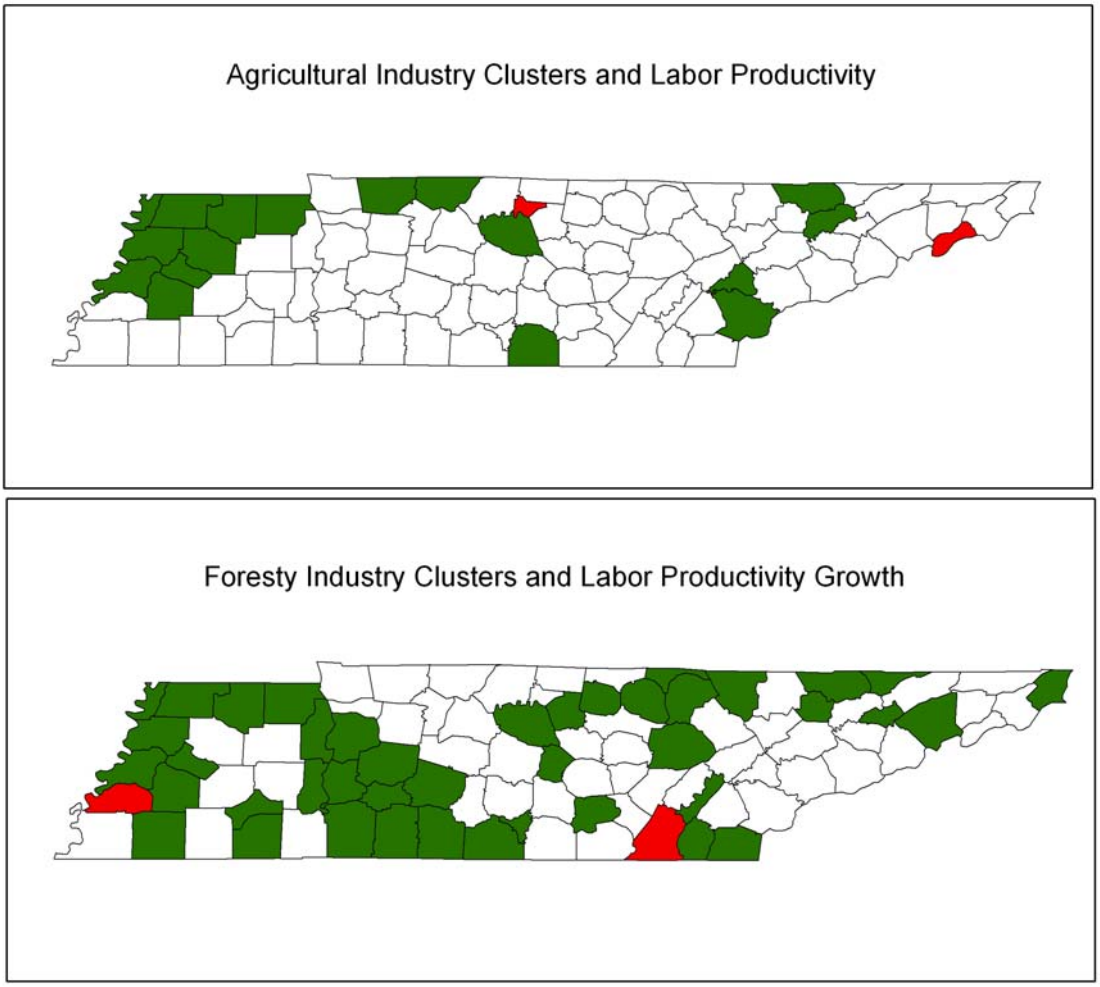


Figure 6.1. Spatial Distribution of County by Industry Cluster Interactions.

Key: Green counties, interaction effect was significant and positive between county and industry cluster; red counties, interaction effect was significant and negative between county and industry cluster. Counties where growth effect was not significant are not colored.

