A LATENT VARIABLE APPROACH TO MIGRATION

By

RAFAEL ESCUTIA VILLAGOMEZ

Bachelor of Arts
University of Illinois
Chicago, Illinois
1971

Master of Arts
DePaul University
Chicago, Illinois
1973

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
May, 1985

A LATENT VARIABLE APPROACH TO MIGRATION

Thesis Approved:

Ronald J Moomaw
Thesis Adviser

John D. Rea

Orley M. Amos Jr.

Luther G. Tweeten

Norman N. Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

I dedicate this thesis to my parents, Raúl Villagomez and Consuelo Escutia. Their love and support have guided me though the years.

I thank my wife, Sarah Toledo, for all that she has given of herself to help me finish. Her help has been incalculable.

I wish to thank Dr. Clifford Young for his help. His contribution was truly indispensable. I am pleased to have had the opportunity to have Dr. John Rea, Dr. Luther Tweeten, and Dr. Orley Amos on my committee. Most of all, I wish to extend a heartfelt thank you to Dr. Ronald Moomaw, the committee chairman. His guidance has taught me much indeed.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The purposes of this study are to investigate the determinants of migration with emphasis on quality of life considerations and to compare the three estimation techniques, multiple regression analysis, latent variable modelling, and index number modelling, used in the investigation. Since the latent variable technique has not been used in migration studies before, the study also provides an opportunity to examine the usefulness of this approach for analyzing migration.

A better understanding of the determinants of migration leads to a better understanding of the effects of policy actions on migration and a more efficient use of our scarce resources. In this regard, if quality of life factors are important determinants of migration, then regional economic development policies may not be as effective as otherwise thought because government has little influence over some quality of life factors, especially climate (Porell, 1982). If environmental amenities have become more important in the location decision, then the preservation of an area's environmental attractiveness can have an important effect on an area's growth. To retain population and to attract new residents, environmental protection and improvement may have to receive more attention. If climate considerations have increased in importance in the migration decision, then jobs alone may not be enough incentive to attract migrants to an environmentally unattractive locale (Long and Hansen, 1979).

The determinants of migration are also of importance to policy makers who have to predict future demand for public services. An explanation of migration flows would improve the accuracy of their predictions (Graves, 1979a).

Although the usefulness of knowledge about the determinants of migration has stimulated many migration studies (for two surveys, see Greenwood, 1975; Ritchey, 1976), only recently have scholars turned their attention to quality of life factors (Cebula and Vedder, 1973; Liu, 1975a; Graves, 1976, 1979a, 1979b, 1980; Kau and Sirmans, 1976; Porell, 1982). These studies demonstrate the importance of quality of life factors in the migration decision by comparing economic with environmental variables. Although the studies differ with regard to the types of models used, the results are similar in that quality of life variables are found to be important determinants of migration.

Both the Kau and Sirmans and the Liu studies looked at migration between states. Since states are heterogenous areas, the Standard Metropolitan Statistical Area (SMSA) or county level used in the other studies cited above is a better choice for analysis of migration. The Cebula and Vedder study (1973) and two of the Graves studies (1976, 1979) used net migration rates as the dependent variable. Gross migration studies, however, tend to have more variables with significant coefficients than do net migration studies because common variables probably cancel out in net migration studies (Greenwood, 1975). This cancellation results because net migration is in-migration minus out-migration. Greenwood (1975) also points out that for some variables, such as income, the use of net migration amplifies their influence.

Graves (1980) and Porell (1982) used gross migration data. The Graves study used regression, but Porell estimated his model using index numbers. Index numbers are one way to group many variables into a smaller number of regressors to simplify the model. Another way to do this is to use a latent variable technique related to factor analysis. This technique may yield more information than the index number alternative, and allows for more flexible modelling.

Plan of the Study

Chapter II will develop the theoretical model used in the study. In Chapter III, the data, methodology, and limitations of the empirical tests will be discussed along with the regression results. Chapter IV explains the latent variable estimation technique and Chapter V presents the empirical results of latent variable estimation. Chapter VI reports and discusses the results of index number regression. Finally, Chapter VII presents a summary, policy implications, and conclusions from the study.

Quality of Life Concept

Central to the work on determinants of migration that focuses on quality of life is the very concept of quality of life itself. Amos (1980) reviews the quality of life literature and finds four points of agreement on the meaning of quality of life. The first point is that satisfaction, be it of wants or needs or of motivations, is a common theme in the literature. The second point is that quality of life is defined in terms of the individual. The third point is that quality of life is subjective. The fourth point is that there are objective dimensions of quality of life which combine with the subjective dimension. For example, playing golf can relax a person. The physical act of playing golf combines with the person's perceptions to produce relaxation. Relaxation is what is important to him. It is his motivation, and the motivation is satisfied by playing golf (Amos, 1980).

Chapter II discusses the quality of life concept. It develops a theoretical framework which ties quality of life and migration. This model provides the base for the empirical work of the study.

CHAPTER II

THEORETICAL FRAMEWORK

Introduction

In this chapter the utility function is presented and discussed. The relationship between quality of life and utility is analyzed. The chapter then ties together utility, quality of life, and migration. A summary and conclusions end the chapter.

Utility Function

Adapting the analysis from Lesourne (1977) and from Kau and Sirmans (1976), the individual is assumed to maximize utility subject to constraints. In general, the utility maximization framework is summarized by the following equations:

$$\text{Maximize } U = f\ (TG, NTG, PS, L, W) \tag{1}$$

Subject to:

$$P_{TG}TG + P_{NTG}NTG = I \tag{2}$$

$$PMT + CT + L + W = 24 \tag{3}$$

where TG = Traded Goods,

NTG = Non-traded Goods,

PS = Personal Situation,

L = Leisure Time,

W = Work Time,

P_{TG} = Prices of Traded Goods,
P_{NTG} = Prices of Non-traded Goods,
PMT = Personal Maintenance Time,
CT = Consumption Time, and
I = Income.

Equation (1) shows that utility depends, in part, on the consumption of goods and services. Some of these are available only locally. These are included in the category of non-traded goods and services. Examples of this category are such items as a Museum of Modern Art, housing, and climate characteristics. Utility also depends on one's personal situation, that is, on family life, social life, proximity to loved ones, or changes in the life cycle. An individual's leisure activities and work experiences also affect his utility.

In equation (2) we see that income puts, along with prices, a limit on the costs the individual can incur. These are the costs of traded goods plus the costs of non-traded goods.

Equation (3) shows that the individual also has a time constraint. Work, leisure, and consumption are constrained by the length of the day and by the demands of other activities on an individual's time. A person needs time as well as income to consume goods and services as well as time for personal maintenance.

Quality of Life and Utility

Quality of life can be considered to be utility. All four points of agreement on quality of life discussed by Amos (1980) can also be made with respect to utility. Utility is generated by satisfaction of needs and wants. Utility is discussed strictly at the level of the individual and not at the group level. Utility is subjective since everyone's tastes and preferences differ. Different tastes lead to different utility functions and different reactions to identical stimuli. Many of these stimuli are objective and combine, as in point four, for quality of life with the subjective dimension to produce utility.

The role of needs and wants is recognized in utility theory, although utility is usually presented as a direct function of the consumption of goods (Lesourne, 1977). The consumption of goods generates utility. For example, watching a baseball game generates utility. But the quality of life approach would emphasize that watching baseball relaxes a person and that it is the relaxation that is important. So relaxation generates utility and not watching baseball per se. The satisfaction of motivations, such as relaxation, generates utility, in other words a certain level of quality of life. Quality of life can be considered to be utility but viewed in a wider context.

Measuring Quality of Life

The work that attempts to measure quality of life can be seen as an attempt to measure utility. This implies a cardinal utility concept as opposed to the common assumption of ordinal utility. Measuring quality of life for an individual is a difficult task because of the subjectivity involved in quality of life. There are no quality of life meters that can be used in the measuring process. Measuring an intangible concept such as quality of life is more difficult than measuring the quantifiable inputs that, through want satisfaction, generate quality of life. The difficulty is even greater when many of the inputs are themselves intangible, e.g., love. Thus, most of the approaches to measuring quality of life have measured the tangible inputs rather than the intangible output.

Since quality of life is multi-dimensional, with many motivations involved, quite a few approaches to measuring it have appeared in the literature. If measuring quality of life for an individual is difficult, measuring quality of life for a group is even more difficult due to the aggregation problems. Quality of life is, as shown by the second of Amos's (1980) four points, defined at the level of the individual. Nonetheless, attempts have been made to compare quality of life at the group level. Zapf (1975) provides a good summary of some representative attempts and the problems involved in trying to measure quality of life.

The measurement process has taken three routes, depending on the unit of measurement used. Measuring can be done in terms of money, physical units, or "psychical units" (that is, through measurement of attitude). Problems arise in defining quality of life, in selecting and assigning weights to the various dimensions, or components, of quality of life, in choosing which measures of quality of life to use, in making these measures workable, and finally, in obtaining new or available statistical data (Zapf, 1975).

Some examples are: Japan's <u>Net National Welfare</u>, which measures marketed and non-marketed production available for consumption; John Wilson's <u>Social Indicators Battery</u>, which uses eight main components to rank the 50 states according to quality of life; Nestor Terleckyj's <u>National Goals Accounting</u>, which uses an input-output matrix to calculate possibilities for improving the quality of life; Andrew and Withey's <u>Perceived Life Quality Scale</u>, which attempts to measure and predict general satisfaction with life, using a survey and a minimum of questions; and finally, Abram's Quality-of-Life Survey, which uses a combination of measures of satisfaction, significance, and aspirations to look at an individual's quality of life (Zapf, 1975).

Liu (1975) developed a production function model for quality of life in which he divided the inputs into the production process into physical and psychological inputs. He first developed an iso-quality curve, analogous to an isoquant, and then an iso-capability curve, analogous to an iso-cost line. Then he viewed the individual as optimizing quality of life subject to the iso-capability constraint.

Liu's study quantified quality of life by measuring the inputs, especially the physical inputs, for which data are more readily available. He combined 123 factors into five component quality of life indexes for 243 SMSAs in the U.S. These five components were then used to describe variations in quality of life among SMSAs in 1970.

The Liu quality of life data consist of index numbers representing the various quality of life dimensions. Indicators of quality of life dimension are combined into one index number. The construction of the index numbers involves assigning weights to each of the component indicators. The assignment of weights can be a serious problem (Aaker and Bagozzi, 1979). Should the indicators receive equal weights or not, and, if not, what weights should be assigned and why? The index numbers are also probably not exact measures of the aspects of quality of life that they represent. That is, they probably have some measurement error. These index numbers are then used as the regressors in empirical analyses of migration and quality of life, but multiple regression assumes no measurement error in the independent variables.

Using the same indicators as the index number approach, the latent variable approach estimates the relationship between various aspects of the quality of life--the latent variables--and migration. Latent variables modelling takes into account measurement error and does not require the construction of index numbers. This approach will be described in detail in Chapter IV.

Studies on Quality of Life and Migration

Due to the difficulty in measuring quality of life, previous migration studies that focus on quality of life just added factors to investigate as determinants of migration, for example, climate. That is, they have looked at differences in inputs rather than differences in quality of life itself. Though differences in quality of life are not directly observable, their effects on migration are. Since inputs affect quality of life which in turn affects migration, researchers have studied the relationship between inputs and migration in order to study indirectly the relationship between quality of life and migration.

