

IMPACT OF TRANSPORTATION ON BUSINESS LOCATION DECISIONS IN
RURAL UPPER GREAT PLAINS

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Mariya Burdina

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department
Agribusiness and Applied Economics

May 2004

Fargo, North Dakota

ABSTRACT

Burdina, Mariya; M.S.; Department of Agribusiness and Applied Economics; College of Agriculture, Food Systems, and Natural Resources; North Dakota State University; May 2004. Impact of Transportation on Business Location Decisions in Rural Upper Great Plains. Major Professor: Dr. Robert R. Hearne.

State and local officials who wish to encourage economic development and who should understand the process of making location decisions do not have a clear understanding of which factors are most important in a firm's location decisions. Hence, there is a great uncertainty in determining the importance of transportation infrastructure in the process of making business location decisions in rural areas.

Location of new manufacturing companies, employment data, and transportation factors from 424 rural counties were analyzed, and possible interactions between location decisions, employment, and transportation were investigated.

The dichotomous dependent variable logit model was used to determine location factors which influence business location decisions. Interstate and other principal arterials variables have positive influence, and distance to the nearest metropolitan statistical area has a negative impact on manufacturing firms with less than 50 employees.

Instrumental variable estimator and variance correction model were used to estimate impact of transportation on manufacturing and total employment, respectively. Interstate and other principal arterials variables are positively associated with total employment. Total lane miles and airport variables have strong positive impact on total and manufacturing employment. Distance to the nearest MSA was negatively and significantly associated with manufacturing employment in the county. Other transportation variables did not show any significance in the models.

ACKNOWLEDGMENTS

I would like to express special appreciation to Dr. Robert Hearne for his guidance and motivation. Recognition is given to my committee members, Dr. John Bitzan, Dr. Won Koo, and Dr. Denver Toliver; your knowledge and willingness to help are greatly appreciated.

A special thank you goes to Gene Griffin and Upper Great Plains Transportation Institute staff who have provided funding for this research and who always provided help.

I would also like to thank all professors who taught me during the two years of my master's program. A special thank you goes to Anatoliy Skripnitchenko for his valuable suggestions and assistance.

Thank you to all of my friends from North Dakota State University, who made these two years of study an unforgettable experience.

This thesis is dedicated to my parents, Tatiana and Ivan Burdin, for their love, support, and encouragement.

TABLE OF CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	iv
LIST OF TABLES.....	viii
LIST OF FIGURES.....	ix
CHAPTER I. INTRODUCTION.....	1
Research problem.....	1
Objectives.....	4
Hypothesis.....	5
Organization	5
CHAPTER II. LITERATURE REVIEW.....	6
Introduction.....	6
Transportation and rural economy.....	6
Industrial location theory.....	11
Evolution of location theory	11
Levels of business location decisions.....	14
Empirical studies	17
Econometric studies.....	17
Survey studies.....	23
CHAPTER III. MODEL DEVELOPMENT AND DATA	29
Introduction.....	29
Methodology.....	29

Business location decisions.....	29
Employment.....	34
Independent variables	35
Transportation factors.....	35
Economic factors.....	38
Social factors.....	39
Data	40
Summary.....	42
CHAPTER IV. RESULTS.....	43
Introduction.....	43
Collinearity diagnostics	43
Results of empirical location model.....	46
Logit analysis.....	47
Probability estimation.....	51
Results of empirical employment model	52
Summary.....	61
CHAPTER V. SUMMARY AND CONCLUSIONS.....	63
Introduction.....	63
Thesis summary.....	63
Results and conclusions.....	65
Location models.....	65
Employment models.....	66
Limitations and future research.....	67

LITERATURE CITED.....	69
APPENDIX A.....	74
APPENDIX B.....	76
APPENDIX C.....	77
APPENDIX D.....	78
APPENDIX E.....	79
APPENDIX F.....	80

LIST OF TABLES

3.1	The nearest Metropolitan Statistical Areas for the states used in the study.....	38
3.2	Predicted estimates of explanatory variables.....	40
4.1	Collinearity diagnostics results	45
4.2	Statistic characteristics of explanatory variables.....	46
4.3	Selected descriptive statistics of manufacturing firm entries: 1999-2004.....	47
4.4	Estimated effects of location factors on new manufacturing firm's locations with less than 50 employees.....	48
4.5	Estimated effects of location factors on new manufacturing firm's locations with more than 50 employees.....	49
4.6	The effect of one unit change on probability of new manufacturing firm with less than 50 employees' location.....	53
4.7	The effect of one unit change on probability of new manufacturing firm with more than 50 employees' location.....	54
4.8	Hausman test results.....	57
4.9	The results of Breusch-Pagan test.....	57
4.10	Estimated effects of location factors on total employment.....	58
4.11	Estimated effect of location factors on manufacturing employment.....	59

LIST OF FIGURES

4.1	Change in location probability for manufacturing companies with less than employees.....	55
4.2	Change in location probability for manufacturing companies with more than employees.....	55

CHAPTER I. INTRODUCTION

Research problem

State, local, and federal officials who wish to encourage economic development should understand the process of making business location decisions in the United States, but different economic specialists agreed that economic development officials and communities still do not have clear understanding of which factors are most important in a firm's decisions to locate in one area rather than another. As a result, there is a great uncertainty in determining the importance of transportation infrastructure in the process of making business location decisions in rural areas.

The relation between transportation and economic development is still not well understood. While a number of studies have attempted to examine the importance of different factors in economic development, only a few of them looked specifically into the role of transportation in a firm's location selection. The primary conclusion from empirical studies is that transportation can be a determining factor when other characteristics of communities are similar. In most of the cases, transportation seems to be less important than other factors like labor costs, and availability of raw materials.

Economic development may be defined as an increase in income of a particular region as a result of newly created economic activity which does not reduce the activity or income of other areas (Hough and Griffin, 1994). Sloggett and Woods (1989) defined reasons for economic development to be referred to the location of new firms:

- 1) Establishment of new businesses results in expansion and diversification of the economic base in state or community;

- 2) New business and industry provide new jobs and lead to employment and income increase; and
- 3) Attracting new business or industry is an accepted, traditional way which has support in the government, state, or other development organizations.

Transportation plays an essential role in economic development of the region. Building new and extending existent roads and widening highway lanes can result in many benefits for rural areas, such as better access to jobs and markets, and reduce costs of production which will lead to regional economic growth and development (Brown and Flake, 1999). Many rural areas play the role of the link between states or metropolitan centers. Commodities produced in the rural areas (agricultural products, food, fuel) should be transported to urban areas where they will be processed and consumed. In general, it can be said that transportation in rural areas is an important connector of rural areas to the nation.

Highways are one of the most important infrastructure links in rural development. The construction of a network of interstate highways has been a major factor in non-metropolitan industrialization. With construction of the highway network, manufacturing companies were able to locate in non-metropolitan areas but still maintain quick and economical shipment of produced goods to markets in metropolitan areas. Investing in highways has often been viewed as an effective economic development strategy, particularly for underdeveloped rural areas (Brown and Flake, 1999). If an improved highway network leads to expansion or diversification of a local area's economic base, it may also bring higher wages for workers and greater net

income for owners of local businesses. According to Brown and Flake (1999), highway expenditures lead to a rise in rural employment, particularly in manufacturing and retail industries.

Rural roads comprise 80 percent of national road miles and carry 40 percent of vehicle miles traveled. In general, transportation infrastructure in rural areas is well developed, and the transportation network of local roads, arterials, and highways is well established, but along with well-developed rural transportation, deteriorating bridges, narrow roads, and inadequate road maintenance occur in many rural areas. According to T. Besser (1997), the Midwest region experiences great migration of the rural population to metropolitan areas; hence, fewer people are responsible for the maintaining and upgrading of rural roads, which causes even greater loss in population and economic activity.

Manufacturing is one of the most important sectors in the economy of the U.S. According to J. Popkin (2003), some of the benefits of manufacturing include the following: every dollar in manufacturing goods generates 1.43 worth of additional economic activity, two thirds of private sector research and development are done in the manufacturing sector, manufacturing average wages are the highest among other industries, and manufacturing contributes one third of all corporate taxes collected by government. These and other benefits of manufacturing encourage local officials to attract new manufacturing plants.

At the end of the 1980s, a shift from urban to rural areas in manufacturing had begun. Lower costs of production (wage rate, taxes, land cost, etc.) are one of the reasons for manufacturers to be more attracted to rural areas than to metropolitan areas.

This shift helped to maintain the importance of manufacturing in rural areas (Wilkerson, 2001).

Over the last 20 years, the manufacturing sector has met many challenges, and as a result, the manufacturing share in national employment and personal earnings considerably decreased. J. Popkin (2003) suggested that “if the U.S. manufacturing base continues to shrink at the present rate and the critical mass is lost, the manufacturing innovation process will shift to other global centers. Once that happens, a decline in U.S. living standards in the future is virtually assured” (p. 2).

Objectives

The overall goal of this thesis is to identify the importance of transportation as one of the determinants of economic development in the Northern Great Plains of the United States. The following specific objectives are identified:

- 1) To identify the factors which significantly influence manufacturing and total employment in the Great Plains.
- 2) To identify the factors which significantly contribute to explaining company location in the Great Plains region.
- 3) To develop and employ an empirical model to estimate the importance of transportation infrastructure in business location decisions.
- 4) To analyze the relative importance of transportation factors in making business location decisions for different sizes of companies.
- 5) To calculate probabilities of location of new manufacturing companies in the Great Plains and compare the results with the observed dataset.

Hypotheses

The following hypotheses will be examined in the thesis:

- 1) There is a significant positive relationship between transportation infrastructure and the total and manufacturing employment rate.
- 2) The factors associated with total employment will differ from factors associated with manufacturing employment.
- 3) There is a significant positive relationship between transportation infrastructure and firm location decisions.
- 4) Factors associated with business location will differ for companies with a different number of employees.

Organization

Chapter II provides an overview of the existing literature on the impact of different factors on business location decisions and economic development. Chapter III describes methods used to estimate the impacts of transportation on business location decisions and evaluate the influence of transportation on employment growth in the Midwest region. Chapter III also contains the description of the data used in the analysis. Results are presented in Chapter IV. Chapter V includes thesis conclusions and limitations and an outline of needs for future research.

CHAPTER II. LITERATURE REVIEW

Introduction

The literature review focuses on the links between transportation and economic development. Issues involved can be put into context by considering literature from the following sources. First, rural economy and transportation approaches are discussed. Second, industrial location theory is reviewed, and finally, recent empirical studies that look at the role of transportation in business location decisions and economic development are presented.

Transportation and rural economy

Hough and Griffin (1994) define economic development as an increase in income or wealth of a state or a community as a result of economic activity that is newly created and that does not eliminate the activity or reduce the income of some other area. Economic development includes enlargement of existing businesses or attracting new business and industry. There are several reasons economic development often refers to recruiting new business or industry to community (Sloggett and Woods, 1989):

- 1) New business and industry will lead to expansion and diversification of the economic base of a community or region;
- 2) New business and industry will provide new jobs and will lead to employment and income increase; and

- 3) Attracting new business or industry is an accepted, traditional approach which has support in the government, state, or other organizations of development.

State economic development policies started to center recruiting of new businesses in the 1930s. Attracting new industry was the most important task for the communities, but for the small rural areas, this goal was unrealistic. Rural communities were not attractive for the large industries due to the lack of resources and amenities. In the 1970s, economic development officials shifted the focus from unrealistic desire to attract large companies to home-grown economic development (Hough and Griffin, 1994).

According to the location theory, which will be discussed later in this chapter, transportation plays an important role in making location decisions for new business or industry. Generally, two transportation objectives are taken into consideration while making a business location decision: 1) low transportation cost and 2) satisfactory transportation service (Sloggett and Woods, 1989).

If transportation costs are of major significance for a company, it will locate to where the cost of supplying raw materials and shipping cost of produced goods are minimized. Transportation costs vary depending on the type of product that needs to be carried. Raw materials that do not require special handling are usually shipped at lower cost than semi-finished or finished materials. However, finished products will have higher value and will increase the cost of transportation. Thus, difference in freight cost for shipping different kinds of products may affect the location choice of a manufacturing company.

For non-manufacturing companies, other transportation cost factors may serve the same basic role as freight. For example, for a corporate office with a great deal of personal travel, the cost and time required for executive travel are a major consideration.

Satisfactory transportation services are sometimes even more important than the cost of transportation (Sloggett and Woods, 1989). Companies are concerned about the availability of and proximity to interstate highways, railroads, and trucking facilities. In addition, air service and water systems are important for some industries. Existence of intermodal facilities will also greatly increase attractiveness of a community or region (Berwick et al., 2002).