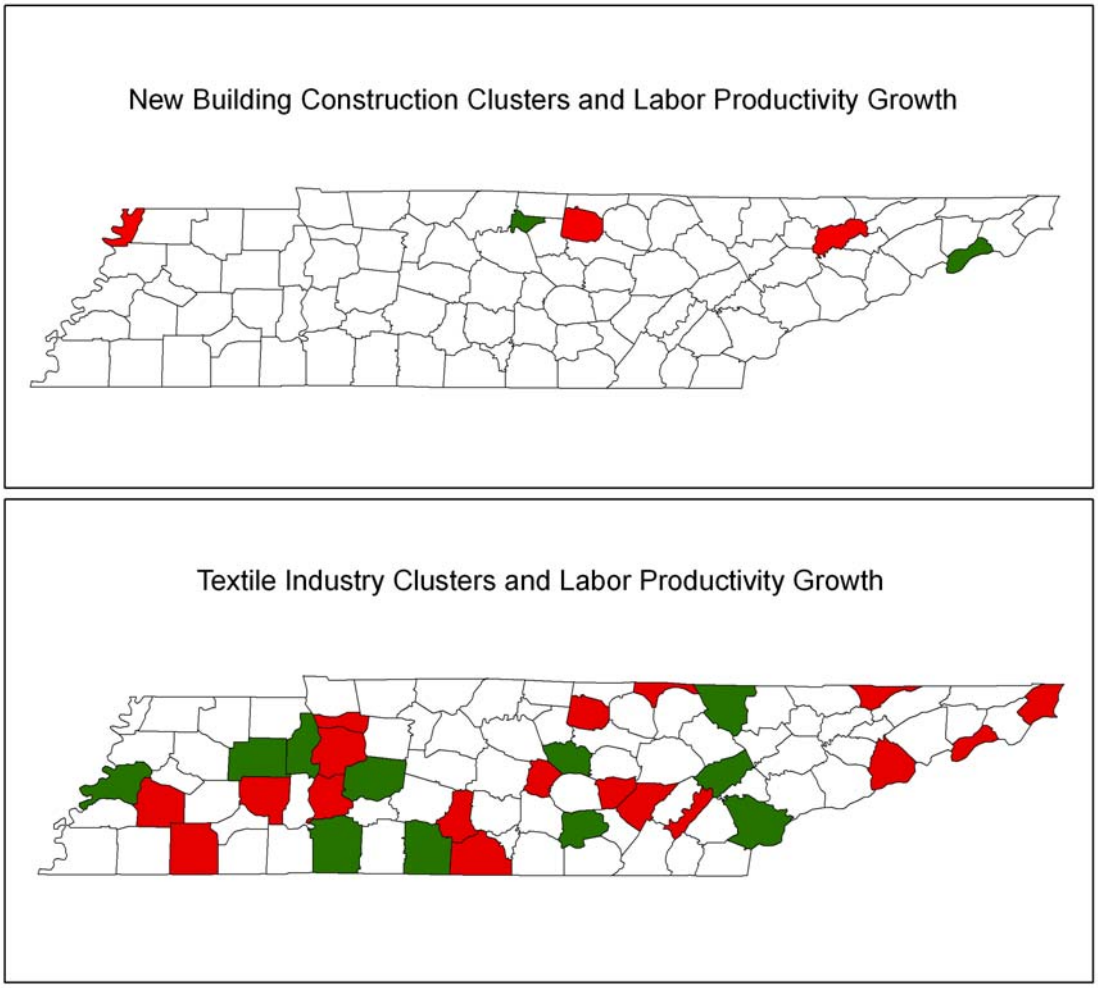


Figure 6.2. Spatial Distribution of County by Industry Cluster Interactions.

Key: Green counties, interaction effect was significant and positive between county and industry cluster; red counties, interaction effect was significant and negative between county and industry cluster. Counties where growth effect was not significant are not colored.