Migration studies which look at quality of life have done so at the group level rather than at the level of the individual. This increases the problem in measuring quality of life. So it is even more difficult to do anything other than compare differences in inputs rather than differences in quality of life itself. Quality of life is, as Amos (1980) shows in his point number one, concerned with non-physical dimensions and cannot be measured in physically observable phenomena. Thus, for example, Liu's (1975a) and Porell's (1982) studies are inputs to quality of life rather than quality of life as determinants of migration.

Overall, then, studies dealing with quality of life and migration have increased the number of determinants of migration under investigation adding "quality of life" variables. These studies have added climate variables, city amenity and disamenity variables, health factors, and social variables (Cebula and Vedder, 1973; Liu, 1975a; Graves, 1976, 1979a, 1980; Kau and Sirmans, 1976; Hall and Licari, 1977; Porell, 1982).

Cebula and Vedder (1973) studied net migration, 1960 to 1970, for SMSAs. They regressed net migration on income, unemployment, income growth, number of physicians, crime, racial composition, temperature, and air pollution. Both economic and quality of life variables were found to be important since unemployment, income growth, temperature, and number of physicians were significant and had the correct signs.

Graves (1976) attempted to reproduce Cebula and Vedder's results using 1960 to 1968 net migration for 39 SMSAs. He then made several alterations in the model: median family income replaced per capita income, heating degree days replaced the average number of days the temperature is below freezing, and crime and air pollution were dropped from the model. Graves argued that only global environmental factors, climate but not air pollution, affect the migration decision. The empirical results upheld the importance of both the economic variables and climate as determinants of net migration.

Miller (1973a) studied 1955 to 1960 out-migration at the state level. He added temperature as a determinant of migration and found that warmer states had lower out-migration. Liu (1975a) looked at 1965 to 1970 net state migration data and added not only temperature but many other quality of life variables to the determinants of migration. He used over 100 variables to construct indexes representing economic status, individual status, economic equality, living conditions, agricultural production, technological development, educational development, health and welfare, and government. Liu concluded that quality of life factors did significantly affect migration and were more important than economic factors.

Kau and Sirmans (1976) used Liu's data to study 1965 to 1970 gross migration flows from the nine census regions to each of the states. They employed a recursive model which incorporated the migrant stock from previous migrations. Both lifetime and current migration were examined as well as total, white, and black migration. There were specific differences both between lifetime and current migrants and between white and black migrants. But, in general, Kau and Sirmans found that migration flows were affected by both economic and quality of life factors.

Graves (1979) examined net 1960 to 1970 SMSA migration adding climate variables to median family income and the unemployment rate as determinants of migration. He disaggregated by race and age. Differences were found both between the races and between age groups with negative effects in one group being offset by positive effects in another group. In general, he found that both economic and quality of life variables were important. Indeed, Graves considered the omission of climate variables a source of downward bias on the income and unemployment coefficients.

Graves (1980) investigated gross rather than net 1965 to 1970 migration for 49 SMSAs. As before, he disaggregated by age and race. Income and unemployment, as sole regressors, were not significant.

This was consistent with his previous (1979) agrument. However, when climate variables were included, economic considerations continued, by and large, to be insignificant. Climate, on the other hand, was important across age groups for both in- and out-migration.

Porell (1982) examined the question of the relative importance of economic and quality of life determinants of migration. He investigated gross 1965 to 1970 migration flows between 25 SMSAs and their relationship to population variables, economic variables, and quality of life indexes representing climate, outdoor amenities, indoor amenities, crime, pollution, and health. He found both economic and quality of life factors to be important determinants of in-migration, but not of out-migration.

Due to these migration studies, the importance of non-economic, quality of life determinants has been recognized and incorporated into the migration model. Thus, the migration model has been made more complete and realistic.

Migration

A location decision facing the individual is where to live. As an individual maximizes utility, subject to constraints, he includes in the decision making process a comparison of the expected utility levels associated with alternative locations. The constraints and the utility function combine to determine the expected utility level at each location. The problem is which location offers the highest expected utility level.

An individual living and working in a certain SMSA experiences a level of utility and forecasts an expected level of utility for the future at that location. Other SMSAs offer alternatives with respect to jobs, climate, cultural amenities, public services, cost of living, distance from friends and relatives, etc. The individual predicts an expected utility level for each location he considers and chooses the one with the highest expected utility level as his next home. This may entail a move to another SMSA. If he moves, once at the new SMSA, and having experienced a certain level of utility, a new comparison between locations is made. Another move may follow, as is shown by the many repeat migrants (Miller, 1973).

Both investment and consumption considerations enter into the decision-making process. Where investment considerations dominate, Sjaastad's (1962) approach is relevant. At some stages of the life cycle, however, the investment aspect is not as important. For example, consumption factors may be relatively more important for the elderly than for the 30-34 year age group. Thus, as Kau and Sirmans (1976) point out, in order to evaluate investment considerations, other factors, such as climate and age, must be held constant.

Locations are ranked by individuals according to expected utility levels. When the current location drops from the top of the list, a move ensues. Moving costs are taken into account in comparing expected utility levels. Expected utility levels depend on traded goods, non-traded goods, and on what is done during the various parts of the day. Changes in inputs affect utility levels and so affect migration since changes in the expected utility levels may drop the current location from the top of the list. In an unobservable relationship, migration depends on expected utility levels. In the observable relationship, migration depends on the levels of inputs or what the literature refers to as the determinants of migration. These determinants can be divided into various categories. This study divides them into an economic group (ECON), a climate group (CLIM), a demographic group (DEMO), a city amenities group (AMEN), and a city disamenities group (DISAMEN). The relationship between gross migration from a given location to other locations and its determinants is shown in equation (4):

$$\text{MIGRATION} = f(\text{ECON, CLIM, DEMO, AMEN, DISAMEN}) \quad (4)$$

Summary

In this chapter it was argued that quality of life and utility are equivalent terms. Measurement approaches to measuring quality of life were presented and problems of measurement were discussed. The relationship between quality of life and migration was spelled out and a general migration framework was given. In the next chapter, regression analysis is used to begin the empirical analysis of migration and quality of life.

CHAPTER III

REGRESSION ANALYSIS

Introduction

In this chapter, the sample used in the study is presented, the variables are discussed, and summary statistics for the sample are given. Multiple regression results are then presented and compared with those of other studies. A summary and conclusions end the chapter.

Sample

The sample consists of 77 SMSAs ranging in population from 84,000 to 11,366,000. SMSAs were chosen since, as Fields (1979) points out, SMSAs reflect the labor market better than other data sources. However, the gross migration data were available for Standard Economic Areas (SEAs) not for SMSAs. This causes a problem since the subject of study is inter-location migration. When the boundaries of an SMSA fall within more than one SEA, intra-location migration appears as inter-location migration. Thus, only a portion of the 243 SMSAs in the U.S. in 1965 could be used. The sample is larger than Graves' (1980) sample of 49 SMSAs in his study of gross SMSA migration. The SMSAs in the sample are presented in Table I along with their 1965 population.

Variables

The five categories represented in equation (4) of the previous chapter need to be given empirical content. Variables are selected to provide indicators of each of the categories. The utility maximization framework and the discussion of the migration decision facing the individual guide the selection of variables. The variables, their definitions, and the data sources are presented in Table II.

Economic Category

Looking at the economic category, the constraints in the utility maximization framework suggest that real income is important. Thus, both the cost of living (COL) and nominal median family income (MFY) influence the variable used for the real level of income, MFY/COL. Median family income is used since migration involves households. In effect, the term "individual" refers to the entire family. Fields (1976) and Cebula (1979) argue for the use of deflated data for the level of income.

When the individual projects into the future to form estimates of expected utility levels in order to compare cities, he needs a forecast for income, that is he needs to forecast changes in income. An indicator of future income change could be past change in median family income (MFYG).

Moreover, the individual will be concerned with the probability of getting or keeping a job, which can be indicated by employment growth (EMPG). Greenwood (1981) agrues that the growth of employment reflects a rise in job opportunities in an area and the growth in labor demand. A frequent measure of employment opportunities is the unemployment rate,

TABLE I

SAMPLE WITH 1965 POPULATION IN THOUSANDS

SMSA	Population	SMSA	Population
Albuquerque, NM	288	Los Angeles, CA	7,872
Altoona, PA	137	Lubbock, TX	185
Ann Arbor, MI	187	Macon, GA	201
Asheville, NC	143	Madison, WI	260
Atlanta, GA	1,216	Miami, FL	1,061
Atlantic City, NJ	179	Minneapolis, MN	1,612
Baton Rouge, LA	255	Monroe, LA	112
Bay City, MI	109	Muncie, IN	117
Billings, MT	84	New Haven, CT	704
Birmingham, AL	644	New York, NY	11,366
Boston, MA	3,205	Oxnard, CA	318
Bridgeport, CT	746	Pensacola, FL	224
Buffalo, NY	1,320	Phoenix, AZ	818
Canton, OH	356	Pittsburg, PA	2,372
Cedar Rapids, IA	148	Portland, ME	197
Charleston, WVA	245	Providence, RI	739
Charlotte, NC	360	Pueblo, NM	119
Chicago, IL	6,689	Raleigh, NC	195
Decatur, IL	122	Reading, PA	283
Des Moines, IA	271	San Diego, CA	1,136
Detroit, MI	3,987	San Francisco, CA	2,918
El Paso, TX	344	San Jose, CA	885
Erie, PA	255	Santa Barbara, CA	243
Eugene, OR	194	Savanah, GA	192
Fall River, MN	411	South Bend, IN	270
Fresno, CA	403	Spokane, WA	167
Gary, IN	596	Springfield, IL	153
Green Bay, WI	137	Springfield, OH	147
Hartford, CT	765	Springfield, MO	140
Jacksonville, FL	250	Syracuse, NY	606
Jersey City, NJ	619	Topeka, KA	149
Kalamazoo, MI	181	Trenton, NJ	296
Kenosha, WI	114	Tucson, AZ	307
Lake Charles, LA	135	Tuscaloosa, AL	118
Lancaster, PA	289	Waco, TX	156
Las Vegas, NV	232	Waterloo, IA	124
Lexington, KY	159	Worcester, MA	608
Lincoln, NE	161	West Palm Beach, FL	281
Little Rock, AR	279		

TABLE II

VARIABLES

Variable	Definition	Source
INMIG	Gross in-migration into an SMSA from 1965 to 1970 divided by 1965 SMSA population	Migration - 1970 Census, "Migration Between State Economic Areas" Population - Bureau of the Census, Current Population Reports
OUTMIG	Gross out-migration from an SMSA from 1965 to 1970 divided by 1965 SMSA population	Migration - 1970 Census, "Migration Between State Economic Areas" Population - Bureau of the Census, Current Population Reports
MFYG	Percentage change in median family income, 1960 to 1970	City and County Data Book, 1960, 1970
MFY/COL	Median family income, 1960 to 1970 average, divided by a 1970 cost of living index	Income - City and County Data Book, 1960, 1970 Cost of Living Index - Liu (1975)
EMPG	Percentage change in employment	City and County Data Book, 1962, 1972
COLD	Heating degree days, 1941-70 normals	U.S. Climatological Data, Annual Summary, 1980
TVAR	The difference between the average July maximum temperature and the average minimum January temperature, 1941-70 normals	U.S. Climatological Data, Annual Summary, 1980
RELH	An average of relative humidity readings at different times of the day for January and July	U.S. Climatological Data, Annual Summary, 1980
WIND	Average wind speed in miles per hour, 1941-70 normals	U.S. Climatological Data, Annual Summary, 1980

TABLE II (Continued)

Variable	Definition	Source
EDUC	Percentage of the population over 25 years of age with one or more years of college	1970 Census, State Volumes
AGE	Percentage of the population in the 20-34 year old age group, 1960 to 1970 average	1960 Census, 1970 Census, State Economic Areas
CUINS	Number of cultural institutions such as museums	Liu (1975)
DDM	Number of dance, drama, and music events, 1970	Liu (1975)
SPORTS	Number of major sports events, 1970	Liu (1975)
CRIME	Total crime rate per 100,000 population, 1970	Liu (1975)
AIRPOL	Mean level of suspended particulates, 1966	National Air Pollution Control Administration, 1968

but, according to Greenwood (1975), it is usually insignificant. This may be because, as Fields (1979) argues, the unemployment rate pertains to the complete set of workers and jobs, including the employed, whereas migrants are more concerned with turnover in the labor market and give more importance to the creation of new jobs or the growth in hiring for new jobs. Thus, growth in employment would be a better indicator of employment opportunities than the unemployment rate.