Attractiveness of the community or region can be increased in different ways. Investment in transportation infrastructure makes communities more attractive through reduced production and distribution cost, improved access for employees, and better access to business customers. Thus, transportation investment can lead to growth in the number of businesses and expansion of already existing businesses in a particular community or region, and business growth implies more jobs. Expansion of existing companies will lead to greater profitability and/or increased market share, and as a result, employment and income will increase as businesses grow. Furthermore, expansion of existing businesses together with just-created ones will demand more raw material and components from their suppliers. Retail and wholesale businesses can expand as employees will spend more as a result of their additional income (Thompson, 1961).

Infrastructure services and improvements can be made in several forms:

- 1) Improved infrastructure quality—new highways, airports, commuter rail lines, etc.;
- 2) Repair and maintenance of existing infrastructure;
- 3) More capacity from existing infrastructure via intelligent transportation systems, better management of traffic flows and breakdowns, etc.; and
- 4) Changed user costs—fuel taxes, tolls, etc.

In the last decades, transportation in rural America experienced many changes by deregulation and devolution of federal responsibilities to state and local governments and traffic growth. All modes of rural transportation—highways, public transport, waterways, rail freight service, and air service—have been affected (Stommes and Brown, 2002). The focus of transportation programs shifted from building systems to adapting and improving already existing transportation facilities.

The interstate highway system expansion and modernization became critical in areas with expanding population and growing industries. The rail system has streamlined as a result of deregulation about 20 years ago, and nowadays, it is restructuring to meet changing rail freight needs. The water system that has been reorganized during the 1930s now experiences heavy demand from its users and increased maintenance and upgraded cost. Airport capacity is strained over all country, but rural airports are suffering because of high per capita cost as a result of serving a small number of customers.

Highways are one of the most important infrastructure links in rural development. Non-metropolitan industrialization would be impossible without construction of a network of interstate highways. With construction of a highway

network, manufacturing companies are able to locate in non-metropolitan areas but still maintain quick and economical shipment of produced goods to markets in metropolitan areas.

The U.S. rail system offers an alternative to highway transportation, particularly for bulk commodities and other low unit-value products. Railroads also play an important role for manufacturing companies, and many non-metropolitan communities benefit from their location on rail lines (Kale and Lonsdale, 1979). The inland waterway system provides a low-cost, effective means of transporting bulk products over long distances.

By serving as a catalyst for local job and income growth, airports can offer rural areas the potential of economic development. The absence of scheduled airline service may discourage new companies from locating in such areas; also, some businesses may refuse to locate in the area beyond reasonable distance from regularly scheduled airline service.

High fuel cost may reduce location attractiveness of a community even if it has advantages brought by interstate highways. Furthermore, high prices for gasoline may make it impossible for non-metropolitan employees to make long-distance trips to work.

Although the role of transportation in rural economy seems very important, this relationship is still not well understood. Different economists were studying this relationship from the beginning of the 20th century but still did not reach a conclusion. Location theory, which will be discussed in the next part of this chapter, provides different perspective to the relation between transportation and rural economy.

Industrial location theory

Communities and regions attract new businesses and industries to increase the local economic base. Location theory, developed with noted contributions from August Losch, Alfred Weber, and Johann von Thunen, offers an explanation for industrial attraction for a particular community or region. There are two parts in this section. First, evolution of location theory in contest with transportation issues will be summarized, and second levels of the business location decision will be discussed.

Evolution of location theory

The first location theories were developed by economists in the beginning of the 19th century. Johann H. von Thunen in 1826 formulated a theory which dealt with optimum location for agricultural crops in relation to a city (Miller, 1977). He assumed that a city is located in a homogeneous plain having uniform resources, and price is uniform, rent decreases, and transportation cost increases with distance from the city; thus, land rents and transportation costs are the determinants that affect the type of agriculture that could be practiced (Miller, 1977). The significance of von Thunen's work was widely acknowledged by other researchers. For example, Hite in 1999 said, "If modern economics began with Adam Smith, modern location economics began with Thunen, and it is to Thunen we first look for the most basic analytical model of the interplay between markets, production, and geography" (p. 1).

Alfred Weber in 1900 formulated a theory of industrial location. Optimal location of industries was based on minimization of transportation costs of raw materials and final products (Kilkenny, 1998). Weber made assumptions regarding the

locations of inputs and market for final products. Inputs were allocated across space discretely, away from output markets that were located in urban centers. Labor was available at fixed price and located in selected localities (Goode and Hastings, 1989). Composite costs of location consisted of transportation costs, labor costs, and agglomeration costs. Weber focused on multiple input sources and ignored outside demand and supply factors.

Weber found that firm which produce goods less heavy than the raw materials used in production, would settle near to the raw-material source. Firms producing heavier goods would settle near their market since they minimize the weight they had to transport. Weber defined transportation cost as the most important factor, and he believed that a plant would locate at the points where transportation cost are minimized. In general, Weber approached the firm location problem explicitly as a transport-cost minimization problem.

Another German economist, August Losch, developed a new location theory in 1930 which was opposite to Weber's least-cost theory. He believed that the best plant location is where revenue is maximized rather than costs minimized. Losch assumed that inputs were uniformly distributed over space and that any good can be transported on a straight-line route between any origin and destination. These assumptions pilfered from Weber, who assumed that inputs are located at selected locations with transportation network comprised of a limited number of trunk routes. The main finding of the Losch model is that producers will be spatially distributed in such a way that a cost of transporting of output to consumers will be minimized (Goode and Hastings, 1989).

Melvin Greenhut in 1956 developed a theory that provided a broader view on factors that influence the location pattern of industry. Greenhut synthesized earlier theories to form a profit maximization model (Hough and Dooley, 1991). He argued that location of any manufacturing establishment depends on different factors such as transportation and production cost, demand conditions, and cost-reducing and revenue-increasing factors (Miller, 1977). Finally, Greenhut was the first who suggested that when different locations have approximately the same revenue, the firm's location decision will be based on personal considerations (Goode and Hastings, 1989). Principles developed in the Greenhut location model were lately used in modern location theory.

As a conclusion, it can be said that in earlier theories, transportation services and their cost were frequently considered as an important factor in making location decisions. According to Goode and Hastings (1989), every prominent location theory of the 20th century dealt with transportation and its relative importance in the location selection process. Furthermore, through the evolution of location theory, transportation cost remained a major factor in making location decisions.

Modern location theory suggests that companies locate in places where cost and revenue considerations are balanced. According to Goode and Hastings (1989), transportation cost, tax rates, labor and land costs, intermediate inputs, and agglomeration effect should be considered on a cost side. On a revenue side, number and location of consumers, competitors' location, and existence of agglomeration effect must be considered. The next section more closely examines factors which influence business location decisions.

Levels of business location decisions

Although theoretical models provide useful background for understanding firm behavior, Salvesen and Renski (2002) argued that they usually fail to appreciate the complexity of the business location decision process. Theoretical models represent simplified reality and build on assumptions that do not exist in real life, like perfect competition, perfect information, and so on. In reality, business location decisions are not always rational, information is imperfect, and large companies can bargain with local government when considering alternative sites (Salvesen and Renski, 2002). Theoretical models usually do not disclose the actual process of making the location choice—how the location decision is actually made and how firms estimate trade-offs among different locations (Salvesen and Renski, 2002).

The basic approach in selecting location is that the company is looking for a region/community with such characteristics that will match the company's needs (Murray and Dowell, 1999). There could be a hundred factors that might be taken into account while making the location decision, but only a few of them are really important. At the same time, it is impossible to recognize and evaluate all possible locations that are suitable for a firm according to chosen criteria. For this reason, the process of making the location decision is intended to limit consideration to the most important factors for the most likely candidate locations while making sure that all viable alternatives are considered (Ritter, 1990).

Making business location decisions is a multistage process. Location factors are different for different industries and even for different firms as well as different stages of the location selection process, but it is generally understandable that location factors

and their relative importance for a company will differ from one industry to another and much less understood that location factors will significantly differ in selection of state, community, or site location. Location factors as well as their relative importance will differ in each level of the process of making the location decision. Some economists like Kinnard and Malinowski (1961) and Thompson (1961) distinguish three levels of business location decision and identify them as primarily factors, secondary factors, and tertiary factors.

Primary/regional choice factors are used by decision makers during the initial stage to delineate the broad area within which the firm will be located. Five factors are usually mentioned when determining the general location of the firm: 1) relation to markets, 2) location relative to raw materials, 3) transportation cost and service, 4) availability and cost of energy resources, and 5) labor cost (Thompson, 1961). Manufacturing companies in their selection of general area should consider the following factors: raw materials, energy resources, labor, transfer and transportation cost, market, and capital. If any of these factors is missing, manufacturing cannot occur in this area (Miller, 1970). Usually, during the first stage, companies choose the state where it will be located.

Primary factors are factors that affect cost, or in other words, the first group of factors includes various inputs that are needed in business. On the other hand, secondary and tertiary factors represent revenue factors and are related to the attributes of the areas (Risto, 1995).

Secondary factors are used to narrow the area to one or several communities within the region that was selected in previous step. Although some of the factors may

overlap primary or tertiary factors, in general, they differ from factors influencing the state of site selection. Thompson in 1961 distinguished the following factors that should be considered in choosing a community for firm location: labor supply, cost, climate, transportation facilities and cost, proximity to markets, community appearance and facilities, local taxation, and availability and cost of utilities. If those factors among different communities are approximately the same, site characteristics may be the determining factor in making the location decision (Goode and Hastings, 1989).

The third stage of the location decision involves tertiary factors. On this level, decision makers select the site where the company will be located. Although tertiary factors may also overlap with primary or secondary factors, there are also factors that influence particularly site selection. Examples of tertiary factors are availability and cost of utilities, soil conditions, drainage, road surface, availability to expand, and so on (Kinnard and Malinowski, 1961). Thompson (1961), in addition to above factors, added following: location in the community, price, possibility of flooding, transportation facilities, taxes, insurance, fire and police protection, and so forth.

In some cases, the number of levels of business location decisions may vary. The decision to expand or relocate is risky and expansive. For this reason, many companies decide to expand/relocate near an existing site or at least within the same state. In this situation, the firm will not pay attention to first-level/primary factors which could have been important for the other company that is choosing its new location among different states.

For some companies, the same factors play important role on all three stages of making the location decision. As an example, transportation cost may first lead the company to a particular region, then to a community, and finally to a specific site.

Empirical studies

In the past decades, there have been many empirical studies directed to identify the factors that influence business location decisions and the factors of relative importance in making location decisions, but only a few of those studies dealt specifically with the impact of transportation on making location decisions.

There exist two types of empirical studies of business location: revealed preference and stated preference studies. Revealed preference studies rely on researchers' measurement of weighting of location factors without considering statements or motivations of companies involved. Survey studies based on a large amount of mailed or other types of questionnaires and importance of location factors are developed from the statements of the participants of this study.

Econometric studies

Goode and Hastings (1989) examined the impact of transportation services on a community's ability to attract manufacturing plants in New England, mid-Atlantic, and southern Atlantic states. In their study, Goode and Hastings (1989) used regression analysis to identify the factors associated with the industrial location. As a dependent variable, they had a value of 1 if the community attracted a new plant during the period 1970-78 and a value of 0 if there was no new plant. For independent variables,

community characteristics were used. Sixty-nine industries identified by four-digit Standard Industrial Class (SIC) were analyzed.

Six transportation factors were included as independent variables:

- 1) Distance to Road (DR) measured the road distance between a community and the nearest non-limited access state or federal paved road,
- 2) Distance to Limited Access Four-Lane Highways (DLH) measured the road distance from the community to the nearest limited-access four-lane highway,
- 3) Number of rail lines serving the community,
- 4) Number of airlines serving the community,
- 5) Potential net input availability, and
- 6) Market access.

Results for the non-metropolitan communities indicated that rail service was an important location factor for only five industries, primarily for food processing and wood products industries. These industries in the Northeast depend on the rail service for delivery of raw materials from Midwest. Industrial location for nine of the industries was positively associated, and five industries are negatively associated with air service in the Northeast. Industries that are positively correlated with air service include large firms with multiple branch plants. Industrial location for five industries was positively associated and for four industries was negatively associated with proximity to limited-access highway. Access to federal roads is positively associated with six industries and negatively associated with five industries and has a similar impact on rural economic development as rail service does. Market access variable was

positively associated for 35 of the industries and negatively associated for 29 of the industries.

Findings of this study showed that results for small metropolitan communities differ from results for non-metropolitan communities. Air service has less impact on economic development in small metropolitan communities. Distance to limited-access highways has a stronger influence on business location in small metropolitan communities than in non-metropolitan communities. Rail service has in general a positive influence in both small metropolitan and non-metropolitan communities.