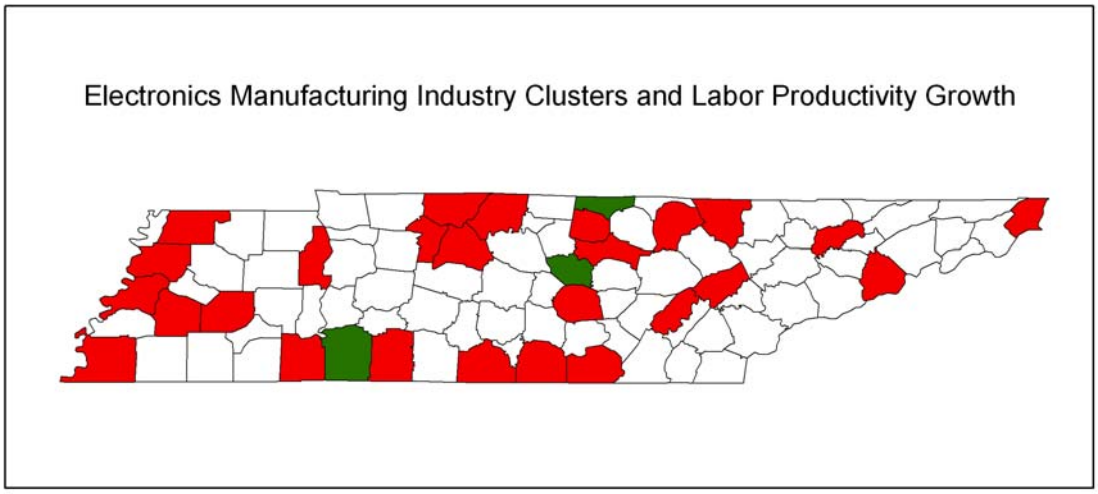


Figure 6.3. Spatial Distribution of County by Industry Cluster Interactions.

Key: Green counties, interaction effect was significant and positive between county and industry cluster; red counties, interaction effect was significant and negative between county and industry cluster. Counties where growth effect was not significant are not colored.