In-migration is expected to be positively related, and out-migration negatively related, other things such as climate being equal, to the level of income, the growth of income, and the growth in employment.

Climate Category

There are several relevant aspects of the climate category. The first is temperature. An individual may seek a warm climate for itself, or because of the outdoor activities it allows, or for his health. Some people may prefer a cold climate and its lifestyle and associated outdoor activities. A related aspect is the variance in temperature. Some people may prefer the full flower of the four seasons and their swings in temperature. Others may prefer to avoid temperature swings and seek a constantly pleasant temperature range. As incomes rise, demand for the preferred climate may rise. To satisfy this increased demand, a person may migrate, as Graves and Linneman (1979) indicate, since climate is location fixed and non-tradeable.

These two aspects may be indicated by heating degree days (COLD) and temperature variance (TVAR). Normals for 1941-70 are used to insure that unusual years do not receive undue importance. As Graves (1976) points out, heating degree days capture the usual notion of cold better than average temperature. Heating degree days are the number by which the average temperature for the day falls short of 65 degrees Fahrenheit. Two cities may have the same average temperature, but if city A's winter high temperature is 25 degrees lower than city B's, city A will be considered by most people to be colder than city B. City A will also have more heating degree days than city B.

Two other aspects of climate are relevant to an individual's assessment of a city's climate. Both relative humidity (RELH) and wind speed (WIND) affect the body's perception of a given temperature. The higher the wind speed, the colder it feels. In addition, relative humidity affects a person's skin and breathing. There is the expectation that relative humidity, wind speed, temperature variance, and heating degree days will be negatively related to in-migration and positively related to out-migration.

Demographic Category

The relevant aspects of the demographic category are education and the life cycle. According to Greenwood (1975), employment information increases with education. More and better job information is available to the better educated. In addition, job opportunities are expected to increase with education. Schwartz (1973) finds that the job market is more national in scope for the better educated. Moreover, Saben (1964) finds that most professional and technical workers are likely to already have a job in their destination when they move. Education may also, according to Greenwood (1975), lessen the hold of custom and family links on the individual. This, coupled with an increased awareness of other places, weakens the bonds of attachment to the current location. Furthermore, Schwartz (1973) concludes that distance has a weaker effect on migration as education increases. Miller (1973) argues that the mobility of the highly educated produces high out-migration from areas inhabited by well-educated residents. This masks the effect of income on out migration thus making it necessary to control for education when looking at the effect of the economic side on migration.

On the other side of the job market, employers who require better educated employees need to search for them in a wider geographic area according to Miller (1973). The higher the educational requirements of the jobs in an area, the higher the chances are that the people who fill those jobs will come from outside the area. With regard to measuring educational requirements, Miller (1973, p. 7) says, "Educational attainment of the population already living in a state serves as a surrogate for the educational requirements of the jobs in the state." Thus, education appears in both the in-migration equation and in the out-migration equation, but its interpretation in the two equations is not the same. Education is expected to be positively related to both in- and out-migration.

The stage of the life cycle the individual is in is likely to have an important influence on his estimates of the expected utility levels at different cities. Greenwood (1981) argues that older persons are less likely to migrate because their shorter working life lowers the rate of return on migration for them. Job security and family ties are also probably more important for the older person. This makes it less likely that he will migrate. Younger people are expected to be more mobile. A study by Long and Boertlein (1977) shows that for the time period of the present study, 1965-70, the most mobile age group was the 20-34 year old age group. The overall peak was at 20-24 years of age. Mobility did decline with age and reached a low at 70-74 years of age. In order to interpret the effects of the other variables on out-migration correctly, age must be taken into account. One indicator of age is the percentage of the population that is in the 20-34 year old age group (AGE), the most mobile age group. Age is expected to be positively related to out-migration.

Amenities Category

Equation (1) of the previous chapter shows utility depending in part on consumption of non-tradeable goods and services and on leisure time. Non-tradeables include cultural, social, sport, and other types of man-made events that are location fixed. Thus, like climate, also location fixed, social man-made amenities may experience an increase in their demands as the general level of income rises. To satisfy the increased demand for leisure time enjoyment of man-made amenities, an individual may migrate to the city where the amenities most attractive to him are located. The argument is the same as that used by Graves and Linneman (1979) to link climate and migration. Dance, drama, and music events (DDM), sports events (SPORTS), and cultural institutions (CUINS) can be indicators of urban amenities. Those SMSAs with more amenities are expected to be more attractive, other things being equal, and to have higher in- and lower out-migration.

Disamenities Category

As there are attractive aspects to city life, so are there also unattractive aspects, disamenities. A safe and secure environment betters the quality of life while an unsafe and insecure one lowers it. Individuals would form lower expected utility levels for those cities they perceive as being less safe and secure. An indicator of relative safety and security can be the total crime rate (CRIME). The physical environment within which the individual will live is also important. The emergence of measures to deal with pollution is witness to that. There is a considerable difference among cities with regard to the physical environment. An indicator of the cleanliness of the environment can be the state of the air that everyone must breathe. Air pollution can be measured by the mean level of suspended particulates (AIRPOL). Those SMSAs with cleaner air, less air pollution, are expected to have more in-migration and less out-migration. Likewise, the other indicator of disamenities, crime, is expected to be negatively related to in-migration and positively related to out-migration.

An interesting question is whether city size is related to migration. According to Alperovich, Bergsman, and Ehemann (1975), surveys show that smaller cities are preferred. Thus, smaller cities are expected to have higher in-migration and lower out-migration than larger cities. Miller (1973) argues that the larger the job market, the lower the need to look outside the area for work and the lower the need to recruit from outside the area. Smaller cities are expected to have higher in-migration and larger out-migration. Both arguments imply that smaller cities will have more in-migration, but it is indeterminate whether they will have more out-migration. The indicator of city size will be the natural logarithm of population (LSIZE) instead of population (SIZE) on the assumption that is it relative rather than absolute size that matters.

Migration Measure

The indicator of the migration variable itself is a gross rather than a net measure. Schuessler (1972) argues that gross measures are superior to net measures because large in-migration flows tend to be offset by large out-migration flows resulting in a low net migration figure which does not indicate the large amount of migration taking place. Greenwood (1975) observes that variables expected to have the same sign for in- and out-migration tend to cancel out, while variables expected to have opposite signs tend to have their effects exaggerated. In addition, as Graves (1980) notes, a model involving individual decision making is better represented by gross rather than net data. The gross migration figures are, as Cebula (1979) advises, divided by SMSA population to yield the indicator for in-migration (INMIG) and the indicator for out-migration (OUTMIG).

Summary Statistics

Summary statistics for the variables for the sample of 77 SMSAs are presented in Table III. The table shows that growth in median family income from 1960 to 1970 (MFYG) was considerable, 65 percent. Employment showed a slower growth on the average, less than half that of income. The economy grew substantially from 1960 to 1970, but this growth was not evenly distributed. The sample also shows diversity with respect to other variables. Both large and small cities are represented with the largest being 135 times as large as the smallest.

The migration data are for the 1965 to 1970 period. The data for independent variables should be from the pre-migration period. But sometimes the data are not available and so, like Greenwood (1981),

TABLE III

SUMMARY STATISTICS

Variable	Mean	Minimum	Maximum	Standard Deviation
INMIG	1.61	0.34	3.22	0.68
OUTMIG	1.57	0.08	3.13	0.50
MFYG	65.90	36.00	96.00	9.96
MFY/COL	76.20	57.50	92.60	8.17
EMPG	24.80	1.00	115.00	20.10
AIRPOL	10.50	4.00	22.00	3.02
CUINS	5.10	0.00	34.00	5.80
DDM	35.60	0.00	84.00	25.00
SPORTS	7.34	0.00	22.00	6.05
CRIME	31.70	9.00	59.00	11.60
SIZE	834.00	84.00	11366.00	1775.00
EDUC	23.40	11.00	40.00	6.83
AGE	22.80	18.00	33.00	2.79
HEAT	2580.00	114.00	4532.00	8.23
TVAR	33.40	16.40	44.00	6.57
RELH	66.40	31.30	77.00	8.23
WIND	4.11	2.80	5.70	0.68

the closest available time period is used. Thus income and employment growth are for the 1960-70 period. Real income, MFY/COL, is the average nominal median family income for 1960 and 1970 divided by a 1970 cost of living index. The assumption is that the 1970 ranking of cities by cost of living is very similar to the 1965 ranking. Using a 1960-70 average to explain a 1965-70 flow is better than using end-of-the-period 1970 data, as does Liu (1975a), since the influence of migration itself on the independent variables is probably suspiciously high with the 1970 data. It is also better than using 1960 data since 1960 data are further removed from the 1965-70 migration period. For variables such as heating degree days, the data used are appropriate since they reflect underlying long term comparisons between cities.

In-Migration Results

Table IV shows the results of estimating the following equation for in-migration:

$$\mathrm{INMIG} = a_0 + a_1X_{i1} + \dots + a_kX_{ik} + e_i \qquad (1)$$

where $i = 1\text{-}77$,

$k = 1\text{-}14$,

a_0, a_k = coefficients,

X_{ik} = independent variables, and

e_i = error terms.

The hypotheses to be tested are that in-migration is positively related to the economic indicators, to education, and to city amenities, and negatively related to city disamenities. In addition, it is expected that smaller cities will have more in-migration, other things being

TABLE IV

FULL SPECIFICATION MODEL IN-MIGRATION

Variable	Coefficient	Standard Error	t-Statistic	Probability	Computed Elasticity
CONSTANT	2.392022	1.030076	2.32	0.0235	
MFY/COL	0.012433	0.005910	2.10	0.0395	0.5884
MFYG	-0.000600	0.004798	-0.13	0.9009	-0.0006
EMPG	0.008648	0.002772	3.12	0.0027	0.1332
COLD	-0.000147	0.000077	-1.90	0.0616	-0.2356
TVAR	-0.003931	0.015372	-0.26	0.7990	-0.0815
RELH	-0.006899	0.007942	-0.87	0.3884	-0.2845
WIND	-0.045868	0.068949	-0.67	0.5084	-0.1171
EDUC	0.045660	0.006718	6.80	0.0001	0.6636
DDM	-0.000790	0.002724	-0.29	0.7727	-0.0175
SPORTS	0.000392	0.012704	0.03	0.9755	0.0018
CUINS	0.001105	0.007667	0.14	0.8858	0.0035
CRIME	0.000997	0.004306	0.23	0.8176	0.0196
AIRPOL	-0.015574	0.015722	-0.99	0.3257	-0.1016
LSIZE	-0.000028	0.000007	-4.05	0.0001	-1.0315

$\bar{R}^2 = 0.7887$

$\bar{S}^2 = 0.0969$

F = 21.26

F-probability = 0.0001

N = 77

equal. With regard to climate, the relationship is indeterminate, since preferences vary. However, the suspicion is that, on the whole, colder cities will have less in-migration, cities with less variance in temperature will have more in-migration, windier cities will have less in-migration, and cities with more humidity will have less in-migration.