Warren Kriesel (1984) identified factors affecting plant location in Virginia and estimated the probabilities of communities attracting one or more manufacturing plants. In his study, Kriesel used the logit model to estimate parameters that influence plant location decisions. The dependent variable was equal to '1' if at least one manufacturing plant was created in community and '0' if no manufacturing plants were created during the period 1979-81. As independent variables, the following characteristics were used:

- 1) Number of employees in manufacturing sector in the jurisdiction,
- 2) Labor force,
- 3) 1-0 presence of an interstate highway in the jurisdiction,
- 4) Existence of a development group,
- 5) Fire protection rating,
- 6) Number of population older than 25 with high school diploma,
- 7) Distance to college,
- 8) Average weekly manufacturing wage, and

9) Site quality index.

The results showed that the number of manufacturing employees in the manufacturing sector and the average manufacturing wages were significant explanatory factors of location at the 0.05 level of significance. Presence of an interstate highway did not have a strong effect on plant location decisions in this model.

James Wheeler and Clifton Pannell (1977) studied the influence of transportation on economic development location in Appalachian Georgia. One of the purposes of this study was to estimate relationships between manufacturing employment and different socioeconomic characteristics of a particular region. First, using shift-share analysis, the comparison of the area's manufacturing growth performance was made which allowed for determining growth and importance of manufacturing in study area. After shift-share analysis, the multiple regression analysis was made. Multiple regression analysis measures the amount of variation in a dependent variable that is explained by the independent variable(s). This methodology was used in order to determine the impact of available land, labor, utilities, and transportation on manufacturing employment in each county.

Total manufacturing employment in each county was used as a dependent variable, and as independent variables, the following counties' characteristics were used:

- 1) Population density,
- 2) Miles of state and federal roads,
- 3) Presence or absence of interstate highways, and
- 4) County land area in national forest.

The results of the study indicated that such factors as adequate labor, urban services, and accessibility are crucial for counties with “weak” manufacturing growth. For the counties with “strong” manufacturing growth, the independent variables explained much less of the variation. The findings showed evidence of the primary importance of highway miles and accessibility in “strong” manufacturing employment.

Walker and Greenstreet (1991) estimated the impact of government incentives and assistance on location and job growth in manufacturing. Data were analyzed through an application of discrete choice modeling. Results of the research showed that access to highways was not a significant factor in business location decisions and job growth during the 1980s in the Appalachia manufacturing sector. Walker and Greenstreet (1991) explained it by existence of negative collinearity between the road variable (existence of interstate in county) and unemployment variable, which might obscure the true effect of road access on business expansions.

Stephanedes and Eagle (1986) investigated interactions between transportation infrastructure and manufacturing and retail employment in 30 non-metropolitan Minnesota counties. No significant correlation between state highway expenditures and either manufacturing or retail employment was found from a cross-sectional analysis of the data. Analysis of 30-county time series data showed that highway expenditures influence manufacturing and retail employment, and employment affects highway expenditures. From the analysis, the authors conclude that employment increases in the first 2-3 years following highway improvements, but in a decade, it returns to its initial rate.

Craine (1993) evaluated the effect of transportation expenditures on income and employment in 254 Texas counties by using the distributed lag model. The results indicated that the impact of transportation expenditures on income and employment will be different for rural and urban areas. Authors conclude that transportation expenditures play an important role in rural America. The results showed that transportation expenditures have a positive impact on non-farm and farm incomes in rural areas and on employment in rural areas.

Carlino and Mills (1987) estimated the effect of economic, demographic, climatic, and policy-related variables on the population and employment growth in 3,000 counties in the U.S. during the 1970s. The two-stage, least-square model was used in order to determine the impact of county characteristics on the county population and employment growth. Dependent variables in the model refer to population and manufacturing employment densities. As independent variables, economic, demographic, climatic, and other characteristics of the county were used. The findings of the study showed strong positive correlation between interstate density and employment and population growth. Elasticity analysis showed that doubling the miles of interstate per square miles of land in a county will lead to an increase in manufacturing employment of about 6.0 percent and an increase in population of about 2.8 percent.

Bartik (1985) examined the relationship between the location decision of a new manufacturing plant in the U.S. and different factors that influence that decision. To examine the influence of a state's characteristics on location decisions, he used a conditional logit model developed by McFadden in 1974. As independent variables, the

following state's characteristics were chosen: land area, unionization percentage, corporate tax rate, property tax rate, unemployment insurance tax rate, workers' compensation insurance rate, road miles, existing manufacturing activity, wage rate, educational level of population, construction costs, population density, and energy price. Results of the study indicated that new manufacturing plants in the 1970s were more likely to locate where adequate roads already existed.

Survey studies

David J. Forkenbrock and Norman S. J. Foster (1996) in their article examined the relationship between highway investments and facility location. Data for the survey were collected from two states, Missouri and Iowa. The questionnaires were mailed to managers of all manufacturers and warehousing facilities in Iowa and Missouri with 50 employees or more. The managers were asked to rank seven factors that can influence their location decisions like labor quality, labor costs, proximity to markets, proximity to input materials, transportation services, utilities, and tax rates.

The results indicate that transportation service was ranked third in Missouri and fifth in Iowa which basically was consistent with the author's opinion about the decreasing role of transportation in business location decisions. Relatively few Iowa respondents felt that it would be good public policy to expand the state's four-lane highway network if user taxes must be raised to do so or if less would be spent to maintain the existing network as a result. Two-thirds of the Iowa respondents also suggested that as long as reasonable access to the four-lane highway exists, the distance to this highway is not important.

The general conclusion of the study was that although many facilities rely quite heavily on highway transportation, close proximity to an interstate or other four-lane highway often is not seen as essential and other factors, such as quality and cost of labor, play a larger role in making business location decisions.

Button and Scott (1995) conducted a survey that was aimed to estimate the importance of transport and related infrastructure in the location/relocation process. Companies were asked to appraise the importance of different factors which they considered when making location/relocation decisions. Eighteen factors were used for estimation, and four of them were transportation factors such as road links, bus links, air links, and rail links. The analysis showed that road links were considered by a firm's representatives as the most important factor in company location. Some general accessibility factors (market access, support services) were rated higher than bus or air links. The further analysis allowed making a conclusion that weak transportation infrastructure does not stimulate the firm's migration, but the nature and quality of the transportation system become a very important factor for companies that make a decision to relocate.

Fox and Matthew (1990) examined the effects of a wide-ranging set of local public policies on interregional business location decisions. In the empirical model, firm entries were aggregated to the county level and divided by the stock of active firms to yield a county entry rate. A pooled time series cross-section database on establishment entries into 95 Tennessee counties for the years 1980 to 1986 was the basis for construction of the dependent variable. Independent variables were presented by the following attributes: tax, per capita value of non-school transfers, county

expenditures on street and highways to total operating expenditures, the county education expenditures, presence or absence of interstate highway and major rail line, distance to the nearest major airport, and other.

The main conclusions of this study were that firm location decisions are sensitive to variations in local public policies, especially to those policies such as the presence of interstate highways and rail infrastructure and educational policy. Also, the results of the current study indicated that firms of different size respond differently to site attributes.

The study by Rainey and McNamara (1992) was conducted to examine factors affecting manufacturing location in Indiana. The variables chosen to represent those factors in this study are listed below:

- 1) Labor characteristics;
- 2) Taxes;
- 3) Government expenditures;
- 4) Agglomeration economies, which refer to the cost savings associated with size and concentration of business activities;
- 5) Quality of life;
- 6) Access to markets;
- 7) Attitude toward business, meaning business climate; and
- 8) Other community-specific costs.

As a dependent variable, the increase in the number of manufacturing establishments within each Indiana county between 1986 and 1989 was used. The findings of the study implied that local policy affects manufacturing location. Authors

conclude that although agglomeration characteristics, highway access, labor availability, and cost are beyond local control, policymakers can influence development potential through tax and educational policies. A firm's perception of a community attitude toward business and quality of life also affects its location decision.

Leitham, McQuaid, and Nelson (2000) investigated the influence of road transport and other factors on industrial location in terms of the decision-making process by applying the stated preference experiment. In order to investigate the relationship between transportation and industrial location and operation, a computer-based "face-to-face" survey was carried out. Companies that had experience in making location decisions in the previous 10 years were under consideration. A total 40 interviews were carried out with senior managers at the premises of firms in the Strathclyde region of Western Scotland, Great Britain. Results of the study indicated that road links do not have a strong effect on business location decisions as specified by Button (1995), who said that road links have the highest rate among other location factors considered by firm. Also, the findings of the study showed that companies that have inter-regional links and wide market areas consider the level of transportation links in their location decisions. Good public transportation and proximity to the motorway were significantly important factors for those companies.

Kinnard and Malinowski (1961) conducted a survey that dealt with influence and significance of the existing highway system on the location decisions. They sent the mail questionnaire to the manufacturing firms which have been involved in a new location or a relocation decision in the Hartford Economic Area between 1953 and 1956. The main findings of that study were as follows:

- 1) It is rather the network of existing roads than any one special highway or road influencing the location decision process,
- 2) The importance of the highway system differs for different industries, and
- 3) Many firms do not really need particularly good access and find avoidance of the congestion that they associate with highway locations more significant as a motivating factor in plant site selection, etc.

Leistriz (1990) identified new or growing export-oriented businesses and industries in the Upper Midwest to determine both their economic contribution and factors critical to their location decisions. The mailed questionnaire was used to obtain information about each firm's current operations and factors that were important in location/relocation or expansion decisions.

Companies were asked to rate the number of factors to determine their influence on a firm's decision to locate or relocate. The transportation factors that were listed among the other factors and the results of the study were as follows: motor freight service was seen as the most important factor among transportation factors, and more than half of the new and relocation firms and almost half of the expanding firms rated motor freight service as critical or very important to their location decision.

Hough and Dooley (1991) analyzed the relationship between transportation and economic development. The data for this study were collected by a mail survey. The survey was sent to economic development specialists from three states in the Upper Great Plains. The perceptions of economic development specialists were compared with perceptions of manufacturers which were obtained by Leistriz and Ekstrom

(1990). The results indicated the difference in the ranks of main location factors between economic development specialists and manufacturers. The difference also existed in perceptions of the small metropolitan and rural economic specialist. The small metropolitan economic development specialist ranked transportation factors higher than the rural specialist. Transportation was defined as an important factor for businesses and economic development.

Helgeson and Zink (1973) assessed the feasibility, costs, and benefits of incorporating a manufacturing sector into a local economy dominated by agriculture. They evaluated unique location factors for study area through questionnaire response from four manufacturing firms located in Jamestown, ND, and four firms that considered the site. They found that factors designated of greatest influence to site selection were community attitudes toward industry, grants and concessions made available to manufacturing firms, and labor-related factors. Positive factor evaluations related to the labor resource were labor costs, willingness of local workers, a large supply of trainable labor, labor laws, and labor unions. Factors considered to be a negative influence were lack of subcontractors, proximity to other industry byproducts, and a shortage of skilled labor.

Analysis of empirical studies showed that transportation can be a determining factor when other characteristics of communities are similar. In most of the cases, transportation seems to be less important than other factors like labor costs, availability of raw materials, and proximity to the markets.

CHAPTER III. MODEL DEVELOPMENT AND DATA

Introduction

This chapter describes the methodology and data used to estimate the impact of transportation factors on business location decisions and evaluate the influence of transportation on employment growth in the Midwest region. In summary, cross-section data on county characteristics in 2000 were gathered, estimated, and organized. A dichotomous logit model was used to estimate the impact of transportation on creating new companies in the county, and multiple regression models are used to evaluate the influence of transportation on employment growth in the county.

The first part of the chapter consists of description of econometric models used in the study. The second part is focused on describing dependent and independent variables used in the models. A full description of data used in the research is given in the final part of the chapter.

Methodology

Business location decisions

There have been several studies with the same objectives as this research. Goode and Hastings (1989) identified the importance of transportation factors in business location decisions. They applied the ordinary least-squares (OLS) model to estimate the model,

$$Y_i = \sum_j a_{ij} X_{ij} + \sum_k b_{kj} Z_{kj} + e_j$$

where dichotomous choice dependent variable $Y_j = 1$ if community j attracted a new manufacturing plant and 0 otherwise. Transportation and socioeconomic independent variables are represented by vectors X_{ij} and Z_{kj} , respectively. OLS probability model has two main shortcomings:

- 1) Heteroscedasticity of error term, and
- 2) Parameter predictions generated may take values from negative to positive infinity, not meaningful for probability explanation.

To avoid such difficulties, multiple choice (conditional logit) or dichotomous choice (logit or probit) models should be used in this analysis.

Carlton (1979) and Bartik (1985) use a conditional logit model developed by McFadden in 1974. The conditional logit model is used to analyze choice behavior, which is described by 1) the objects of the choice and alternatives available for the decision maker, 2) observed attributes of the alternatives, and 3) actual choice and behavior of decision makers (McFadden, 1974). Although information on attributes of the counties being selected and actual choice of companies was known, there was no available information on alternatives that the firm considered while making the location decision. In other words, if a new company decided to locate in Colorado, it is doubtful that North Dakota counties, which are also included in the study area, were considered as an alternative to the chosen location. For this reason, a dichotomous dependent variable logit model is used to determine location factors which influence business location decisions.