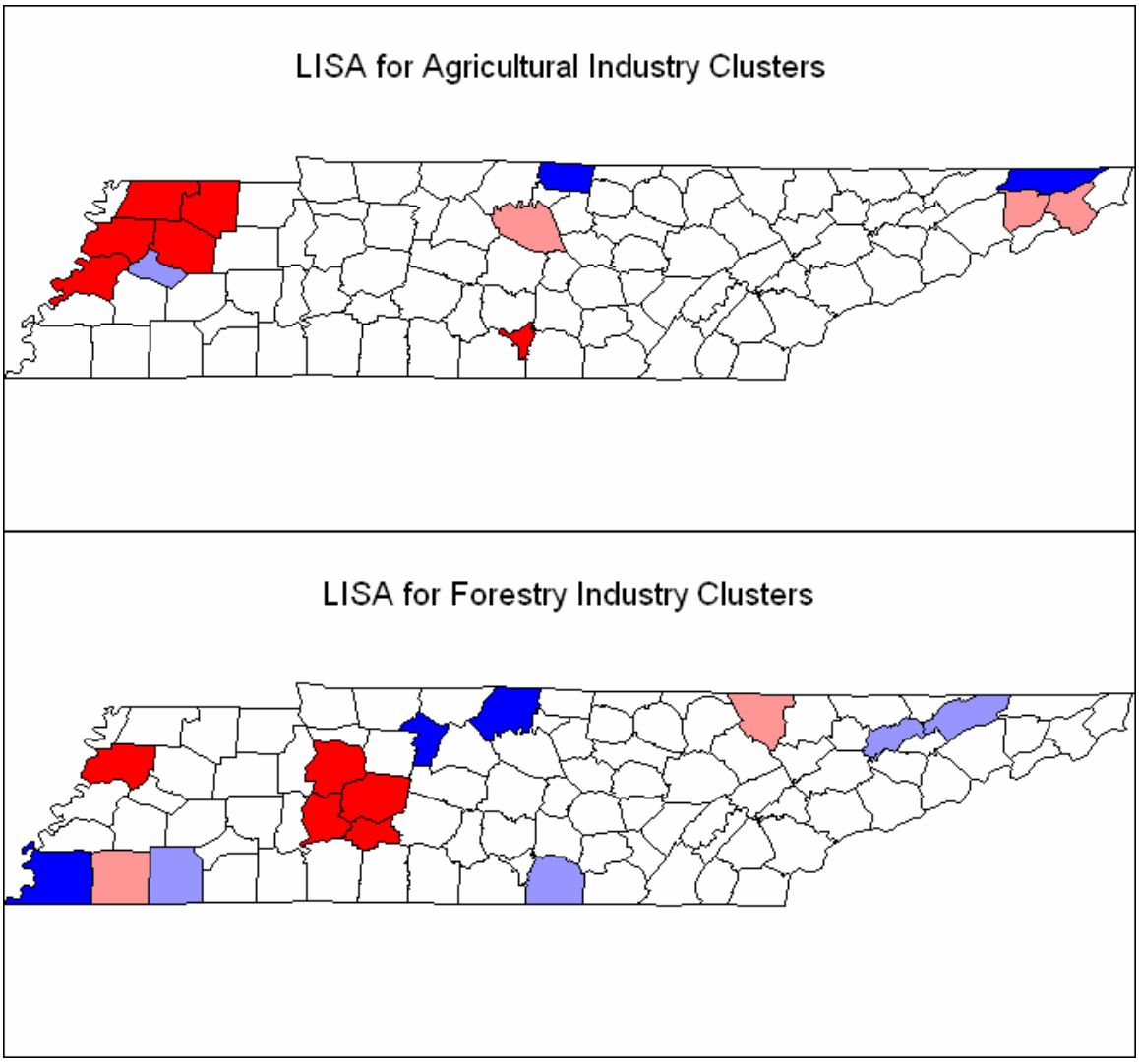


Figure 7.1. Local Indices of Spatial Association (LISA) for industry clusters (significant at the 5% level)

Key: Red counties, spatial clusters of counties relative where the interaction between industry clusters and the county fixed effect is relatively high; Blue counties, spatial clusters of counties relative where the interaction between industry clusters and the county fixed effect is relatively low; etc. Counties not belonging to spatial clusters are not colored.