Table IV shows that the model explains more than three fourths of the variation of in-migration. The equation is significant as indicated by an F-value of 21.26, significant at the 0.0001 level. Real median family income (MFY/COL) shows a positive relationship with in-migration, as hypothesized. Its estimated coefficient is significant at the five percent level. This result is consistent with that of Fields (1979) who also found that median family income, deflated by a cost of living index, had a positive and significant relationship with in-migration during the 1965-70 time period. Porell (1982) obtained the same result for the same period using real wages in manufacturing as his income measure. Table IV shows that the elasticity for in-migration with respect to real income computed at the mean indicates that a one percent increase in real income above its mean would have led to a 0.58 percent increase in migration. The computed elasticity for each variable was obtained by dividing the mean of the independent variable by the mean for in-migration and multiplying the result by the regression coefficient (Pindyck and Rubinfeld, 1981).

Growth in income has a negative coefficient and is not significant, having a t-value of only -0.13. It was expected a priori that SMSAs experiencing greater growth in median family income would show relatively higher in-migration. However, the computed elasticity of -0.0006 means that a one percent increase or decrease in income growth was associated with practically no difference in the in-migration rate. In-migration appears to have been responsive to differences in real income but not to differences in growth of income. An explanation may be that, whereas migrants have information on the history of income in their area and can compare current real income in other areas, they do not have information on the history of income in other areas.

The same may be true about employment growth. That is, migrants may be unaware of differences in the growth of employment in other areas. However, Miller (1973) argues that migrants may be drawn to areas with expanding employment even if they have no knowledge of which areas have growing economies. Workers apply for jobs in various areas and the firms that hire them are more likely to be located in areas with relatively rapid growth in employment. Employment growth does show a highly significant positive relationship with in-migration. The study by Alperovich, Bergsman, and Ehemann (1975) for 1965-70 and the study by Miller (1973) for 1955-60 found employment growth to be positively related to in-migration and significant at the five percent level. The computed elasticity for employment growth indicates that a one percent increase in employment growth above its mean was associated with a 0.13 percent increase in migration. In-migration appears to have been more responsive to higher levels of real income than to higher levels of growth in employment.

In-migration was lower for colder cities, as expected, with an estimated coefficient that is significant at the 10 percent level. Graves (1980) used the same variable for the same time period and found a negative relationship for all age groups. The relationship was significant at the five percent level for five out of the seven age groups, significant at the 10 percent level for one of the other two groups, and insignificant for the last group, white males 55-64 years of age. Table IV shows that a one percent increase in heating degree days was associated with a 0.23 percent decrease in migration. Put another way, a temperature one degree lower for six months menas a seven percent colder city and a decrease of 1.6 percent for in-migration.

Cities with pronounced swings in temperature had less in-migration. However, the coefficient is not significant. Graves (1980), using the same measure on a smaller sample, found temperature variance to be positively related to in-migration for white males 15-54 years old with a significant coefficient for four of five age groups and negatively related to in-migration for white males 55 years and older with a significant coefficient for the 65 and up age group. Alperovich, Bergsman, and Ehemann (1975), using a variable that is defined as the deviation from a moderate climate, found that extreme climates were significantly related with less in-migration. Porell (1982) used the first two principal components of a group of climate indicators for which he expected a negative sign. The indices did have significant negative coefficients.

Relative humidity, like temperature variance, has a negative sign and an insignificant coefficient. Graves (1980) found a negative sign for five out of seven age groups of white males. But only the coefficients for the 55-64 and 65 and over groups were significant. Wind speed likewise has a negative and insignificant coefficient. This is consistent with Graves' study since he found wind speed insignificant for all age groups. Its sign was negative only for the 55-64 and 65 and over groups. In-migration appears to have been responsive to better climate. Specifically, warmer cities had relatively more in-migration.

It was hypothesized that cities whose industries required better educated workers would tend to hire relatively more workers from outside the area than other cities. The educational attainment of the population in the city was used as a surrogate for the educational requirements of the industries in the city. It was expected that education would be positively related to in-migration. Miller (1973) found that education did have a positive and highly significant coefficient for the 1955-60 time period. Using the same measure of education as Miller, the present study finds that for 1965-70 migration the relationship found by Miller still held. This is shown by the positive coefficient for education in Table IV, significant at the one percent level. The computed elasticity for education shows that an increase of three percent above its mean of 23.4 percent for education was associated with an increase of two percent for in-migration. This is the second highest elasticity, exceeded only by that for the logarithm of population.

It was expected that SMSAs with more amenities, DDM, SPORTS, and CUINS, would have relatively more in-migration. However, while two of the three indicators, SPORTS and CUINS, were positively related to in-migration, none of the indicators were significant. In addition, the computed elasticities are close to zero. Porell (1982), using an index of city amenities, reported a positive but insignificant relationship between city amenities and in-migration. City amenities do not appear to have influenced in-migration appreciably during 1965 to 1970.

Likewise, city disamenities do not appear to have discouraged in-migration much. Indeed, the total crime rate was positively associated with in-migration but was not significant. This was also found by Porell who used an index of crime indicators and reported a positive and insignificant coefficient for crime. Crime may not have been a deterrent to in-migration because migrants may assume that they will move into a safe area. After all, all cities have both safe and unsafe areas. Air pollution did have the expected negative sign but was not significant either. Porell used an index of air pollution indicators to measure air pollution and found that air pollution was negatively related to in-migration and insignificant.

It was suspected that smaller cities would be preferred and, indeed, smaller cities were associated with more in-migration. The coefficient of the logarithm of population was negative and significant at the one percent level. The elasticity of migration with respect to the logarithm of size shows that a one percent decrease in the logarithm of population was associated with a one percent increase of the in-migration rate. That is, an SMSA of about 334,000 residents would have about 325 more in-migrants than an SMSA of about 354,000 people. Miller (1973) investigated 1955 to 1960 state migration and found the logarithm of population to be negatively related to in-migration and significant at the five percent level. Alperovich, Bergsman, and Ehemann (1975), studying metropolitan areas and 1965 to 1970 migrations, also reported a negative coefficient for the logarithm of population, significant at the five percent level. They found that the preferred city size was 141,000 inhabitants. Their results indicate that the most preferred cities are the moderately small satellite cities surrounding the largest urban centers.

The equation presented in Table V consists of the significant variables from the full model. Compared to the full model, the adjusted R^2 and F-statistic are higher. The five variables of the equation: real income, employment growth, cold, education, and city size, explain more than 80 percent of the variation of in-migration. The qualitative results are the same as those of the full specification model. The variables are significant at the one percent level, except for real income which is significant at the five percent level. The F test for the omission of variables tests the null hypothesis that the coefficients of the omitted variables are equal to zero. Since the calculated F value is not in the critical region, the null hypothesis is not rejected.

Out-Migration Results

Table VI shows the results of the full specification model for out-migration. The out-migration model adds age to the variables included in the in-migration model. The adjusted R^2 is slightly lower than that for in-migration, explaining almost three quarters of the variation in out-migration. Greenwood (1981) and Miller (1973a) also reported lower R^2s for out-migration than for in-migration. The F-statistic is also lower, but the equation is still significant at the one percent level, as was the in-migration equation.

The hypotheses to be tested are that out-migration is lower where economic conditions are better, larger for SMSAs with a higher percentage of residents with one year or more of college, higher where the percentage of people 20 to 34 years of age is larger, smaller where there are more city amenities, and higher where there are more city disamenities. The question is asked, as for in-migration, whether

TABLE V

BEST FIT MODEL IN-MIGRATION

Variable	Coefficient	Standard Error	t-Statistic	Probability	Computed Elasticity
CONSTANT	1.52551	0.361441	4.22	0.0001	
MFY/COL	0.011552	0.005242	2.20	0.0308	0.5467
EMPG	0.009726	0.002188	4.45	0.0001	0.1498
COLD	-0.000178	0.000038	-4.73	0.0001	-0.2844
EDUC	0.047517	0.005911	8.04	0.0001	0.6906
LSIZE	-0.000029	0.000003	-8.81	0.0001	-1.0461

$\bar{R}^2 = 0.8062$

$S^2 = 0.0889$

$F = 64.24$

F-probability = 0.0001

$N = 77$

$F_{\text{joint test for omitted variables}}(9,62) = 0.35 < F_{\text{table, } 0.95} = 2.04$

TABLE VI

FULL SPECIFICATION MODEL OUT-MIGRATION

Variable	Coefficient	Standard Error	t-Statistic	Probability	Computed Elasticity
CONSTANT	0.514648	0.902266	0.57	0.5705	
MFY/COL	0.000899	0.005319	0.17	0.8663	0.0436
MFYG	-0.018678	0.004169	-4.48	0.0001	-0.7840
EMPG	0.000039	0.002412	0.02	0.9871	0.0006
COLD	-0.000193	0.000070	-2.77	0.0074	-0.3172
TVAR	0.025012	0.013579	1.84	0.0703	0.5321
RELH	0.004502	0.007019	0.64	0.5237	0.1904
WIND	0.125990	0.059960	2.10	0.0398	0.3298
AGE	0.060215	0.017365	3.47	0.0010	0.8745
EDUC	0.022043	0.007126	3.09	0.0030	0.3285
DDM	-0.001652	0.002389	-0.69	0.4920	-0.0375
SPORTS	-0.007368	0.011038	-0.67	0.5070	-0.0344
CUINS	0.007887	0.006662	1.18	0.2410	0.0256
CRIME	0.002261	0.003754	0.60	0.5493	0.0457
AIRPOL	0.000282	0.013738	0.02	0.9837	0.0019
LSIZE	-0.000014	0.000006	-2.29	0.0256	-0.5233

$\bar{R}^2 = 0.7100$

$S^2 = 0.0731$

F = 13.40

F-probability = 0.0001

N = 77

smaller or larger cities are preferred. The relationship between the size of the city and out-migration, other things being equal, is indeterminate beforehand. Out-migration is also expected to be higher for colder cities. Although some people may prefer the outdoor activities associated with a cold climate, it is assumed that for most people a cold climate is inferior to a warm one. Likewise, it is assumed that out-migration will be higher for SMSAs with relatively large changes in temperature, with higher humidity, and with higher winds.