The dichotomous choice logit model is used when the location decision is cast in terms of whether or not a given location attracted new business. The maximum

likelihood estimation procedure is used to estimate logit parameters. The probability of choosing location i is given by

$$P(Y = 1|X) = e^{\sum b_n X_n} / (1 + e^{\sum b_n X_n}).$$

This, in turn, implies that probability of not choosing location i is

$$P(Y = 0|X) = 1 - P(Y = 1|X) = 1 / (1 + e^{\sum b_n X_n}).$$

Then the likelihood function is formed as

$$L = \prod [(e^{\sum b_i X_i} / (1 + e^{\sum b_i X_i})) \prod [1 / (1 + e^{\sum b_j X_j})],$$

where i refers to companies located in the county, and j refers to companies that did not locate in the county. Maximizing this likelihood with respect to vector b will obtain that the probability of locating in the particular county is estimated as

$$P(Y = 1|X) = \exp(\sum b_i X_i) / [1 + \exp(\sum b_i X_i)],$$

where P is predicted probability, vector X denotes the set of n independent variables, and vector b denotes maximum likelihood estimated parameters.

In the logit model, the dependent variable is the natural log of the odds of location, and the best information we can gain from the estimates is their relative importance. Furthermore, given change in explanatory variables will have a different impact on location probability, depending on their estimated probabilities. Observations with probabilities close to 0 or 1 will be much less affected by change in explanatory variables than observations with probabilities close to 0.5.

The model examined in this study is of the following form:

$$Z_j = f(L_j, T_j, S_j) + e, \text{ where}$$

Z_{ij} = dichotomous variable equal to 1 new establishment located in county j and 0 otherwise,

L_j = a vector of location-specific economic factors (rates of labor force participation and unemployment, median home value, per capita income),

T_j = transportation infrastructure (major arterial mileage, number of railroads in a county, airline dummy, intermodal connector dummy, average time to work, average daily traffic in the county), and

S_j = social indicators (population density, county size, etc.).

The primary hypothesis in the study is that the dependent variable new firm (NF) in any county during a specified period of time is a function of several independent or explanatory variables. A cross-section database on establishment entries into Midwest counties was the basis of analysis. The dependent variable is a dichotomous variable which has value 1 if the county attracted one or more manufacturing firms in period 1998-2002 and a value of 0 if there was no new firm established. This time period is used because it is recent and therefore is more relevant than earlier periods. Also, this time period is long enough to allow a reasonable number of new businesses to open. In addition, separate analyses were performed for different firm sizes.

Entering firms were defined as those that have been established in the last 5 years and have an employment size equal to 10 or more. This low employment limit is chosen because it assumes that 1) small businesses are an important ingredient in rural economy and 2) smaller size companies are more likely locally owned.

The data were separated into two groups by number of employees in the entering firm. Small firms consider less location alternatives; thus, it is likely that they will be less affected by major differences in county attributes (Fox and Murray, 1990). Two separated models were run to investigate the difference in influence of transportation factors on business location decisions of firms with different sizes. The dependent variable for the first model was defined as a dichotomous variable which has value 1 if the county attracted a new firm with more than 10 but less than 50 employees and 0 otherwise. The second model has a dependent dichotomous variable which has value 1 if the county attracted a new firm with more than 50 employees and 0 otherwise.

The study was concerned with impact of transportation on location decisions of companies in rural areas; thus, counties with one or more metropolitan statistical areas (MSA) or small metropolitan statistical areas (SMSA) were eliminated from the dataset. According to the United States Office of Management and Budget Standards (2003), each metropolitan statistical area must have at least one urbanized area of 50,000 or more inhabitants. Using this definition, there were 543 rural counties used in the study area.

SM is a dichotomous variable which has value 1 if the county attracted 1 or more manufacturing firms with less than 50 employees in period 1998-2002 and a value of 0 if there was no new firm established

LG is a dichotomous variable which has value 1 if the county attracted 1 or more manufacturing firms with more than 50 employees in period 1998-2002 and a value of 0 if there was no new firm established

Employment

Although the dichotomous variable logit model allows estimating the probability of a new firm to locate in the particular region, it does not estimate the number of new jobs created in the county. Hence, an additional model was used to define the correlation between transportation factors and employment in the counties.

The multiple regressions model was used in this study to evaluate the impact of transportation on employment/economic development of the region. The multiple regression technique measures the amount of variation in a dependent variable which is explained by independent variables. Ordinary least-squares estimation was originally planned to be used to obtain the estimates.

One of the key assumptions of regression is that the variance of the errors is constant across observations. If the errors have constant variance, the errors are called homoscedastic. Standard estimation methods are inefficient when the errors are heteroscedastic or have non-constant variance. The test that can help to establish whether the heteroskedasticity exists or not was proposed by Breusch and Pagan (Griffiths et al., 1993).

Another problem that might arise during the estimation is collinearity or multicollinearity of explanatory variables. Collinearity and multicollinearity might cause diminished ability of tests to find significant results. The simplest way to identify collinearity is an examination of the correlation matrix for the predictor variables. The presence of high correlations (generally .90 and above) is the first indication of substantial collinearity, but collinearity may be due to the combined effect of two or

more other independent variables, and the correlation matrix shows only simple correlations between two variables.

Two measures for assessing both pair-wise and multiple variables collinearity are the tolerance and the variance inflation factor (VIF). Tolerance values approaching zero indicate that the variable is highly collinear with the other predictor variables. Large VIF values (a usual threshold is 10.0) indicate a high degree of collinearity or multicollinearity among the independent variables.

The dependent variable in this model refers to the number of employees in counties. For the first model, the dependent variable was presented by the total number of employees (EMPLT), and for the second model, the dependent variable was presented by the number of employees in the manufacturing sector in each county (EMPLM).

The data for this variable were obtained from the Bureau of Economic Analysis local area personal income database.

Independent variables

Transportation factors

The major question addressed in this analysis is whether or not transportation infrastructure influences business location decision.

Highway mileage is a measure of the transportation network and thus the cost and availability of materials inputs. Access to the major highways may reduce transportation costs for the firm, which will lead to higher profits. Therefore, highway

variables are assumed to have a positive impact on business location decisions and employment rate in the county.

Three highway mileage variables that might have a significant impact on economic development in the county were identified. These are

- 1) Lane miles of interstate,
- 2) Lane miles of other principal arterial, and
- 3) Total lane miles in the county.

The term lane miles was determined as the length of a roadway or roadway system multiplied by the number of lanes that are open for traffic.

LMI is lane miles of interstate,

LMO is lane miles of other principal arterials, and

LM is total lane miles in the county.

Airline variable was defined as 1 if the county contains one or more major airports and 0 otherwise. Access to the airports means reduction in business travel costs and air freight cost. Therefore, availability of the airport in the county providing service in 1998-2001 is expected to have a positive effect on business location decisions of the manufacturing firms and total and manufacturing employment rates.

AIR is existence of airport in the county.

Railroad variable was presented as a dichotomous variable which has a meaning 1 if the railroad is presented in the county and 0 otherwise. Railroad service is essential for attracting manufacturing companies in the rural areas (Goode and Hastings, 1985). Thus, it is expected that effect of railroad existence on business location and employment rate in the county will be positive.

RAIL is existence of railroad in the county.

Companies usually seek the place to locate which is close to the major markets (Fox and Murray, 1990). By locating close to the market, companies reduce their transportation costs and hence increase their profits.

Access to the markets variable was presented as distance from county center to the nearest Metropolitan Statistical Area. The information on county centers was obtained from the National Association of Counties (2003).

Cities with a population of 400,000 or more people and located in participating research states or in nearby states were considered as potential nearest Metropolitan Statistical Areas. The list of major cities used for this variable is presented in Table 3.1. For each of the nine selected states in the study, the distance from potential MSA to county center was estimated. The smallest distances were used in the model estimation.

DIS is distance to the nearest MSA. This variable is expected to have a negative correlation with business location decisions and employment rate.

TIME variable represents average time to work (in minutes) that a person spends to get to his or her job. The time that a person spends for getting to the work place was assumed to be negatively associated with business locations and employment in the county. Traveling time will depend on distance to the place of employment, road conditions, and traffic on the roads.

AVTIME is average time traveled from home to the place of work.

Table 3.1. The nearest Metropolitan Statistical Areas for the states used in the study.

State	Closest Major Cities
Colorado	Denver
Iowa	Omaha, Minneapolis, Chicago
Kansas	Kansas City, Denver
Minnesota	Minneapolis
Montana	Seattle, Denver
Nebraska	Omaha, Denver
North Dakota	Minneapolis
South Dakota	Denver, Omaha, Minneapolis
Wyoming	Denver

Economic factors

Per capita income variable is used as a measurement of consumer tastes and purchasing power (Fox and Murray, 1990).

PCPI is per capita personal income in the county, 2000.

Access to labor is critical to a locating firm. Only a few firms will be indifferent to labor considerations. Locating in the particular county, companies want to be assured that an adequate labor supply exists. Labor characteristics can be divided into three categories:

- 1) Labor availability, which shows availability of potential employees to the firm;
- 2) Labor quality, which represents productivity of labor; and
- 3) Labor cost.

Firms tend to seek locations with a well-trained labor force. The first factor that affects the quality of labor force is education. A high-quality work force is usually more productive and leads to lower production and training costs for the new firm. In this research, the skilled labor force is presented by the number of people enrolled in college or graduate school. The population density variable represents labor availability in the county.

COLL is percentage of people enrolled in college or graduate school, 2000.

POPDEN is population of the county divided by area square miles, 2000.

Social factors

The following variables represent qualitative attributes of the county that may be important to the industry location decision.

RENT is median rent asked in the county, 2000.

AREA is county area in square miles.

The importance of quality of life factors is not well defined in economic literature. It is hard to recognize the importance of these factors because one person's idea of life quality may differ from another's.

In this study, quality of life factors are represented by two variables, such as means of transportation to work (percentage of people who used car, truck, or van to get to work) and unemployment insurance benefit payments.

CAR is % of people who used car, van, or truck as transportation to the work place.

INSUR is unemployment insurance benefit payments, 2000.

A summary of independent variables used in the models is given in Table 3.2.

Data

The study area includes the Northern Great Plains states of Colorado, Iowa, Kansas, Minnesota, Montana, Nebraska, North Dakota, South Dakota, and Wyoming. The units for analysis were all counties from the study area. Using America's Labor Market Information System, Employer Database file (2003), a dataset on new establishments in each industry was created.

Table 3.2. Predicted estimates of explanatory variables.

Parameter	Variable	Sign to expect			
		Total Employment	Manuf. Employment	Location <50 empl.	Location >50 empl.
RAIL	Railroad availability	+	+	+	+
AIR	Airport availability	+	+	+	+
DIS	Distance to market	-	-	-	-
LMI	Interstate	+	+	+	+
LMO	Other principal arterial	+	+	+	+
LM	Total lane miles	+	+	+	+
INSUR	Unempl. insurance benefit payments	+	+	+	+
PCPI	Per capita personal income	+	+	+	+
CAR	Car transportation	+	+	+	-
AVTIME	Av. time to work	-	-	-	-
RENT	Median rent	-	-	-	+
COLL	College enrolment	+	+	+	+
AREA	Total area	+	+	+	+
POPDEN	Population density	+	+	+	+

America's Labor Market Information System, Employer Database, provides information on location, employment, sales, and primary four-digit SIC codes for every employer in the U.S.A.

Data for highway variables were obtained from Highway Performance Monitoring System Core data. This dataset consists of a complete inventory of length by functional system, jurisdiction, geographic location (rural, small urban, urbanized), intermodal connectors, and other selected characteristics. SAS programming was used in order to calculate the length of lane miles and identify intermodal connectors for each county of the study region.

Data for the distance to the nearest MSA variable were collected from the MapQuest Driving Directions (www.mapquest.com), which allowed finding the shortest way from county center to nearest MSA.

Data for the railroad variable were obtained from Bureau of Transportation Statistics (BTS) Geographic Information Services (GIS). Using GIS maps, the information on availability of the railroads in the county was obtained.

Data for the airport variable were obtained from the Bureau of Transportation Statistics, The Intermodal Transportation Database. The dataset contains a list of domestic and foreign airport codes and their associated city codes, world area codes, city or airport names, state or country names, and latitude and longitude information. The information about domestic airport codes and associated cities was collected first, and then the cities were assigned to their counties.

Economic characteristics of the counties were mostly obtained from the Bureau of Economic Analysis Local Area Personal Income file for 2000. Social attributes of

the counties were collected from the U.S. Census Bureau, Summary File 3 sample data for 2000.