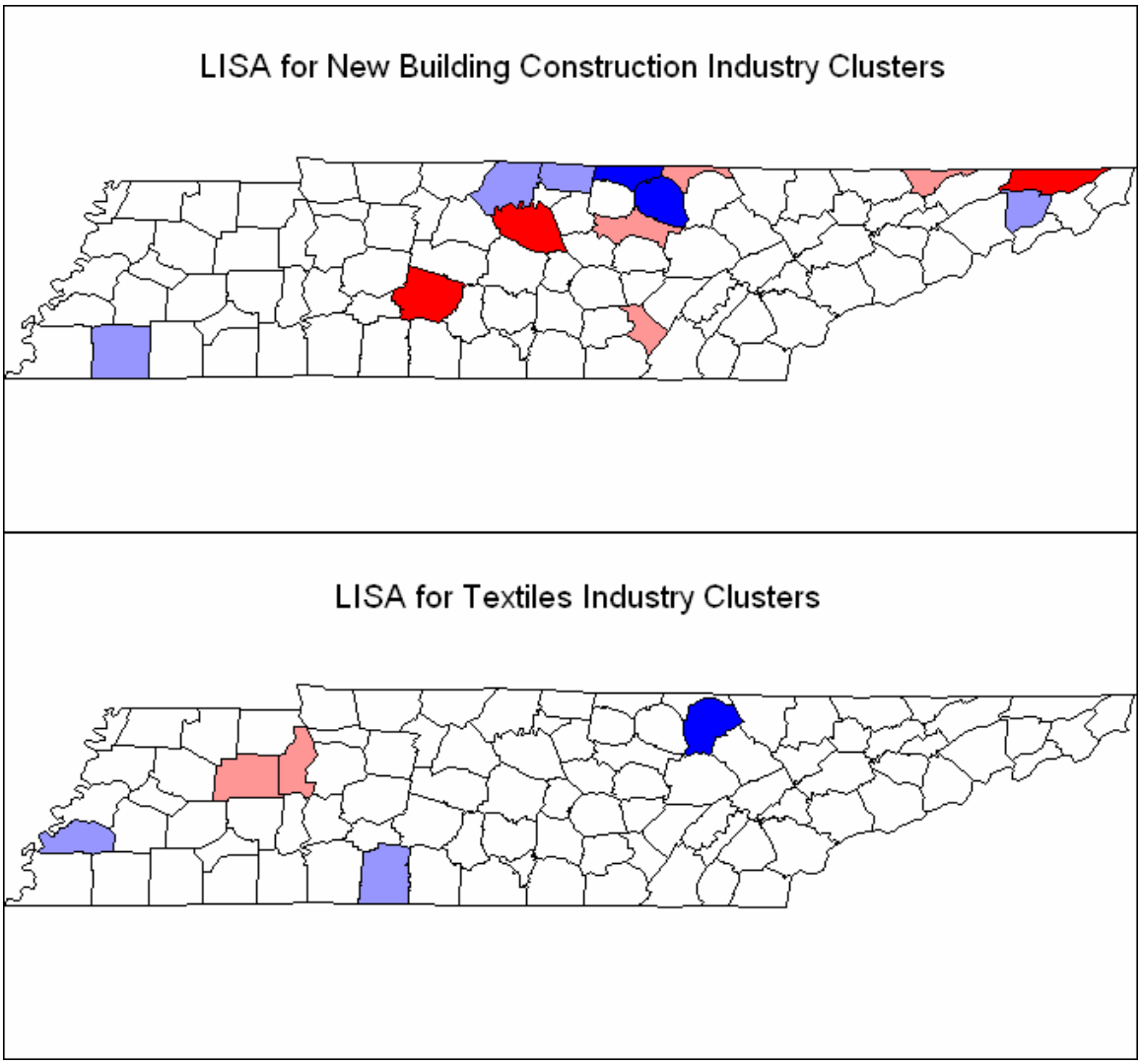


Figure 7.2. Local Indices of Spatial Association (LISA) for industry clusters (significant at the 5% level)

Key: Red counties, spatial clusters of counties relative where the interaction between industry clusters and the county fixed effect is relatively high; Blue counties, spatial clusters of counties relative where the interaction between industry clusters and the county fixed effect is relatively low; etc. Counties not belonging to spatial clusters are not colored.

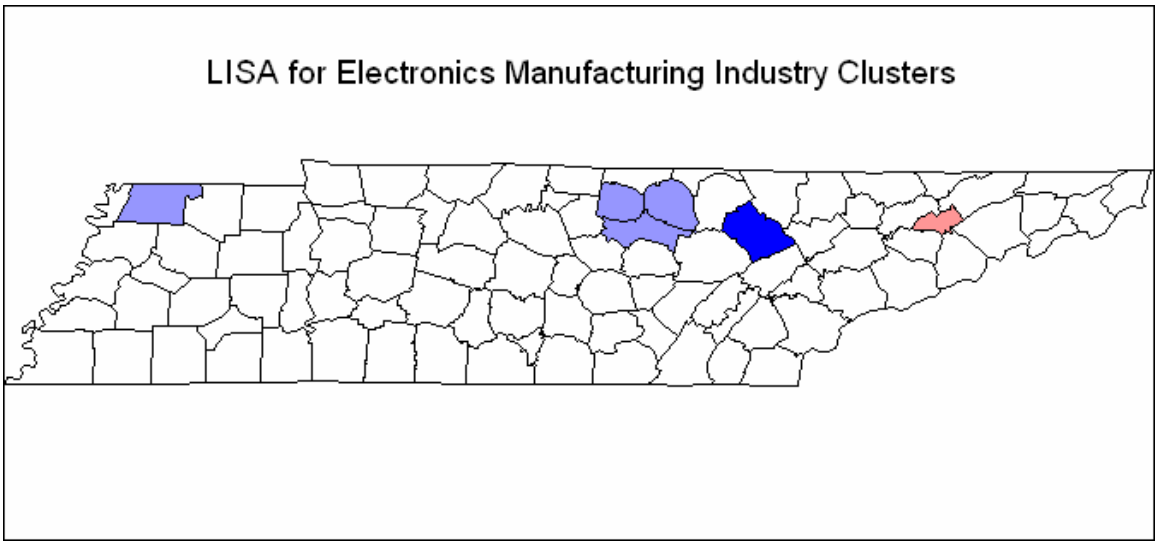


Figure 7.3. Local Indices of Spatial Association (LISA) for industry clusters (significant at the 5% level)

Key: Red counties, spatial clusters of counties relative where the interaction between industry clusters and the county fixed effect is relatively high; Blue counties, spatial clusters of counties relative where the interaction between industry clusters and the county fixed effect is relatively low; etc. Counties not belonging to spatial clusters are not colored.