Looking at the results in Table VI, real income is not significantly associated with out-migration, as shown by its t-statistic of 0.17. Its computed elasticity of 0.043 indicates that out-migration was not responsive to a change in real income. This finding, that real income has a positive and insignificant coefficient, is consistent with that of Fields (1979) who also studied gross migration for 1965 to 1970. Income growth, on the other hand, is significant and has a negative sign as expected. The computer elasticity of -0.784 indicates that a rate of growth one percent higher than the mean was associated with a 0.78 percent decrease in out-migration. Greenwood (1981) also found income growth significant and negatively related to out-migration. It was argued before, with respect to in-migration, that the reason why income growth was insignificant and had an unanticipated sign may have been a lack of information on the part of potential migrants of the history of income in other areas. Potential migrants, however, were assumed to have knowledge of income growth in their area. Thus, larger rates of income growth would be expected to reduce out-migration. This is what Table VI shows. Out-migration appears to have been influenced by the rate of growth of income and not by the level of real income. Neither was it influenced by the other economic indicator, employment growth. The t-statistic for employment growth is only 0.02. In addition, a one percent increase in employment growth did not elicit a noticeable change in out-migration. The computed elasticity is close to zero, 0.0006. In- and out-migration appear to respond differentially to the economic side. In-migration was influenced by real income and employment growth while out-migration was influenced by income growth.

Out-migration was expected to be lower for warmer cities and higher for colder cities. Heating degree days are significantly related to out-migration, at the one percent level, but the variable has an unexpected negative coefficient. The computed elasticity of -0.3172 indicates that a three percent decrease in heating degree days was associated with about a one percent decrease in out-migration. Graves (1980) also found a negative relationship between heating degree days and out-migration for white males, regardless of age.

Wider swings in temperature were expected to promote more out-migration. Table VI shows that temperature variance was positively related to out-migration and significant at the 10 percent level. The computed elasticity indicates that a one percent increase in temperature variance was associated with a 0.53 percent increase in out-migration. Out-migration was more responsive to extremes in temperature than to any other climate characteristic. Graves (1980) likewise found temperature variance positively related to out-migration for all age groups and significant for five out of seven groups.

Relative humidity, expected to be associated with more out-migration, does show a positive coefficient. However, it is insignificant.

Graves (1980) found relative humidity to have a positive sign for four out of seven age groups. He did not find relative humidity to be significantly related to out-migration for any age group.

Higher wind speed, expected to encourage out-migration, was significantly associated with more out-migration. The computed elasticity indicates that a one percent increase in average wind speed above its mean was associated with about a third of a percent increase in out-migration. Graves (1980) found a positive but insignificant relationship between wind speed and out-migration for white males at all ages. Porell (1982), using two climate indexes expected to be positively related to out-migration, also found the relationship to be positive but insignificant. The results in Table VI indicate a slightly more important role for climate in determining out-migration than do those of Graves and a much stronger one than do those of Porell. Temperature variance, relative humidity, and wind speed were all insignificant for in-migration, but only relative humidity was insignificant for out-migration. Out-migration appears to have been more responsive to climate differences than was in-migration.

Out-migration was expected to be higher for cities with relatively more residents in the 20-34 year old age group since this group is a relatively mobile group. The results show a very significant positive relationship between age and out-migration. The computed elasticity indicates that a one percent increase in the percentage of people 20-34 years old in a city was associated with a 0.87 percent increase in out-migration. It appears that, other things being equal, a population at risk that contains a relatively higher percentage of residents who are investing in education and initiating a career will emit relatively more migrants.

Education was also expected to be positively related to out-migration reflecting a larger geographic job market for the better educated. The coefficient for education is positive and significant at the one percent level. Miller (1973a) found the percentage of the population in a state with one year or more of college to be positively related to 1955-60 out-migration and significant at the five percent level. The computed elasticity for education indicates that a one percent increase in education was associated with a 0.33 percent increase in out-migration. The mobility of the 20-34 year old age group appears to be more important in promoting migration than educational attainment since the computed elasticity for age is over two and one half times that of education.

More city amenities were expected to reduce out-migration. Two of the three amenity indicators, dance, drama, and music events and sports events do have a negative coefficient, with cultural institutions having a positive coefficient, but all are insignificant. The computed elasticities are also all close to zero. Porell (1982) found an unexpected positive sign for an index of city amenities, however, the relationship was insignificant. City amenities do not appear to have influenced out-migration much during 1965 to 1970.

City disamenities were expected to be positively related to out-migration. Crime and air pollution have positive coefficients, but they are insignificant. The computed elasticities show out-migration was unresponsive to differences in disamenities. Porell (1982) found an index of crime measures to have an unexpected negative and insignificant coefficient while an air pollution index had a positive sign and was also insignificant. Thus, disamenities do not seem to have promoted out-migration.

Relative city size is negatively related to out-migration with a coefficient significant at the five percent level. The computed elasticity of -0.53 would indicate a one percent increase in out-migration would result from a two percent decrease in the logarithm of population. Miller (1973a) also found the logarithm of population to have a significant negative relationship with out-migration. This result is consistent with Miller's argument that the larger an area, the less need to look outside the area for a job.

The equation presented in Table VII consists of the significant variables from the full specification model. Compared to the full model, the adjusted R^2 and the F-statistic are higher. The seven variables of the equation: income growth, cold, temperature variance, wind, age, education, and the logarithm of city size explain almost three fourths of the variation in out-migration. Most variables are significant at the one percent level, with temperature variance and wind significant at the five percent level. The signs are the same as in the full model and the computed elasticities are almost the same. Only temperature variance has a noticeably different, lower, elasticity. The F test for the omission of variables indicates that the null hypothesis that the coefficients of the omitted variables are equal to zero cannot be rejected.

Summary and Conclusions

In this chapter, the sample, the variables, summary statistics, and

TABLE VII

BEST FIT MODEL OUT-MIGRATION

Variable	Coefficient	Standard Error	t-Statistic	Probability	Computed Elasticity
CONSTANT	1.278535	0.496867	2.57	0.0122	
MFYG	-0.018248	0.003067	-5.95	0.0001	-0.7660
COLD	-0.000173	0.000043	-3.98	0.0002	-0.2836
TVAR	0.017350	0.007387	2.35	0.0217	0.3691
WIND	0.128658	0.053397	2.41	0.0186	0.3368
AGE	0.060248	0.015037	4.01	0.0002	0.8749
EDUC	0.020544	0.006101	3.37	0.0012	0.3062
LSIZE	-0.000017	0.000003	-5.63	0.0001	-0.6503

$\bar{R}^2 = 0.7269$

$S^2 = 0.0689$

$F = 29.89$

F-probability = 0.0001

$N = 77$

$F_{\text{joint test for omitted variables}}\ (8,61) = 0.49 < F_{\text{table, } 0.95} = 2.10$

regression results were presented and discussed. The results were compared and contrasted with those of other studies. Best fit models were presented, incorporating the significant variables, for in- and out-migration.

In- and out-migration appear to have different determinants. Economic factors affect both, however, in-migration is responsive to real income and to employment growth while out-migration is responsive to income growth. Climate is an important determinant of migration, but out-migration seems to be responsive to more dimensions of climate than in-migration. The stage of the life cycle was found to be an important qualifier for out-migration, one that is not always taken account of. Another important qualifier is education which reflects the educational attainment of workers in the out-migration equation and the educational requirements of employers in the in-migration equation. Environmental considerations are important to migrants, but only those, like climate, that are the same throughout the SMSA, and not those, like crime, that are unevenly distributed. As Graves (1976) argues, non-global factors are more likely to affect intra-SMSA migration than inter-SMSA migration. City quality of life amenities do not appear to be important determinants of migration. If what Alperovich, Bergsman, and Ehemann (1975) report is true, people migrate to the smaller satellite cities of the large urban centers. Thus, they stay within tolerable driving distance of sport and cultural amenities.

Overall, the results imply a more effective role for city policy actions with regard to migration, since the importance of economic factors in the migration decision was confirmed, than do the results of some other studies, such as Graves' (1980). In this they agree with the results of Porell (1982) who found both economic and quality of factors important determinants of migration.

CHAPTER IV

THE LATENT VARIABLE MODEL

Introduction

The model in Chapter II, represented by equation (4), shows the categories influencing the migration decision of the population at risk. For in-migration, the categories describe destination characteristics that attract migrants. For out-migration, the categories describe origin characteristics that promote migration. The demographic category relates directly only to the population at risk, that is, at the origin. Thus, the demographic category applies directly only to out-migration. The age structure of the population at destinations does not affect the migration decision of the population at the origins. Age, then enters into the out-migration equation but not into the in-migration equation. However, the national job market that goes along with higher educational attainment of workers has a counterpart in the national job market that goes along with jobs that require more highly educated workers. Thus, the out-migration that higher education promotes is complemented by the in-migration that higher educational work requirements induce. Since educational requirements of jobs at the destination are represented by the educational attainment of the population at the destination, education appears in both the in- and out-migration equations. Thus, the demographic category enters directly into the out-migration model but only partially and indirectly into the in-migration model.

The economic category was represented in Chapter III by real income, growth in income, and employment growth. The climate category was represented by cold, temperature variance, relative humidity, and wind speed. The amenities category was represented by dance, drama, and music events, by sports events, and by cultural institutions. The disamenities category was represented by crime and air pollution.

The relationship between migration and the general categories has not been estimated in Chapter III. Instead migration was regressed on the indicators representing each category and on city size as a control variable. The ideal estimation procedure, however, would not regress migration on the indicators, but on the categories themselves. Estimated coefficients could then be obtained for the categories as represented by the indicators. It would be recognized that the indicators do not represent the multi-dimensional categories perfectly, but rather measure them with some error. An idea could be obtained of the relationship between the indicators and the categories they represent. The latent variable model, where the categories of equation (4) are called latent variables, allows this type of estimation. It will be employed in the following chapter.

This chapter provides a general overview of the latent variable model illustrating the different aspects of specification and estimation via migration examples. Advantages and disadvantages of the technique are presented. The latent variable model is described in detail by Bagozzi (1980). Most of the following discussion is adapted directly from Bagozzi.

Causal Diagrams

Causal diagrams are helpful in summarizing the specification of a latent variable model. Figure 1 helps describe both the causal diagram and the latent variable model. It shows relationships at two levels. First, it shows the theoretical relationship between climate and migration:

$$\text{MIGRATION} = c_1 \text{ CLIMATE} + \varepsilon \tag{1}$$

where c_1 = coefficient and

ε = error term.

The error term signifies the variation in migration not accounted for by climate. At a second level, Figure 1 shows the relationship between the observable indicators and the latent variables, CLIMATE and MIGRATION. Therefore, it shows the indirect relationship between the climate indicators, temperature and temperature variance, and the migration indicator, gross out-migration. The relationship between the indicators and the latent variables can be represented by:

$$\text{gross out-migration} = \text{MIGRATION} \tag{2}$$

$$\text{temperature} = a_1 \text{ CLIMATE} + e_1 \tag{3}$$

$$\text{temperature variance} = a_2 \text{ CLIMATE} + e_2 \tag{4}$$

where a_1, a_2 = coefficients, and

e_1, e_2 = measurement error terms.