Summary

This chapter has established the statistical models for attaining two research objectives. An empirical location model has been proposed which uses logit analysis to explain business locations during 1998-2002 in rural Midwest counties. This model can also estimate the probability of a jurisdiction attracting businesses.

An empirical model of the impact of transportation factors on total and manufacturing employment in the county also has been proposed. The dependent variables are total and manufacturing employment at the county for the year 2000.

CHAPTER IV. RESULTS

Introduction

This chapter reports results, presents analysis of the statistical models and hypothesis tests of Chapter III, and offers interpretations. The chapter consists of three sections. The first section includes collinearity diagnostics of independent variables. In the second section, results of the employment model are presented and discussed, and in the third section, results and discussion of the location model are presented.

Nineteen variables were considered for inclusion in employment and firm location models. Models with many independent variables are likely to have collinearity; therefore, estimation of location and employment models should include collinearity tests. There is a total of four model runs; two models were run for location decisions of manufacturing companies with different sizes (more or less than 50 employees), and two models were run for total employment and manufacturing employment in the county.

Collinearity diagnostics

Collinearity indicates that there is a strong linear relationship between different variables. The presence of collinearity or multicollinearity between independent variables might have inefficient effects on multiple regressions, such as estimation problems and misinterpretation of model results. In particular, collinearity affects parameter estimates in their standard errors and, consequently, **t** ratios. In other words, regression coefficients are not stable, and tests to find significant results are not robust.

One of the ways to detect collinearity between variables is to examine the correlation matrix for the predictor variables. A high value (usually equal or above .90) of the correlation coefficients indicates high correlation between two independent variables to which it refers.

The variance-covariance matrix of the variables used in this research showed several high correlations between predictor variables. To avoid the multicollinearity problem, one of the correlated variables from each pair was omitted from the model. The variance-covariance matrix of the primary considerable variables is presented in Appendix A.

The variance-covariance matrix showed that only one pair of variables (total area–average time to work) had high correlation. Those variables were believed to have statistical significance in the model and were not excluded from further estimation. Instead, additional tests were made to assure that variables included in final models are not collinear; in particular tolerance and variance inflation factors were estimated.

The results for the collinearity diagnostic (tolerance and variance inflation factors) are presented in Table 4.1. None of the variables had a tolerance factor close to 0, and none of the variables had a variance inflation factor more than 10, or in other words, tolerance and variance inflation factors indicated that none of the variables had high correlation with other variables. There is no evidence of a significant collinearity in the model.

Collinearity diagnostics were performed by using SAS programming, and codes are presented in Appendix B.

Table 4.1. Collinearity diagnostics results.

Label	Variable	Parameter Estimate	Standard Error	Pr > t	Tolerance	Variance Inflation
Intercept	Intercept	-906.3018	831.0135	0.2761	.	0
RAIL	Availability of the railroad	-114.7114	230.3800	0.6188	0.83355	1.19969
AIR	Availability of the airport	475.8614	132.8192	0.0004	0.62953	1.58849
DIS	Distance to major city	-0.93594	0.36521	0.0107	0.64531	1.54965
LMI	Interstate	1.59143	0.85536	0.0635	0.56525	1.76911
LMO	Other principal arterial	1.95325	0.95472	0.0414	0.40922	2.44369
LM	Total lane miles	0.13802	0.0585	0.0188	0.56324	1.77544
INSUR	Unemployment insurance benefit payments	0.36527	0.07638	<.0001	0.19526	5.12141
PCPI	Per capita personal income	-0.0015	0.0152	0.9213	0.5503	1.81718
CAR	Car transportation	2105.328	919.7365	0.0226	0.64835	1.54237
AVTIME	Av. time to work	-46.88441	18.01595	0.0096	0.66832	1.49629
RENT	Median rent	0.89007	0.63158	0.1595	0.46259	2.16176
COLL	College enrolment	921.1864	1187.600	0.4384	0.81549	1.22626
AREA	Total area	-0.2695	0.06376	<.0001	0.4187	2.38832
POPDEN	Population density	17.32	2.10652	<.0001	0.24009	4.16506

A description of the variables included in the model was presented in Chapter III. The simple statistics of the independent variables included in the final models are presented in Table 4.2

Table 4.2. Statistic characteristics of explanatory variables.

Parameter	Variable	Mean	Standard deviation	Min	Max
RAIL	Availability of the railroad	0.93868	0.2402	0	1
AIR	Availability of the airport	0.35613	0.47942	0	1
DIS	Distance to major city	264.1791	172.20834	22.19	901.85
LMI	Interstate	38.15016	78.56214	0	567.95
LMO	Other principal arterial	104.38705	82.7241	0	631.1
LM	Total lane miles	2365	1151	0	7706
INSUR	Unemployment insurance benefit payments	1072	1497	0	22370
PCPI	Per capita personal income	22975	4487	5475	69960
CAR	Car transportation	0.84721	0.06822	0.48438	0.94076
AVTIME	Av. time to work	16.68397	3.43034	8.88158	34.62005
RENT	Median rent	305.22406	117.61451	0	1225
COLL	College enrolment	0.0489	0.04711	0.00649	0.38672
AREA	Total area	1270	1225	376.82	10491
POPDEN	Population density	20.72217	48.94747	0.3	946.1

Results of empirical location model

The empirical model for estimation transportation effect on business location described in Chapter III specifies 14 different variables. The total of 424 observations was used for the models estimation. During the 5-year period, 36.27% of the localities attracted manufacturing plants with less than 50 employees, and 17.49% of the counties

attracted manufacturing companies with more than 50 employees. The summary of entering manufacturing companies is given in Table 4.3.

Table 4.3. Selected descriptive statistics of manufacturing firm entries: 1999-2004.

State	% counties 10-50	% counties 50 and more
Colorado	41.30	8.51
Iowa	53.16	30.37
Kansas	37.02	17.44
Minnesota	51.51	33.33
Montana	26.92	7.69
Nebraska	29.76	15.47
North Dakota	16.32	14.28
South Dakota	18.64	6.77
Wyoming	57.51	9.52
Total	36.27	17.49

Logit analysis

The results of the logit model present the information that might be interesting to economic development specialists who are interested in attracting new manufacturing firms. Estimations were made using SAS programming; the codes for these model estimations are given in Appendix C. Table 4.4 presents results for the location of new manufacturing companies with less than 50 employees, and Table 4.5 presents results of the estimation location model of manufacturing companies with more than 50 employees. The estimated coefficients, standard errors, and chi-squares and the significance level of each variable used are provided in that table.

Table 4.4. Estimated effects of location factors on new manufacturing firm's locations with less than 50 employees.

	Variable	Estimate	St Error	Chi-Square	P value
Intercept	Intercept	-4.2317	2.3297	3.2994	0.0693*
RAIL	Availability of the railroad	0.4442	0.5874	0.5719	0.4495
AIR	Availability of the airport	0.2596	0.3389	0.5865	0.4438
DIS	Distance to major city	-0.0017	0.0009	3.4813	0.0621*
LMI	Interstate	0.0068	0.0022	9.2334	0.0024***
LMO	Other principal arterial	0.0050	0.0024	4.1724	0.0411**
LM	Total lane miles	-0.0001	0.0001	0.4777	0.4895
INSUR	Unemployment insurance benefit payments	0.0005	0.0003	2.7998	0.0943*
PCPI	Per capita personal income	0.0000	0.0000	0.6422	0.4229
CAR	Car transportation	3.7514	2.4822	2.2840	0.1307
AVTIME	Average time traveled to the place of work	-0.0937	0.0428	4.7830	0.0287**
RENT	Median rent	0.0018	0.0018	1.0235	0.3117
COLL	College enrolment	-6.6648	3.1719	4.4150	0.0356**
AREA	Total area	0.0003	0.0002	1.6188	0.2033
POPDEN	Population density	0.0636	0.0215	8.7595	0.0031***
<p>* significant at 10% level. ** significant at 5% level. *** significant at 1% level.</p>					

Table 4.5. Estimated effects of location factors on new manufacturing firm's locations with more than 50 employees.

	Variable	Estimate	St Error	Chi-Square	P value
Intercept	Intercept	-7.9495	3.8136	4.3452	0.0371**
RAIL	Availability of the railroad	0.6898	1.1120	0.3848	0.5351
AIR	Availability of the airport	-0.3370	0.3738	0.8129	0.3673
DIS	Distance to major city	-0.0017	0.0012	1.8643	0.1721
LMI	Interstate	0.0032	0.0026	1.5394	0.2147
LMO	Other principal arterial	0.0041	0.0029	2.0246	0.1548
LM	Total lane miles	0.0000	0.0002	0.0001	0.9931
INSUR	Unemployment insurance benefit payments	0.0004	0.0002	4.7445	0.0294**
PCPI	Per capita personal income	0.0000	0.0001	0.1754	0.6754
CAR	Car transportation	6.2021	3.9088	2.5177	0.1126
AVTIME	Average time traveled to the place of work	-0.0714	0.0558	1.6350	0.201
RENT	Median rent	0.0001	0.0019	0.0032	0.9549
COLL	College enrolment	1.8186	3.1582	0.3316	0.5647
AREA	Total area	-0.0001	0.0002	0.1764	0.6745
POPDEN	Population density	0.0315	0.0134	5.5523	0.0185**
<p>* significant at 10% level. ** significant at 5% level. *** significant at 1% level.</p>					

In the first model for the manufacturing companies with less than 50 employees, several variables showed significant results. In the second model, almost all variables did not show any significance.

Distance to the nearest MSA is negatively and significantly (at 10% level) associated with location of new manufacturing firms with less than 50 employees. Railroad and airport availability did not have any significance in the first model. Interstate and other principal arterials had a positive and significant impact on the location of manufacturing companies with less than 50 employees. Interstate mileage (LMI) variable was significant at the 1% level, and the other principle arterials (LMO) variable was significant at the 5% level. These results indicate that larger manufacturing companies do not rely on the interstates and other principal arterials. Other transportation variables did not show any significance.

Two other variables are significant in the location model for manufacturing companies with less than 50 employees. First, college enrollment showed a negative significant relationship at the 5% level. This result might indicate that manufacturing companies in making location decisions consider the number of residents enrolled in college as a lack of labor force in a particular county.

Population density variable had a positive significant influence in the model, which was expected because this variable represents labor availability in the county. This variable also has a positive influence for larger manufacturing companies.

For manufacturing companies with more than 50 employees, 2 variables were found to have a significant impact.

Unemployment insurance benefit payments (INSUR) showed a positive significant relationship at the 10% level. This variable represents quality of life factors, so it was expected to have a positive sign, and as it was mentioned before, the population density variable had a positive significant influence on business location decisions at the 5% level.

All other variables specified in both models did not have a significant influence on the location decisions of new manufacturing firms.

Probability estimation

To interpret location model results, the probability of location of new firms should be estimated. Under the logit analysis, the dependent variable is a natural log of the probability of location

$$Z_i = \ln(P_i/(1-P_i)),$$

where P_i is a probability of a new firm location, and Z_i is the dichotomous dependent variable, specified earlier. Taking the exponential of both sides of this equation and transforming it, the probability of location of new firm can be expressed by the following equation:

$$P_i = \exp(Z_i)/(\exp(Z_i) + 1).$$

By using this formula, it is possible to calculate the change in probability for one unit change in an independent variable. First, mean probability is calculated, setting all variables at its mean and giving 0 values to RAIL and AIR variables. Second, new probabilities for each variable were calculated by adding to each variable additional unit, ceterus paribus. After that, new probabilities are estimated, and the change in the

probabilities is calculated. For example, to find change probability for interstate variable, Z_i mean value was calculated, and then mean of interstate variable was increased by one, while everything else was left the same. New Z_i and new P_i values are calculated, and the difference from mean probability and new P_i value is taken, which represents the change in probability of location manufacturing firm with 1 unit change in interstate lane mileage. The results of location probabilities estimated are given in Tables 4.6 and 4.7.

The location model is able to generate the predicted probabilities of plant location for each of the counties in the dataset. The predicted probabilities were calculated for each county, and the results were compared with observed location variables. If the county had the predicted probability to locate in the county more than 50%, the value 1 was assigned to it, and if the county had the predicted probability less than 50%, the value 0 was assigned to the county. Overall the model was able to correctly predict 78.69% of the observations for the manufacturing location model with less than 50 employees and 84.91% of the observations for the manufacturing location model with more than 50 employees.

The graphical representation of change in probability with one unit change in explanatory variable for manufacturing companies with less than and more than 50 employees is given in Figures 4.1 and 4.2, respectively.

Results of empirical employment model

The manufacturing employment variable had missing values in several counties, so adjustments were made and counties that had missing observations were deleted

Table 4.6. The effect of one unit change on probability of new manufacturing firm with less than 50 employees' location.