IV: Summary and Conclusions

Rural communities struggle to keep pace with the changing economy which includes restructured industries, technological advances, and increasing globalization. Rural economic development strategies focusing on identifying a given community's comparative advantage may offer the promise of development for some rural areas. This thesis focused on identifying and evaluating county comparative advantage using two complementary approaches.

The first part of this thesis focused on the factors that influenced grain-based ethanol plant investment in a given county in the contiguous 48 United States from 2000 to 2007. Bivariate probit and spatial clustering methods measured the influence that local product and input market factors, transport and utility infrastructure, labor, state policy, local demographic characteristics, and farm resource regions had on attracting a grain-based ethanol production facility to a given county. The regression model was used to jointly estimate the location probabilities of plants currently producing ethanol, and plants that have announced interest of location in a county, or under construction.

Results from the bivariate probit model indicated that access to feedstock sources of corn, and the lack of preexisting ethanol production facilities were most important in the site selection decision for these plants. Other factors, such as counties with significant agricultural activities, prevalence of trucking establishments, extensive road networks, and states with bans on MTBE influenced ethanol plant investment as well. The effects of the location determinants are spatially heterogeneous, as suggested by the interaction effect with a rurality measure, suggesting that comparative advantage varies along the urban-rural spectrum. Results also suggest that extremely rural areas are unlikely recipients to attract ethanol plant investment. However, some rural areas may offer

comparative advantage with respect to attracting ethanol plant investment, given that these areas have access to resources such as grain storage facilities, some skilled labor, relatively diverse labor pools, rail road networks, or a large number of cattle, which may offer access to DG markets. Spatial analysis indicated that marginal or secondary sites are becoming attractive to recently announced plants as the ethanol industry approaches saturation, and the first-best sites are occupied.

The second part of this thesis applied new techniques to uncover linkages between purchasers and suppliers at the county and regional level for 447 economic sectors using IMPLAN data for Tennessee. The cluster analysis was extended to estimate the extent to which specific value chains contributed to economic growth between 2001 and 2006 (an important period spanning a brief recession) across Tennessee's 95 counties. County and regional comparative advantage was determined by testing whether the presence of a particular value chain in a given county increased labor productivity during this period, controlling for other local determinants of economic performance. Identification of industry clusters is important in the design and implementation of regional economic development strategies. Value chains establish and support the composition of industry clusters. By determining the sector composition of the value chains defining an industry cluster, more accurate information regarding recruitment of traditional value-added activities (such as manufacturing) or incubation of innovative endeavors (such as niche agricultural marketing) is possible.

Results focused on five specific industry clusters that exhibited strong linkages and are important to Tennessee's economy; agricultural, forestry, new building construction, textiles, and electronics manufacturing. Counties supporting the agricultural

and textile clusters experienced negative growth with respect to labor productivity. Counties supporting the forestry, new building construction, and electronics manufacturing clusters grew between 2001 and 2006, with the electronics manufacturing cluster experiencing the most significant growth during this time period. High probability clusters were located in various portions of Tennessee for the agricultural, forestry, and new building construction industry clusters. However, no significant high probability clusters were identified for the textile and electronics manufacturing clusters. Focusing on counties supporting these clusters may allow regional planners to develop an inventory of resources to focus on business recruitment or retention strategies targeting sectors belonging to clusters that contributed to growth between 2001 and 2006.

This thesis provided further insight into the concept of comparative advantage and regional development. The findings of the research add to the existing body of literature in a significant way by explaining factors that attract grain-based ethanol plants as well as explaining the economic potential and composition of industry clusters. Information gained through this study may be used to further educational efforts, as well as research and development efforts involving comparative advantage, ethanol site selection, and industry clusters.

Vita

Lance Andrew Stewart was born in Nashville, Tennessee, on May 23, 1984 to Tony and Linda Stewart. He graduated from Goodpasture High School in Madison, Tennessee in 2002. He attended Carson-Newman College in Jefferson City, Tennessee and received a B. S. in Financial Economics and a minor in Fine Arts in 2006. He married his wonderful wife, Christy on August 4, 2007. Lance was granted his M. S. in Agricultural Economics from the University of Tennessee, Knoxville in December 2008.