In this model, migration is assumed to be measured without error. Climate, however, is measured with error by temperature and temperature variance. Variation in climate will be a source of variation in

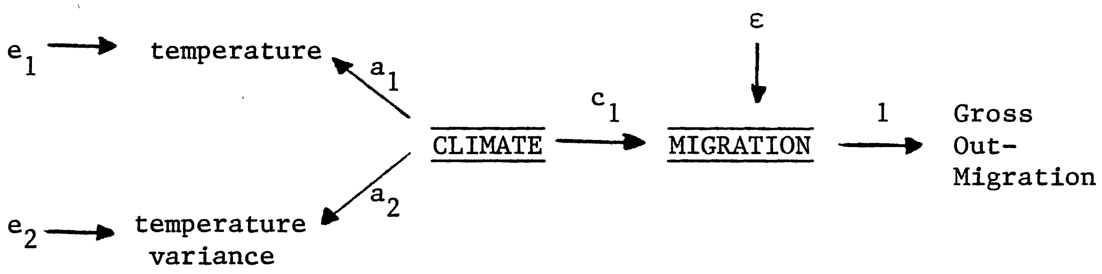

Figure 1. Measurement Error in the Independent Variable

temperature and temperature variance. The closer the degree of correspondence between climate and temperature, the higher will be the coefficient, a_1, and the lower will be the measurement error term, e_1, which reflects other sources of variation in temperature. The closer the correspondence between climate and temperature, the better temperature serves as a measure of climate. The coefficients, a_1 and a_2, show how well, relatively, each indicator measures climate (Aaker and Bagozzi, 1979).

Specification of the Latent Variable Model

The addition of an independent latent variable measured by two indicators, say an economic variable, to the model in Figure 1 gives the model in Figure 2. The model in Figure 2 can be summarized as follows:

$$N = c_1L_1 + c_2L_2 + \varepsilon \tag{5}$$

$$x_1 = a_1L_1 + e_1 \tag{6}$$

$$x_2 = a_2L_1 + e_2$$

$$x_3 = a_3L_2 + e_3$$

$$x_4 = a_4L_2 + e_4$$

$$y = d_1N + u_1$$

where $d_1 = 1$ and

$u_1 = 0.$

Equation (5) shows the relationship between migration, N, and the climatic and economic latent variables, L_1 and L_2, respectively.

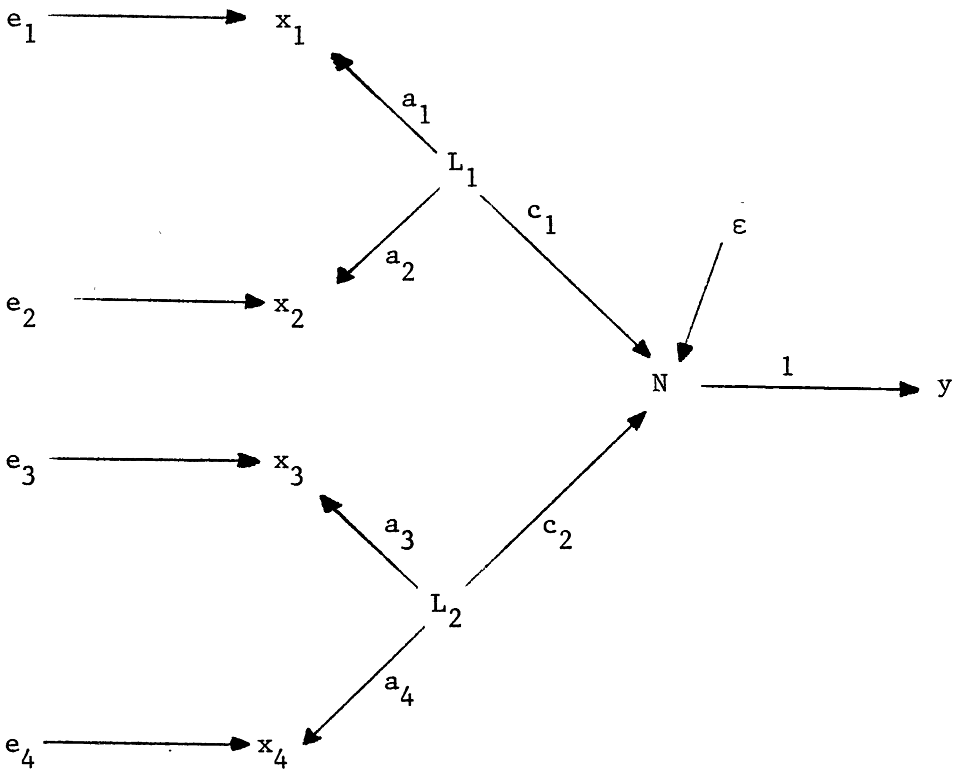

Figure 2. A General Latent Variable Model

Equations (6) link the indicators to the latent variables. Climate is measured by x_1 and x_2 with error; the economic factor is measured by x_3 and x_4 with error; and migration is measured by y without error. In matrix form the system can be written as:

$$(N) = (c_1 c_2) \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} + (\varepsilon) \tag{7}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} a_1 & 0 \\ a_2 & 0 \\ 0 & a_3 \\ 0 & a_4 \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix} \tag{8}$$

$$(y) = (d_1)'(N) + (u_1) \tag{9}$$

A general model would include the relationships between one or more dependent latent variables and one or more independent latent variables. It would also describe the relationships between the latent variables and their indicators. The general model representing a set of simultaneous linear equations is:

$$\underset{mxm}{B}\ \underset{mx1}{N} = \underset{mxn}{C}\ \underset{nx1}{L} + \varepsilon \tag{10}$$

where B is a matrix of coefficients showing the relationship between the dependent latent variables. The model can include more than one dependent variable and allow feedback effects between the dependent variables. C is a matrix of coefficients describing the relationships between the independent and the dependent latent variables and L is a vector of independent latent variables. ε is a vector of residuals. The latent variables can be linked to observations by:

$$\underset{qx1}{X} = \underset{qx1}{V_x} + \underset{qxm}{A_x}\ \underset{mx1}{L} + \underset{qx1}{e} \qquad (11)$$

$$\underset{px1}{Y} = \underset{px1}{W_y} + \underset{pxm}{D_y}\ \underset{mx1}{N} + \underset{px1}{u} \qquad (12)$$

where X and Y are vectors of observed indicators of independent and dependent latent variables, V_x and W_y are the respective vectors of means for X and Y, A_x and D_y are regression matrices, and e and u are vectors of measurement errors in X and Y respectively. An assumption is made that the E(e) = E(u) = E(Le) = E(Nu) = 0, and that E(e'e) = θ_e^2, E(u'u) = θ_u^2 where θ_e^2 and θ_u^2 are diagonal matrices (Bagozzi, 1980).

The varianc-covariance matrices of L and ε are Φ(nxn) and Φ(mxm) respectively. The general form of the variance-covariance matrix (VC) is (Bagozzi, 1980):

$$VC = \begin{bmatrix} D_y(B^{-1}C\Phi C'B'^{-1} + B^{-1}\Phi B'^{-1})D_y' + \theta_e^2 & DDD_yB^{-1}C\Phi A_x' \\ A_x\Phi C'B'^{-1}D_y' & A_x\Phi A_x' + \theta_u^2 \\ & (p+q)x(p+q) \end{bmatrix} \qquad (13)$$

Estimation

The parameters are estimated by the maximum likelihood technique. The vector of observations z = (x',y') is assumed to have a multivariate normal distribution with mean vector (w',v') and variance-covariance matrix VC. With M observations of z(z_1, z_2, ..., z_m) and $\bar{z} = (\bar{y}',\bar{x}')'$ representing the maximum likelihood estimates of the mean vector, the sample variance-covariance matrix can be written as:

$$S = \frac{1}{N}\sum_{i=1}^{N}(z_i-\bar{z})(z_i-\bar{z})' \qquad (14)$$

where N = M-1.

The logarithm of the likelihood function (omitting a constant term) can be written as:

$$\log L = -\tfrac{1}{2} N[\log|VC| + tr(S\ VC^{-1})] \qquad (15)$$

where tr = trace.

The goal is to find values for the independent parameters in VC that maximize the value of log L. A more convenient and equivalent method is to minimize the following function F, which is -2/N times log L (plus the constant term) (Bagozzi, 1980).

$$F = \log|VC| + tr(S\ VC^{-1}) - \log|S| - (p+q) \qquad (16)$$

The values of the parameters that minimize F cannot be found analytically. An iterative procedure was developed by Jöreskog and colleagues. The computer program LISREL (Jöreskog and van Thillo, 1973) calculated the maximum likelihood and standardized estimates of the parameters in VC as well as their standard errors.

Maximum likelihood estimators have some advantages. First, they are asymptotically efficient. Second, the maximum likelihood method is independent of the scales of measurement of the variables in one's model (Lawley and Maxwell, 1971). Third, maximum likelihood estimates are robust over nonnormality. Fourth, a convenient statistic, described in the following section, exists for testing one's model.

Hypothesis Testing

In specifying a model, theory is the proper guide. The model should then be tested to see if it fits the data. An overall goodness-of-fit test is provided by the maximum likelihood estimation method

(Aaker and Bagozzi, 1979). The null hypothesis is that the specified model, with its restrictions on the variance-covariance matrix is correct. The alternative hypothesis is that there are no restrictions on the true population variance-covariance matrix (Aaker and Bagozzi, 1979). Let L_{H_0} be the maximum of L in VC under H_0. Then (Bagozzi, 1980):

$$\log L_{H_0} = -\tfrac{1}{2} N[\log|\widehat{VC}| + \mathrm{tr}(S\ \widehat{VC}^{-1})] \tag{17}$$

where $\widehat{VC}$ stands for the value of parameters that maximize the value of L. For the alternative hypothesis H_1, that VC is any positive definite matrix:

$$\log L_{H_1} = -\tfrac{1}{2} N[\log|S| + p\ q] \tag{18}$$

because log L reaches a minimum under H_1 when VC = S. The likelihood ratio $\lambda = L_{H_0}/L_{H_1}$ can be used to form a chi square statistic since $-2 \log \lambda$ is distributed approximately chi square for large samples if H_0 is true (Thiel, 1971). In addition,

$$-2 \log \lambda = N\ F_0 \tag{19}$$

where F_0 is the minimum value of F from the previous section. The chi square test is distributed with degrees of freedom equal to:

$$\text{d.f.} = \tfrac{1}{2}(p + q)(p + q + 1) - t \tag{20}$$

where t is the number of parameters to be estimated under H_0. The one-tailed test of significance is:

$$X^2 \geq X^2_{\alpha:\text{d.f.}} \tag{21}$$

where H_0 is accepted at the α level if the above equation holds, while it is rejected if the above equation does not hold.

However, a correction factor can be used to improve the chi square approximation (Bagozzi, 1980). Application of the correction factor changes $X^2 = N F_0$ to:

$$X^2 = [N - \frac{1}{6}(2p + 2q + 5) - \frac{2}{3}(m + n)]F_0 \tag{22}$$

$$= N^* F_0$$

where $N = M-1$,

M = number of observations,

p = number of observable endogenous variables,

q = number of observable exogenous variables,

m = number of unobservable endogenous variables, and

n = number of unobservable exogenous variables.

The chi square approximation is more trustworthy if $N - (p+q) \geq 50$.

In general, then, to see how well the hypothesized model fits the data, a comparison is made between $\hat{VC}$ and S. If the fit is good, the residual matrix VC-S is small. This information is then used with the standard errors of parameter estimates to evaluate the model (the t-test applies) (Bagozzi, 1980).

If a model does not show a good fit, theory may suggest alterations that can be made to improve the model. If a model performs well, then a simpler model might be investigated to see if it also fits the data. Changes would be guided by theoretical considerations. Two models, one a more general model than the other, can be compared by estimating each one separately and looking at the difference in their chi square statistics. This difference is itself a chi square with degrees of freedom equal to corresponding difference in degrees of freedom (Bagozzi, 1980).