	Variable	Z	Probability	Change in probability
	Mean	0.6285	0.6522	
RAIL	Availability of the railroad	1.0727	0.7451	0.0929
AIR	Availability of the airport	0.8881	0.7085	0.0564
DIS	Distance to major city	0.6268	0.6518	-0.0004
LMI	Interstate	0.6354	0.6537	0.0016
LMO	Other principal arterial	0.6335	0.6533	0.0011
LM	Total lane miles	0.6284	0.6521	-0.0000
INSUR	Unemployment insurance benefit payments	0.6290	0.6523	0.0001
PCPI	Per capita personal income	0.6286	0.6522	0.0000
CAR	Car transportation	0.6661	0.6606	0.0085
AVTIME	Average time traveled to the place of work	0.5348	0.6306	-0.0216
RENT	Median rent	0.6304	0.6526	0.0004
COLL	College enrolment	0.5619	0.6369	-0.0152
AREA	Total area	0.6288	0.6522	0.0000
POPDEN	Population density	0.6921	0.6664	0.0142

Table 4.7. The effect of one unit change on probability of new manufacturing firm with more than 50 employees' location.

	Variable	Z	Probability	Change in probability
Intercept	Intercept	-2.1330	0.1059	
RAIL	Availability of the railroad	-1.4433	0.1910	0.0851
AIR	Availability of the airport	-2.4701	0.0780	-0.0279
DIS	Distance to major city	-2.1347	0.1058	-0.0002
LMI	Interstate	-2.1298	0.1062	0.0003
LMO	Other principal arterial	-2.1290	0.1063	0.0004
LM	Total lane miles	-2.1331	0.1059	0.0000
INSUR	Unemployment insurance benefit payments	-2.1326	0.1060	0.0000
PCPI	Per capita personal income	-2.1330	0.1059	0.0000
CAR	Car transportation	-2.0710	0.1119	0.0060
AVTIME	Average time traveled to the place of work	-2.2045	0.0994	-0.0066
RENT	Median rent	-2.1330	0.1059	0.0000
COLL	College enrolment	-2.1149	0.1077	0.0017
AREA	Total area	-2.1331	0.1059	-0.0000
POPDEN	Population density	-2.1016	0.1090	0.0030

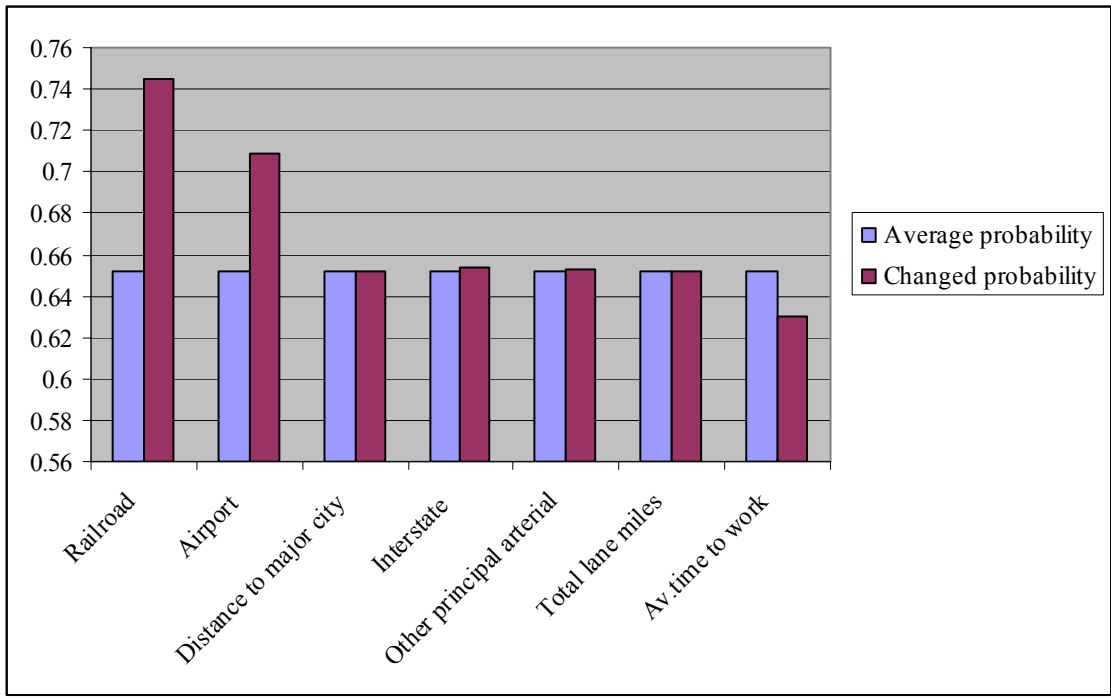


Figure 4.1. Change in location probability for manufacturing companies with less than 50 employees.

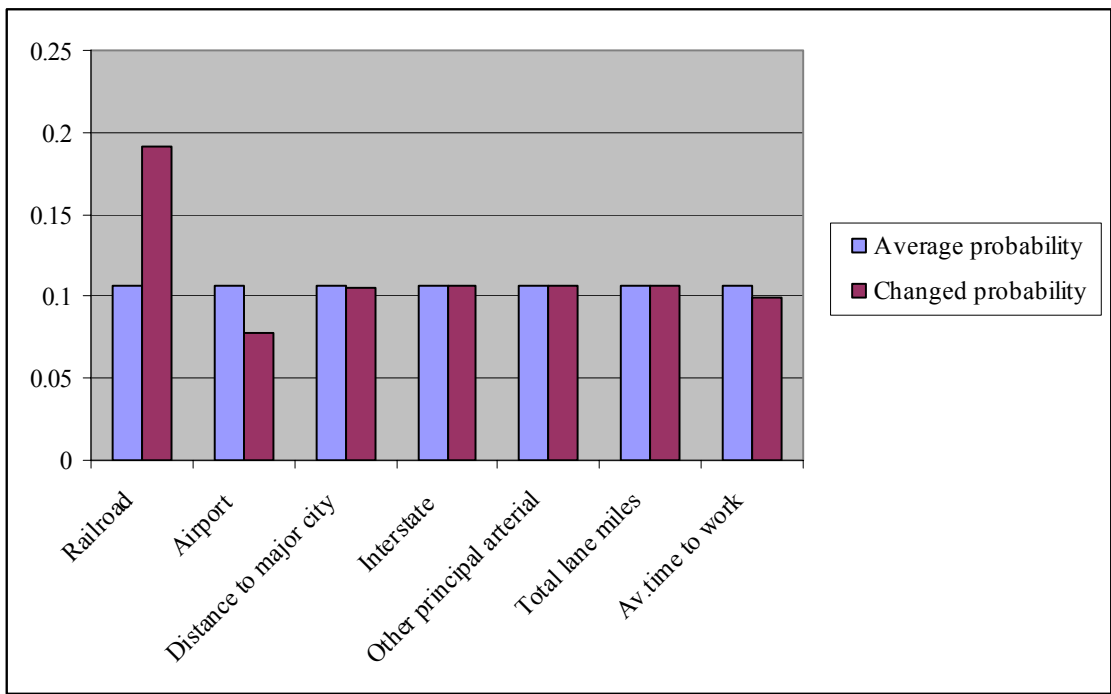


Figure 4.2. Change in location probability for manufacturing companies with more than 50 employees.

from the model. Also, counties which have one or more SMSA or MSA were excluded from the dataset. Overall, the total employment model and manufacturing employment model contain 424 observations.

By analyzing the set of independent variables, it was concluded that the per capita personal income variable might be jointly determined with employment variables, or in other words, those variables are endogenous. Including an endogenous factor as an explanatory variable in an OLS model may create problems for accurately understanding relationships of interest because of the possibility of biased results.

One of the possible ways to solve the endogeneity problem is using the instrumental variable estimator. Three variables, average wage in the county, median rent asked, and percentage of people with a higher degree, were selected as instrumental variables. The Hausman test was performed to determine if the OLS model is inconsistent and whether or not the instrumental variable estimator should be used. The results of Hausman test are presented in Table 4.8. The value of Hausman test in the total employment model is less than 3.84, the 5% critical value from a $\chi^2_{(1)}$ distribution; hence, it was concluded that the least square estimator is consistent, and there is no contemporaneous correlation between total employment and per capita personal income. Also, the Hausman test indicated that in the manufacturing employment model, the per capita personal income variable is endogenous, and the instrumental variable estimator should be used. The code for the instrumental variable model described above is presented in Appendix D.

The Breush-Pagan test (Griffiths, 1993) was used to check the heteroskedasticity of the error term for the total employment model. Results of the test

Table 4.8. Hausman test results.

	Hausman test	Conclusion
Total employment	1.9337	OLS
Manufacturing employment	4.7298	2SLS

are presented in Table 4.9. The results of the BP test were compared with critical values, and it was concluded that the total employment model had the heteroskedusticity of error term. Thus, the variance correction model was used for this model estimation. The codes for the heteroskedusticity and variance correction model described above are presented in Appendix E and Appendix F, respectively.

Table 4.9. The results of Breusch-Pagan test.

Model	Number of observations	BP test	Critical value	P value
Total Employment	424	535.115	124.342	0.0000

The results for the total employment and manufacturing employment models are presented in Tables 4.10 and 4.11, respectively. The estimated coefficients, standard errors, and t-values of each variable used are provided in that table.

Availability of the airport in the county variable has a positive significant influence on the total and manufacturing employment rate in rural counties of the Midwest region. Distance to the major MSA variable in both models showed a negative and significant correlation with the manufacturing employment rate and has no significance in the total employment model.

Table 4.10. Estimated effects of location factors on total employment.

	Variable	Parameter	St. Error	T - value
Intercept	Intercept	-4138.56	2938.985647	-1.316361
RAIL	Availability of the railroad	-990.528	873.6737899	-1.136465
AIR	Availability of the airport	1764.794	467.4878479	3.512104**
DIS	Distance to major city	1.294277	1.751945831	0.936729
LMI	Interstate	13.1811	4.094089323	4.073201**
LMO	Other principal arterial	12.42288	4.388310038	3.439381**
LM	Total lane miles	0.911281	0.251434652	4.117220**
INSUR	Unemployment insurance benefit payments	1.680613	0.531761256	5.816314**
PCPI	Per capita personal income	0.111109	0.052511318	1.934779*
CAR	Car transportation	-5592.44	3028.995824	-1.607209
AVTIME	Av. time to work	-79.3949	68.98613601	-1.164848
RENT	Median rent	13.66227	3.465300382	5.717804**
COLL	College enrolment	7634.703	7081.098837	1.699245*
AREA	Total area	0.924845	0.326626363	3.834291**
POPDEN	Population density	316.9302	16.49168505	39.767862
* significant at 10% level (critical value = 1.67). ** significant at 5% level (critical value = 2.04). $R^2_{adj} = 0.9684$.				

Table 4.11. Estimated effect of location factors on manufacturing employment.

	Variable	Parameter	St. Error	T - Value
Intercept	Intercept	-	1070.209	-2.52**
RAIL	Availability of the railroad	2700.320000	254.7207	-0.41
AIR	Availability of the airport	-105.346000	146.809	3.25**
DIS	Distance to major city	476.791500	0.403858	-2.54**
LMI	Interstate	-1.026790	0.954926	1.25
LMO	Other principal arterial	1.194062	1.058114	1.64
LM	Total lane miles	1.735133	0.063862	2.44**
INSUR	Unemployment insurance benefit payments	0.156073 0.341369	0.084195	4.05**
PCPI	Per capita personal income		0.04265	2.98**
CAR	Car transportation	0.126954	1020.755	1.57
AVTIME	Av. time to work	1606.262000	19.14775	-2.17**
RENT	Median rent	-41.630900	0.962362	-1.52
COLL	College enrolment	-1.458320	1336.992	-0.06
AREA	Total area	-80.386200	0.070523	-3.92**
POPDEN	Population density	-0.276700 16.772060	2.301353	7.29**
* significant at 10% level (critical value = 1.67). ** significant at 5% level (critical value = 2.04). $R^2_{adj} = 0.7594$.				

The interstate and other principal arterials variables are shown to have a positive and statistically significant influence on total employment in rural counties. For manufacturing employment, these variables also showed a positive influence, but the variables are not significant. These results might indicate that for mileage of road, principal arterials are more important for the total employment in the county than for the manufacturing industry.

The total lane miles variable was positively associated with total and manufacturing employment. According to these results, it might be concluded that rural roads play an important role in economic development in rural counties of the Upper Great Plains region.

The average time to work variable was assumed to have a negative impact on employment, and results indicated that this variable is statistically significant and negatively associated with manufacturing employment. This variable is also negatively associated with total employment, but it did not have any significance in the model.

The per capita personal income variable had a positive and statistically significant influence on both total and manufacturing employment. As it was expected, the unemployment insurance benefits variable is positively associated with both total and manufacturing employment rate in rural counties of the Midwest. College enrollment variable had a positive influence on total employment in the county, but did not show any significance in the manufacturing employment model.