A model may give a poor fit because the hypothesized set of causal paths is inappropriate. It may also give a poor fit because some measurement error terms are not independent. If a certain relationship between error terms is suspected, then this relationship can be made explicit in a model and this model can be compared to the model without correlated error terms. A significant difference in chi square values would provide empirical support (but not prove) that a link should be made between the error terms (Aaker and Bagozzi, 1979).

Advantages and Disadvantages of the Latent Variable Model

The latent variable model has the advantage that it can represent the most complicated set of relationships simultaneously at both the theoretical level and the level of observations. The latent variable model forces the theorist to make explicit all the relationships involved in the theory and, by doing so, can aid in theory construction. With the latent variable approach one can obtain estimates of the relationships between the unobservable variables and their indicators, of the error terms associated with the dependent variables, and of the indicator error terms. According to Bagozzi (1980, p. 107), "no other approach in the behavioral sciences yields as much information."

With regard to the disadvantages, sufficient conditions for identification of the latent variable model have not so far been established (Bagozzi, 1980). General rules for identification have been derived only for special cases such as the MIMIC model (Jöreskog and Goldberger, 1975), for a linear dynamic system with measurement error in both endogenous and exogenous variables (Hsiao, 1975, 1976, 1979), and for a small number of other models (Geraci, 1975; Wright, 1970).

Another disadvantage is that parameter estimates are efficient only for large samples, i.e., $50-60 < N < 300$. Also, the chi square goodness-of-fit test, a large sample approximation, is directly sensitive to sample size (Bagozzi, 1980). If the sample is large enough, the chi square test will lead to rejection of the model. But several things can be done to deal with this problem. The residual matrix can be used to evaluate the model. A series of models, each a special case of the preceding one, can be compared by using the difference in chi square test (Bagozzi, 1981). Bentler (1981) developed an incremental fit index to check for an improvement in fit between any two models (Bagozzi, 1981).

Summary and Conclusions

This chapter provided a summary description of the latent variable model. First, causal diagrams were presented and the specification of the latent variable model was discussed. The maximum likelihood technique used to estimate the parameters of the model was presented next. Hypothesis testing was then discussed. The chapter ended with comments on the advantages and disadvantages of the model.

The general categories of the migration model in Chapter II are probably not measured perfectly by any one indicator. The general latent variable model thus appears to be appropriate in analyzing the general determinants of migration.

CHAPTER V

LATENT VARIABLE RESULTS

Introduction

In this chapter, the latent variable approach is used to estimate the migration model of equation (4) in Chapter II. The model is for the population at risk, that is, for the population at the origin and thus for out-migration. All the categories of equation (4) are included in the out-migration equation estimated in Chapter III. The in-migration equation, however, does not include the demographic category since the demographic category applies to the population at risk which is not the population at the destination. Since the full migration model of the study, with all the categories, is an out-migration model, and since the reason for using the latent variable approach is to estimate the relationship between migration and the categories--the latent variable--only out-migration will be modelled and estimated in this chapter. The full specification out-migration model presented in Table VI of Chapter III will be placed in the context of the latent variable model and estimated. The empirical results will be presented, model evaluation criteria described, and the results compared and contrasted with those of the multiple regression analysis. A summary and conclusions end the chapter.

Full Specification Latent Variable Model

The full specification out-migration model of the previous chapter can be placed in the context of the latent variable model by combining the economic, climatic, demographic, amenity, and disamenity groups of indicators, considering each group a latent variable, and holding city size constant as measured by the logarithm of population. The model can be specified as in Figure 3. The model shows how each latent variable affects migration, as is indicated by the coefficients c_1-c_6 and it shows the error term ε associated with the relationship. The error term ε is also the error term for the model as a whole. The model also reveals how well each indicator, relative to the other indicators, measures the latent variable, as indicated by the coefficients a_1-a_{14}, and the error terms e_1-e_{14}. Migration and city size are single indicator latent variables, as opposed to the rest which are multiple indicator latent variables. Therefore, they are assumed to be measured without error as shown by the indicator coefficients set equal to one and the omission of error terms.

The economic variable is represented by real income, income growth, and employment growth. Better economic conditions are expected to be associated with lower out-migration. The climate variable is indicated by cold, temperature variance, relative humidity, and wind speed. Better climate is expected to be related to less out-migration. Thus climate is expected to have a positive coefficient. The demographic variable consists of age and education. Higher percentages of residents in the more mobile age and education groups are expected to promote more out-migration. The amenities variable is measured by dance, drama, and music events, by sports events, and by cultural institutions.

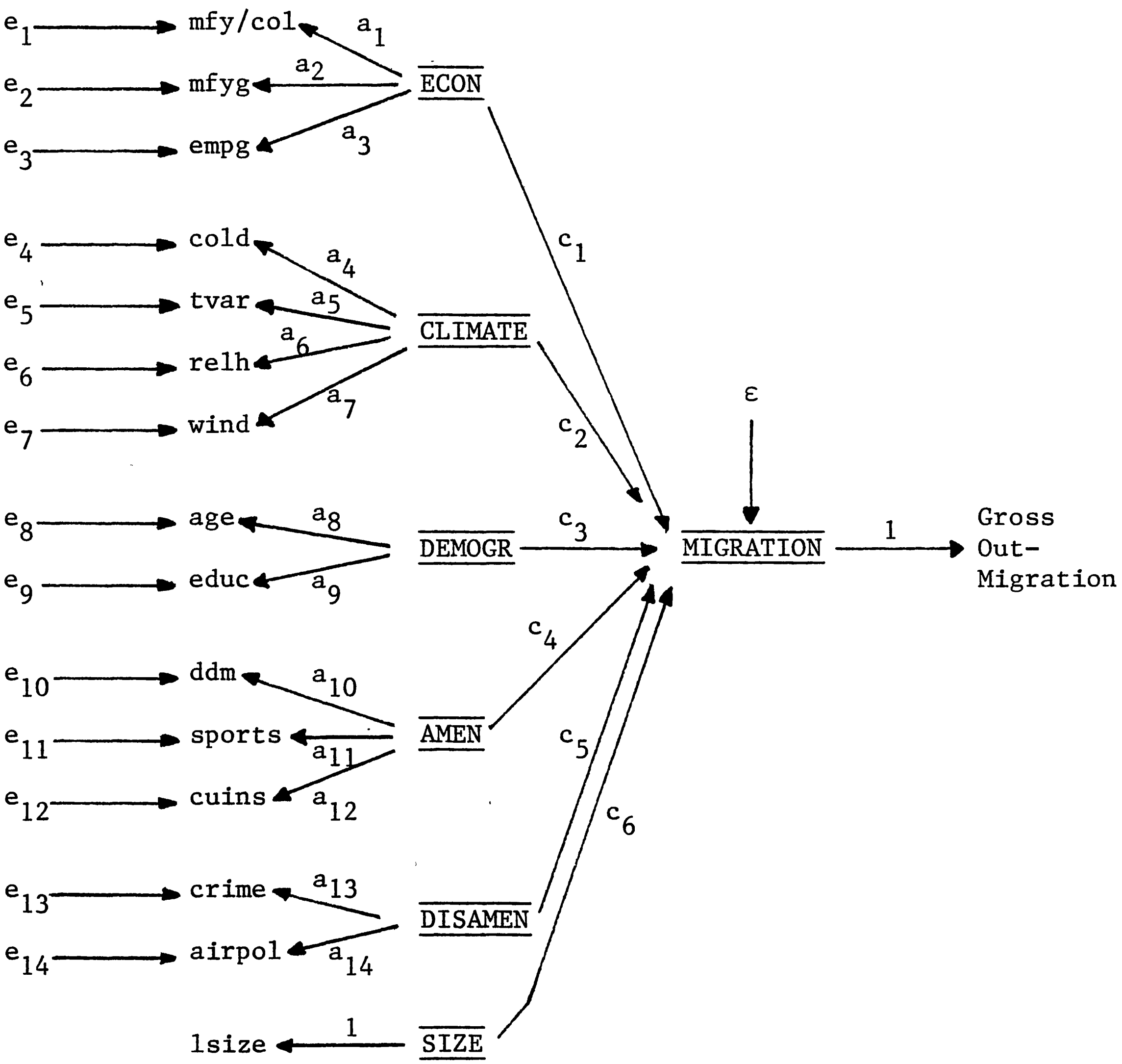

Figure 3. Full Latent Variable Model

Cities with more amenities are expected to have lower out-migration. The disamenities variable is represented by crime and air pollution. More disamenities are expected to be associated with more out-migration. The relationship between city size and out-migration is indeterminate *a priori*.

Table VIII shows the empirical results of estimating the model in Figure 3 using the LISREL program (Jöreskog and van Thillo, 1973). The model does not fit the data well as shown by the high chi square statistic and correspondingly low probability of fit of 0.0000. As mentioned in Chapter IV, a model is accepted as having an adequate fit if the probability of fit is at least 0.10. In addition, the error term for the model as a whole is high, 0.7616, and insignificant. Since the error term is equal to $1-R^2$, $R^2 = 0.2384$, which is low. Notice that the degrees of freedom, 85, are equal to the number of variances and covariances of the observables, 120, minus the number of parameters and error terms to be estimated, 35 (Bagozzi, 1980). That is, the degrees of freedom are equal to the number of elements used to estimate minus the number of elements to be estimated.

The beta coefficients of Table VIII show the change in standard deviations of the dependent variable due to a change of one standard independent variable. The economic variable was expected to be negatively related to out-migration. Instead, it has a positive coefficient and is insignificant. The multiple regression results in Table VI of Chapter III, show that real income and employment growth both had positive signs and were insignificant. Income growth, however, was significant and negatively related to out-migration as expected. The coefficients for the indicators in Table VIII show that employment

TABLE VIII

FULL LATENT VARIABLE MODEL OUT-MIGRATION

Variable	Beta Coefficient	Standard Error	t-Statistic
ECON	0.2427	6.2150	0.1593
CLIMATE	0.0577	1.7092	0.0858
DEMOGR	-0.0887	4.0623	-0.0501
AMEN	-0.5497	4.6481	-0.3334
DISAMEN	0.3721	4.4893	0.3539
SIZE	0.0638	3.9255	0.0279
Error Term	0.7616	1.1377	1.9771

Chi Square Statistic = 474.8802
Probability of Fit = 0.0000
D.F. = 85
N = 77

Indicator	Beta Coefficient	Standard Error	t-Statistic	Error Term
Economic Latent Variable:				
mfy/col	0.4212			0.8110
mfyg	0.2513	0.4234	1.4091	0.9872
empg	0.5163	0.5784	2.1194	0.7111
Climate Latent Variable:				
cold	0.6758			0.5247
tvar	0.4972	0.2701	2.7237	0.6703
relh	-0.0908	0.2241	-0.5996	1.0972
wind	0.4269	0.2477	2.5498	0.7215
Demographic Latent Variable:				
age	0.7488			0.4203
educ	0.5748	0.2410	3.1854	0.5798
Amenities Latent Variable:				
ddm	0.6095			0.5857
sports	0.7424	0.2473	4.9260	0.5029
cuins	0.5487	0.2207	4.0783	0.6496
Disamenities Latent Variable:				
crime	0.4025			0.8260
airpol	0.1293	0.3150	1.0197	1.1361

growth is the most important indicator of the economic latent variable with a coefficient of 0.5163, followed by real income and income growth with coefficients of 0.4212 and 0.2513 respectively. Employment growth is significant with a t-statistic of 2.1194, while income growth is not significant, having a t-statistic of 1.4091. Table VIII does not show a standard error or a t-statistic for real income. This is because the LISREL program used to estimate the model needs a scale of reference for the economic latent variable. This is done by choosing the units of measurement of one of the indicators, real income in this case, as the scale of measurement for the latent variable (Jöreskog and van Thillo, 1973). As a result, the program does not provide the standard error or t-statistic for that indicator, but does provide a coefficient and error term. Within a group of indicators, the higher the coefficient, the lower the error term, and the higher the t-statistic. Thus, employment growth has a higher coefficient than income growth, a lower error term, and a higher t-statistic. Real income, the scale of reference for the economic variable, has a larger coefficient than income growth. Its error term, which is shown, is smaller than that of income growth, and its t-statistic, which is not shown, is larger than that of income growth. Real income is probably also significant.