Surprising results were obtained for the median rent variable. It was expected that this variable would be negatively associated with total and manufacturing employment, but results indicated that this variable had a positive influence on the total

employment in the county and did not have any significance in the manufacturing employment model.

Summary

The influence of transportation variables on total employment, manufacturing employment, and business location for different sized firms was estimated. First, several tests detecting collinearity were made. According to results of the tests, 14 independent variables were selected for final model runs.

Second, business location models for manufacturing companies with different employment sizes were estimated. Interstate and other principal arterials variables were positive and statistically significant for manufacturing companies with less than 50 employees. Distance to the nearest MSA was negatively correlated with location decisions of manufacturing companies with less than 50 employees, and finally, the average time to work variable showed a negative and statistically significant correlation with the location of manufacturing companies with less than 50 employees. Other transportation variables did not show any significance in both models. Then, probabilities of new firm location were calculated, and change in probability with unit change in one variable was estimated. At last, predicted probabilities for location of new firms were calculated, and according to those results, current models predicted 79% of a firm's location with smaller size and 85% of a firm's location with bigger size.

Third, total employment and manufacturing employment models were run. Interstate and other principal arterials variables showed significant correlation with

total employment in the county, and total lane miles showed a positive significant relation with both total and manufacturing employment in the county. Airport availability has a positive effect on the total and manufacturing employment rate, and distance to the nearest MSA has a negative effect on the manufacturing employment rate. Other transportation variables were not significant.

General conclusions, limitations of the study, and propositions for future research will be discussed in Chapter V.

CHAPTER V. SUMMARY AND CONCLUSIONS

Introduction

Chapter V provides an overview of the thesis, a summary of the procedures used, and conclusions drawn from the results. Limitations of the study are discussed, and directions for future research are proposed.

Thesis summary

Four different models were developed to evaluate the impact of transportation on business location decisions and the impact of transportation on employment growth. Total numbers of 14 variables were used in the models' estimation.

The location of new manufacturing companies, employment data, and transportation factors from 424 non-metropolitan counties were analyzed, and possible interactions between location decisions, employment, and transportation were investigated. The study used regression analysis to identify the factors associated with business location and employment growth.

Four different models were developed to evaluate the impact of transportation on business location decisions and the impact of transportation on employment growth. Total numbers of 14 variables were used in models' estimation.

The dichotomous dependent variable logit model was used to determine location factors which influence business location decisions. The maximum likelihood estimation procedure was used to estimate logit parameters. Probabilities of new firms'

location decisions were calculated as well as changes in probabilities with unit change in one variable, *ceterus paribus*.

The ordinary least-squares model was used in this study to evaluate the impact of transportation on employment rate of the study region. The Hausman test was performed to determine if the per capita income variable is jointly determined with the dependent variable in total and manufacturing employment models. The results indicated that for the manufacturing employment model, OLS estimation is inconsistent and the instrumental variable estimator should be used. For the total employment model, the results of the Breusch-Pagan test showed the heteroskedusticity of the error term, so the variance correction mode was used for the estimation.

Data for this research were collected for 645 counties from the Northern Great Plains states of Colorado, Iowa, Kansas, Minnesota, Montana, Nebraska, North Dakota, South Dakota, and Wyoming. Counties containing SMSA or MSA were excluded from the study, and a total number of 424 observations was used for the analysis.

Data on new establishments were collected from America's Labor Market Information System, Employer Database file (2003). Data for transportation variables were obtained from Highway Performance Monitoring System Core data. Economic characteristics of the counties were mostly obtained from the Bureau of Economic Analysis Local Area Personal Income file for 2000. Social attributes of the counties were collected from the U.S. Census Bureau, Summary File 3 sample data for 2000.

Results and conclusions

Three out of four hypotheses of the study were supported by this analysis. First, the transportation factors were significant for total and manufacturing employment in the Upper Great Plains. Second, the results of the estimated logit model for new manufacturing establishments showed that influence of transportation differs for companies with less than 50 employees and companies with more than 50, and finally, factors which have been associated with total employment in the county were different from those associated with manufacturing employment.

Location models

The significant finding of this research was that only a few variables showed significance in logit location models. A strong and positive relationship was found for interstate mileage and other principal arterials variables in the model for manufacturing companies with less than 50 employees. It was concluded that smaller companies are more likely to locate near major principal arterials than larger companies. Distance to the nearest MSA showed a negative impact on business location decisions for the small manufacturing firms, which means companies are looking for the location place not so far from the major market. All other transportation variables specified in both models did not have a significant influence on the location decisions of new manufacturing firms.

After logit model estimation, predicted probabilities of locating new manufacturing firms were calculated for each county, and the results were compared with observed location variables. Overall, the model was able to correctly predict 79%

of the observations for the manufacturing location model with less than 50 employees and 85% of the observations for the manufacturing location model with more than 50 employees.

Employment models

For the total employment model, several transportation variables, like interstate and other principal arterials mileage, showed a positive significant relationship. In the manufacturing employment model, those variables did not have any statistical significance. A variable total lane mile was found to have a strong positive impact on total and manufacturing employment in the Midwest region. These results suggest that manufacturing companies in rural areas are more concerned about roads in general than about availability of interstates or other principal arterials.

The results indicated that existence of the airport in the county had a positive and significant impact on total and manufacturing employment in the Great Plains. Distance to the nearest MSA variable was, as expected, negatively and significantly associated with manufacturing employment in the county. Other transportation variables did not show any significance in the models.

The main conclusion that can be made according to the results of this research is that improvements in transportation infrastructure would benefit the economic situation in the counties of the rural Midwest. However, if the community's main goal is to attract large manufacturing companies, then it must take the risk of not receiving any benefits from investments in transportation infrastructure.

Limitations and future research

First, due to the data limitations, transportation infrastructure variables were presented as a mileage of roads of different functional systems. Although these variables in some way represent transportation investment results, it is not possible to determine where the changes were made and if the changes made any difference in the employment rate or in attractiveness of the county for new manufacturing firms.

The available data allowed identifying the year when the company was established in a particular county but did not allow identifying whether the company was newly created in this county, moved within the county, or moved from a different county.

Finally, the great uncertainty exists in determining whether transportation investments lead to economic development in the rural counties or if it is the other way around.

Limitations of this study indicate implications for future research. First, although these findings provided some explanations of the relationship among transportation factors and employment and business location decisions, several questions remain unanswered and need to be investigated.

For instance, the current study in general was concerned with manufacturing companies and employment in the manufacturing sector, but the interactions between transportation and other sectors of the economy should also be estimated. Another suggestion is that in this research, new firms were disaggregated on less and more than

50 employees; in further research, it would be useful to disaggregate the companies in more categories.

Another suggestion for the future research is to use time series cross-section data for the analysis. Investment in transportation does not give the immediate effect on the economic growth, so with using several lags in the model, it is more likely to get better estimates and more significant results.

LITERATURE CITED

Bartik, Timothy J., Business location decisions in the United States: Estimates of the effects of unionization, taxes and other characteristics of states, *Journal of Business and Economic Statistics*, 5, Jan. 1985, 14-22.

Besser, T., Scenic roads, death traps and the rural transportation system, Iowa State University Extension to Communities, 1997.
<http://www.extension.iastate.edu/communities/news/ComCon61.html>. April, 2004.

Berwick, Mark, Bitzan, John, Chi, Junwook, and Lofgren, Mark, North Dakota strategic freight analysis: The role of intermodal container transportation in North Dakota, UGPTI report No. DP-150, North Dakota State University, November, 2002.

Brown, Dennis M., and Flake, Oliver L., Rural transportation: An annotated bibliography, Economic Research Service U.S. Department of Agriculture, March, 1999.

Button, Kenneth J., and Leitham, Scott, Transport and industrial and commercial location, *The Annals of Regional Science*, 29, 1995, 189-206.

Carlino Gerald A., and Mills, Edwin S., The determinants of county growth, *Journal of Regional Science*, 27, 1987, 39-54.

Carlton, Dennis, W., Why do new firms locate where they do: An econometric model, *Interregional Movements and Regional Growth*, COUPLE Paper Vol. 2 Washington DC.: The Urban Institute, 1979, 13-50.

Craine, Laurence M., Distributed lag effect of transportation expenditures on rural income and employment. *The Review of Regional Studies*, 23/2, 1993, 163-181.

Forkenbrock, David J., Norman, S., and Foster, J., Highways and business location decisions, *Economic Development Quarterly: The Journal of American Economic Revitalization*, 10, Aug. 1996, 239-248.

Fox, William, and Murray, Matthew, Local public policies and interregional business development, University of Tennessee, *Southern Economic Journal*, 57, Oct. 1990, 413-427.

Goode, Frank M., and Hastings, Steven E., The effect of transportation service on the location of manufacturing plants in nonmetropolitan and small metropolitan communities, *Profitability and Mobility in Rural America*, ed. Gillis, William R. The Pennsylvania State University Press, 1989.

Griffiths, W., Carter Hill, R.C., and Judge, G.C., *Learning and practicing econometrics*, New York: John Wiley and Sons, 1993.

Helgeson, Delmer L., and Zink, Maurice J., A case study of rural industrialization in Jamestown, ND, *Agricultural Economics Report No. 95*. North Dakota State University, 1973.

Hite, J., The Thunen model as a paradigm for rural development, Strom Thurmond Institute of Government and Public Affairs at Clemson University, 1998. www.strom.clemson.edu/opinion/hite/thunen.html. June, 2003.

Hough, Jill, and Griffin, Gene, North Dakota's transportation needs to further economic development: A user's perspective, UGPTI Report No. SP-127, North Dakota State University, 1994.

Hough, Jill, and Dooley, Frank, Transportation and rural economic development: A comparative analysis of important location factors, MPS Report No. 91-4, North Dakota State University, 1991.

Kale, Steven R., and Lonsdale, Richard, Factors encouraging and discouraging plant location in nonmetropolitan areas, *Nonmetropolitan Industrialization*, 1979, 47-78.

Kinnard, William N., and Malinowski, Zenon, S., Highways as a factor in small manufacturing plant location decisions, The University of Connecticut, 1961.

Kilkenny, Maureen, Economics of location, A selective survey, Department of Economics, Iowa State University, February 1998.

Kriesel, Warren, The estimation of benefits, costs and probabilities of manufacturing plant location in rural Virginia, Master Thesis, Virginia Polytechnic Institute and State University: Blacksburg, Virginia, January 1984.

Laulajainen, Risto, and Stafford, Howard A., Corporate geography, *Business Location Principles and Cases*, Dordrecht, The Netherlands, 1995.

Leistriz, Larry F., and Ekstrom, Brenda, L., Characteristics of new or expanding, export oriented firms in the Upper Great Plains, *Agricultural Economics Report No. 256*, North Dakota State University, 1990.

Leitham, S., McQuaid, R. W., and Nelson, J. D., The influence of transportation on industrial location choice: A stated preference experiment, *Transportation Research Part A*, 34, 2000, 515-535.

McFadden, D., Conditional logit analysis of qualitative choice behavior, Zarembka, P. (ed.), *Frontiers in econometric*, Academic Press: New York, 1974, 105-142.

Miller, E. Willard, *Manufacturing. A study of industrial location*, Pennsylvania State University, 1977.

Miller, E.W., *A geography of industrial location*, Brown, 1970, 28-38.

Murray, Matthew N., and Dowell, Paula, *The location decision of automotive suppliers in Tennessee and the Southeast*, Center for Business and Economic Research, University of Tennessee, Knoxville, 1999.

National association of counties, 2003.

http://www.naco.org/Template.cfm?Section=Find_a_County&Template=/cfiles/counties/usamap.cfm. April, 2004.

Popkin, Joel, *Securing America's future: The case for a strong manufacturing base*, Joel Popkin and Company, Washington DC, June 2003.

Rainey, Daniel V., and McNamara, Kevin T., *Determinants of manufacturing location: An analysis of locations in Indiana between 1986 and 1989*, *Indiana Business Review*, Bloomington, 1992.

Ritter, J. A., *The industrial location decision: A practitioner prospective*, *Economic Development Quarterly: The Journal of American Economic Revitalization*, 4, No. 2, 1990, 154-156.

Salvesen, David, and Renski, Henry, *The importance of quality of life in the location decisions of new economy firms*, *Reviews of Economic Development Literature and Practice*: No. 15, University of North Carolina, 2002.

Sloggett, G., and Woods, Mike D., Critical factors in attracting new businesses and industry in Oklahoma, Cooperative Extension Service, Oklahoma State University, 1989.

Stephanedes, Yorgos, and Eagle, David M., Time series analysis of interactions between transportation and manufacturing and retail employment, Transportation Research Record, 1074, 1986, 16-24.

Stommes, Eileen S., and Brown, Dennis M. Transportation in rural America. Issues for the 21st century, Rural America, 16, 2002, 2-12.

Thompson, James H., Methods of plant site selection available to small manufacturing firms, West Virginia University Bulletin, Series 62, No. 3-3, September 1961.

Walker, R., and Greenstreet, J., The effect of government incentives and assistance on location and job growth in manufacturing, Regional Studies, 25/1, 1991, 13-30.

Wheeler, James O., and Pannell, Clifton W., The impact of an expanding transportation network on potential economic development locations in the North Georgia area, University of Georgia, June 1977.

Wilkerson, Chad, Trends in rural manufacturing, The Main Street economist: Commentary on the rural economy, Center for the Study of Rural America, Federal Reserve Bank of Kansas City, December, 2001.

APPENDIX A

Variance-covariance matrix.

	RAIL	AIR	DIS	LMI	LMO	LM	INSUR
RAIL	1.0000	-0.0357	-0.1298	0.0202	0.0158	0.2252	0.1003
		-0.4631	-0.0074	-0.6789	-0.7461	(<.0001)	-0.0390
AIR	-0.0357	1.0000	0.1431	0.2186	0.3660	0.0700	0.2289
	-0.4631		-0.0031	(<.0001)	(<.0001)	-0.1499	(<.0001)
DIS	-0.1298	0.1431	1.0000	0.0877	0.1220	0.1154	-0.1551
	-0.0074	-0.0031		-0.0712	-0.0119	-0.0175	-0.0014
LMI	0.0202	0.2186	0.0877	1.0000	-0.1564	0.1389	0.1458
	-0.6789	(<.0001)	-0.0712		-0.0012	-0.0042	-0.0026
LMO	0.0158	0.3666	0.1220	-0.1564	1.0000	0.3967	0.4228
	-0.7461	(<.0001)	-0.0119	-0.0012		(<.0001)	(<.0001)
LM	0.2252	0.0700	0.1154	0.1389	0.3967	1.0000	0.3472
	(<.0001)	-0.1499	-0.0175	-0.0042	(<.0001)		(<.0001)
INSUR	0.1003	0.2289	-0.1551	0.1458	0.4228	0.3472	1.0000
	-0.0390	(<.0001)	-0.0014	-0.0026	(<.0001)	(<.0001)	
PCPI	0.0163	0.1805	-0.1380	0.1160	0.1397	0.0977	0.2441
	-0.7375	-0.0002	-0.0042	-0.0169	-0.0040	-0.0444	(<.0001)
CAR	0.2616	0.2383	-0.3314	0.0992	0.1978	0.1949	0.3240
	(<.0001)	(<.0001)	(<.0001)	-0.0412	(<.0001)	(<.0001)	(<.0001)
AVTIME	0.0062	-0.0820	-0.2193	0.0612	-0.0565	-0.1663	0.1823
	-0.8988	-0.0917	(<.0001)	-0.2088	-0.2452	-0.0006	-0.0002
RENT	-0.1164	0.2876	-0.1440	0.1587	0.1694	-0.1789	0.3050
	-0.0165	(<.0001)	-0.0030	-0.0010	-0.0005	-0.0002	(<.0001)
COLL	0.0489	0.2663	-0.1231	0.1335	0.1160	0.0438	0.1446
	-0.3151	(<.0001)	-0.0111	-0.0059	-0.0169	-0.3689	-0.0028
AREA	-0.2142	0.3290	0.4363	0.3523	0.4197	0.1812	0.0553
	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	-0.0002	-0.2557
POPDEN	0.0800	0.0625	-0.2196	0.1270	0.1919	0.1855	0.8168
	-0.0999	-0.1993	(<.0001)	-0.0089	(<.0001)	-0.0001	(<.0001)

	PCPI	CAR	AVTIME	RENT	COLL	AREA	POPDEN
RAIL	0.0163	0.2616	0.0062	-0.1164	0.0489	-0.2142	0.0800
	-0.7375	(<.0001)	-0.8988	-0.0165	-0.3151	(<.0001)	-0.0999
AIR	0.1805	0.2383	-0.0820	0.2876	0.2663	0.3290	0.0625
	0.0000	(<.0001)	-0.0917	(<.0001)	(<.0001)	(<.0001)	-0.1993
DIS	-0.1386	-0.3314	-0.2193	-0.1440	-0.1231	0.4363	-0.2196
	-0.0042	(<.0001)	(<.0001)	-0.0030	-0.0111	(<.0001)	(<.0001)
LMI	0.1160	0.0992	0.0612	0.1587	0.1335	0.3523	0.1270
	-0.0169	-0.0412	-0.2088	-0.0010	-0.0059	(<.0001)	-0.0089
LMO	0.1397	0.1978	-0.0565	0.1694	0.1160	0.4197	0.1919
	-0.0040	(<.0001)	-0.2452	-0.0005	-0.0169	(<.0001)	(<.0001)
LM	0.0977	0.1949	-0.1663	-0.1789	0.0438	0.1812	0.1855
	-0.0444	(<.0001)	-0.0006	-0.0002	-0.3689	-0.0002	-0.0001
INSUR	0.2441	0.3240	0.1823	0.3050	0.1446	0.0553	0.8168
	(<.0001)	(<.0001)	-0.0002	(<.0001)	-0.0028	-0.2557	(<.0001)
PCPI	1.0000	0.1754	-0.1355	0.5462	0.0725	0.0061	0.2923
		-0.0003	-0.0052	(<.0001)	-0.1360	-0.8997	(<.0001)
CAR	0.1754	1.0000	0.2164	0.0985	0.1121	-0.1098	0.2294
	-0.0003		(<.0001)	-0.0426	-0.0210	-0.0237	(<.0001)
AVTIME	-0.1355	0.2164	1.0000	0.1874	-0.1332	-0.0038	0.0516
	-0.0052	(<.0001)		-0.0001	-0.0060	-0.9365	-0.2891
RENT	0.5462	0.0985	(0.1874	1.0000	0.2143	0.1411	0.2481
	(<.0001)	-0.0426	-0.0001		(<.0001)	-0.0036	(<.0001)
COLL	0.0725	0.1121	-0.1333	0.2143	1.0000	0.0214	0.1570
	-0.1360	-0.0210	-0.0060	(<.0001)		-0.6606	-0.0012
AREA	0.0061	-0.1098	-0.0039	0.1411	0.0214	1.0000	-0.1644
	-0.8997	-0.0237	(0.9365*)	-0.0036	-0.6606		-0.0007
POPDEN	0.2923	(0.22936	0.0516	0.2481	0.1570	-0.1643	1.0000
	(<.0001)	(<.0001)	-0.2891	(<.0001)	-0.0012	-0.0007	

APPENDIX B

```
PROC IMPORT OUT= WORK1.masha1
    DATAFILE= "C:\Documents and Settings\burdina\My Documents\My
Documents Copy\My documents\Thesis\Sas.xls"
    DBMS=EXCEL2000 REPLACE;
    GETNAMES=YES;
RUN;

/*Collinearity check 1 */
proc corr data=work1.masha1;
var    RAIL AIR DIS LMI LMO LM INSUR PCPI CAR AVTIME RENT COLL
AREA POPDEN;
proc print;
run;

/*Collinearity check 2 */
proc reg data=WORK1.masha1;
model EMPLT= RAIL AIR DIS LMI LMO LM INSUR PCPI CAR AVTIME RENT
COLL AREA POPDEN / tol vif collin ;
run;
```

APPENDIX C

```
PROC IMPORT OUT= WORK1.masha1
    DATAFILE= "C:\Documents and Settings\burdina\My Documents\My
Documents Copy\My documents\Thesis\Sas.xls"
    DBMS=EXCEL2000 REPLACE;
    GETNAMES=YES;
RUN;

/*Location model for manufacturing companies with less than 50 employees*/
proc logistic DESCENDING data=WORK1.masha1 ;
    model SM= RAIL AIR DIS LMI LMO LM INSUR PCPI CAR AVTIME RENT
COLL AREA POPDEN;
    run;

/*Location model for manufacturing companies with more than 50 employees*/
proc logistic DESCENDING data=WORK1.masha1 ;
    model LG= RAIL AIR DIS LMI LMO LM INSUR PCPI CAR AVTIME RENT
COLL AREA POPDEN;
    run;
```

APPENDIX D**/*ENDOGENEITY EMPLT*/****proc syslin** data=work.masha1 covout 2sls;

endogeneous PCPI

instruments WAGE RENT EDU

first : model EMPLT = RAIL AIR DIS LMI LMO LM INSUR PCPI CAR AVTIME
RENT COLL AREA POPDEN;**run;****/*ENDOGENEITY EMPLM*/****proc syslin** data=work.masha1 covout 2sls;

endogeneous PCPI

instruments WAGE RENT EDU

first : model EMPLM = RAIL AIR DIS LMI LMO LM INSUR PCPI CAR AVTIME
RENT COLL AREA POPDEN;**run;**

APPENDIX E

```

PROC IMPORT OUT= WORK1.masha1
      DATAFILE= "C:\Documents and Settings\burdina\My Documents\My
Documents Copy\My documents\Thesis\Sas.xls"
      DBMS=EXCEL2000 REPLACE;
      GETNAMES=YES;
RUN;

/*Heteroskedusticity Total Employment*/
Intercept=1;
Proc reg data=work1.masha1;
model EMPLT= Intercept RAIL AIR DIS LMI LMO LM INSUR PCPI CAR AVTIME
RENT COLL AREA POPDEN;
output out = sasout residual=ehat;
data step1;
set sasout;
ehat2 = ehat**2;
proc means mean;
var ehat2;
output out= vardat mean = sig2;
proc reg data=step1;
model ehat2 = intercept  RAIL AIR DIS LMI LMO LM INSUR PCPI CAR AVTIME
RENT COLL AREA POPDEN;
output out=regout residual = e;
proc means uss css;
var ehat2 e;
output out=ssqout uss=ssehat2 sse css=sst ce;
data step3;
merge ssqout vardat;
ssr = sst- sse;
sig4 = sig2**2;
bp=ssr/(2*sig4);
pval = 1 - probchi(bp,1);
proc print;
var ssr sig4 bp pval;
run;

```

APPENDIX F

```

PROC IMPORT OUT= WORK1.masha1
  DATAFILE= "C:\Documents and Settings\burdina\My Documents\My
Documents Copy\My documents\Thesis\Sas.xls"
  DBMS=EXCEL2000 REPLACE;
  GETNAMES=YES;
RUN;
proc IML;

/*gravity model estimation*/
/*coefficient matrix, OLS*/
start white(x,y);
b=inv(x`*x)*x`*y; /*OLS coefficients*/
print b;
e=y-x*b; /*residuals*/
e2_=e*e`;
identity=i(nrow(e2_));
e2=e2_#identity;
s0=inv(x`*x)*(x`*e2*x)*inv(x`*x); /*White's S0*/
  /*Var-Covar etimated matrix robust to heteroskedasticity*/
print s0;

s0stnd=sqrt(vecdiag(s0));
print s0stnd;

s_ols=(1/(nrow(x)-ncol(x)))*e`*e*inv(x`*x); /*var-covar of standard ols*/
print s_ols;
s_olsstnd=sqrt(vecdiag(s_ols));
print s_olsstnd;

s_diff=s0stnd-s_olsstnd; /*estimate differences*/
print s_diff;

trs=b/s0stnd; /*t-ratios*/
print trs;

/*R-square*/
M0=i(nrow(y))-(1/nrow(y))*vecdiag(i(nrow(y)))*t(vecdiag(i(nrow(y))));
R2=1-e`*e/(y`*M0*y);
print R2;

/*Adjusted R-square*/
R2a=1-(nrow(x)-1)*(1-R2)/(nrow(x)-ncol(x));
print R2a;

```



```

output=shape(0,nrow(b),5); /*output matrix coeff, white std dev, t-ratios, ols std dev*/
output[,1]=b;
output[,2]=s0stnd;
output[,3]=trs;
output[,4]=s_olsstnd;
output[2,5]=R2;
output[4,5]=R2a;
print output;
create WORK1.outputwhite from output; /*creating SAS output dataset*/
append from output;

```

```

create WORK1.whitevariance1 from s0; /*creating SAS output dataset instrument
variance is in whitevariance*/
append from s0;

```

```

finish white;

```

```

use WORK1.masha1 var{ Intercept EMPLT RAIL AIR DIS LMI LMO LM INSUR
PCPI CAR AVTIME RENT COLL AREA POPDEN };

```

```

read all var{intercept RAIL AIR DIS LMI LMO LM INSUR PCPI CAR AVTIME
RENT COLL AREA POPDEN } into X1;
read all var{EMPLT} into Y1;

```

```

run white(X1,Y1);

```

```

proc export data= WORK1.outputwhite
outfile="C:\Documents and Settings\burdina\My Documents\My
Documents Copy\My documents\Thesis\results1.xls"
REPLACE;
run;

```