This comparison of coefficients only holds within groups of indicators and not across groups. Temperature variance, for example, has a lower coefficient than employment growth, 0.4972, but a higher t-statistic, 2.7237, and a lower error term, 0.6703.

For the economic latent variable, the finding is that employment growth is the best indicator, followed by real income, with income growth appreciably worse than the other two and insignificant. The reliability of the indicator finding can be questioned in view of the poor fit of the model.

Climate was expected to be positively related to out-migration and does have a positive coefficient, however, it is insignificant. The results from Chapter III, Table VI, show that three out of the four indicators of climate, except relative humidity, were significant and three, except cold, had a positive sign. Both the regression and latent variable models show a positive relationship between out-migration and climate. When the climate indicators were entered as separate regressors, they were found to be significantly related to out-migration. However, when these indicators represented climate in the latent variable model, climate was not found to be significantly related to out-migration. The poor fit of the model casts doubt on the latter finding. The indicator coefficients show that cold is the best indicator of climate, followed by temperature variance and wind speed with similar coefficients. Since the t-statistic for cold is larger than 2.7237, all three are significant. Relative humidity does not measure well at all. It has a -0.0908 coefficient and is insignificant.

The demographic variable was expected to have a positive sign, but has a negative sign instead and is insignificant. This contrasts sharply with the multiple regression results since age and education both had positive coefficients and were highly significant. Both have significant indicator coefficients in Table VIII with age being the better indicator of the two.

The amenities latent variable was expected to be negatively related to out-migration with out-migration being lower for SMSAs which have more amenities. The estimated coefficient is negative and insignificant.

The amenities indicators were also insignificant in Chapter III as separate regressors and two out of the three had negative coefficients. The indicator coefficients are significant for all three with sports events having the highest coefficient followed by dance, drama, and music events and by cultural institutions, respectively, with similar coefficients.

The disamenities latent variable is positively related to out-migration, as expected, but is insignificant. Its two indicators, crime and air pollution, were also positively related to out-migration and insignificant in the multiple regression analysis. According to the indicator coefficients, crime is the only indicator of the two that is a significant measure of disamenities.

The relationship between city size and out-migration was indeterminate a priori. The logarithm of population has a positive coefficient and is insignificant. This is a very different result from that of multiple regression analysis which found city size to be negatively related to out-migration and highly significant. The finding from Chapter III agrees with Miller's (1973) argument that areas with larger job markets should have less out-migration since workers have less need to look outside the area for a job. The latent variable finding agrees with public opinion studies which show that smaller cities are preferred (Alperovich, Bergsman, and Ehemann, 1975) and imply that larger cities, other things being equal, should have more out-migration.

The extremely poor fit of the model puts the above results in question, however. The poor fit indicates that the specification is not correct. Darden (1981) points out that a low probability of fit tells more than a high probability of fit. Its implication is clearcut. A high probability of fit does not mean that the specification is necessarily correct, but a low probability of fit does mean that the specification is incorrect.

All models in which the five latent variables were multiple indicator variables produced a poor fit. The goal of estimating the relationship between out-migration and the economic, climatic, demographic, amenity, and disamenity categories could not be achieved. Because of this, comparison with the multiple regression results is difficult.

A Model with Two Multiple Indicator Variables

A cause of the poor fit might be that extraneous indicators are included in the model (Bagozzi, 1980). The incorrect specification of paths linking indicators to latent variables may lead to the poor fit. That is, the full model may contain too many multiple indicator variables. This possibility was examined by respecifying the model. To evaluate the respecified model, the chi square statistic is used. In addition, the difference in chi squares between the simpler model and the full model can be used since, as Bagozzi (1980, p. 105) quotes Jöreskog and Sörbom, "A large drop in X^2, compared to the difference in degrees of freedom, supports the changes made."

In the respecified model the number of multiple indicator variables was reduced from five to two: the demographic and climatic latent variables. Even then, only two indicators were used for climate. Cold and temperature variance were chosen since they reflect the most salient aspects of climate. When studies use only one measure of climate, it is usually the level of warmth or cold, as in Cebula and Vedder (1976), or some measure of swings in temperature, as in Alperovich, Bergsman, and Ehemann (1975). These indicators also were the two best indicators of climate in Table VIII. Income growth was chosen as the sole economic indicator since it was the only significant economic measure for out-migration in Chapter III. Dance, drama, and music events represent city amenities since it is the most inclusive of the amenity indicators. A disamenity indicator is not included since an adquate fit could not be obtained when one was included. The same thing is true for city size. The model is shown in Figure 4.

Table IX presents the results of estimating the model. It performs much better than the full latent variable model in Table VIII in terms of the chi square statistic, the probability of fit, and the error term. The probability of fit is good at 0.3725, well above the 0.10 minimum for an adequate fit. The error term for the model is 0.2983, much lower than the error term for the full latent variable model, 0.7616. The implied unadjusted R^2 for the model of 0.7017 is much higher than the 0.2384 of the full model. The reduction in the chi square statistic is 467 which is almost six times the difference in degrees of freedom, 78. Following Jöreskog and Sörbom's criterion, the reduction in the chi square statistic, being large compared to the difference in degrees of freedom, provides support for the changes made.

Cities with higher income growth had significantly less out-migration, as expected. This contrasts with the positive and insignificant coefficient of the economic latent variable in the full model but agrees with the negative and significant coefficient for

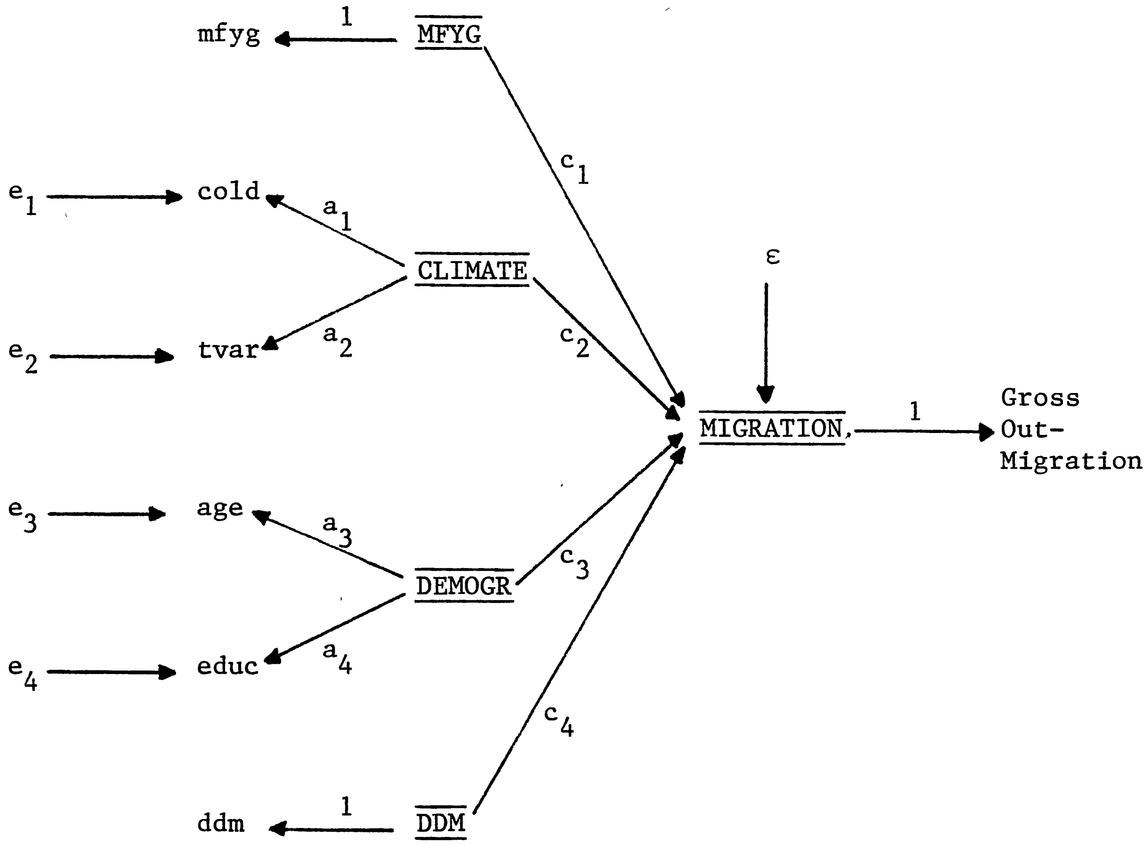

Figure 4. A Model with Two Multiple Indicator Variables

TABLE IX

OUT-MIGRATION MODEL WITH TWO MULTIPLE INDICATOR VARIABLES

Variable	Beta Coefficient	Standard Error	t-Statistic
MFYG	-0.3593	0.0739	-4.8851
DDM	-0.3259	0.0778	-4.2084
CLIMATE	-0.0540	0.1150	-0.5752
DEMOGR	0.7261	0.1262	6.3966
Error Term	0.2983	0.0756	3.9752

Chi Square Statistic = 7.5647
Probability of Fit = 0.3725
D.F. = 7
N = 77

Indicator	Beta Coefficient	Standard Error	t-Statistic	Error Term
Climate Latent Variable:				
cold	0.8196			0.3011
tvar	0.6019	0.4483	1.6382	0.6036
Demographic Latent Variable:				
age	0.9026			0.1997
educ	0.7158	0.1229	6.4518	0.4869

income growth in the multiple regression analysis of Chapter III. The beta coefficient indicates that an increase of one standard deviation in income growth, 10 percentage points, was associated with a decrease of 0.36 standard deviations in out-migration, about one migrant per 3,000 residents.

Better climate was expected to be associated with less out-migration, but instead, it is associated with more out-migration. This is unlike the full model in which climate was positively associated with out-migration. In both models, however, the association is insignificant. The two indicators of climate, cold and temperature variance, were significant in multiple regression analysis with cold having a negative coefficient and temperature variance a positive one. The latent variable results do not agree with the multiple regression results since they show that climate is not an important determinant of out-migration. The indicator coefficients in Table IX show that, as in the full model, heating degree days is a more important climate indicator than temperature variance. Indeed, as in the full model, the indicator for heating degree days is 36 percent higher than that for temperature variance. However, the importance of the indicator findings for a latent variable that is not significant is open to question.

The demographic latent variable was found to be positively related to out-migration, as expected, and significant. This is different from the negative and insignificant relationship found in the full model and agrees with the multiple regression results in which age and education had positive and significant coefficients. As was the case for cold and temperature variance, age and education show the same relationship as indicators of the latent variable in Table IX as they did in Table

VIII. Again, the levels of the coefficients are higher than before, but the relative size is about the same. The poor fit of the full latent variable model does not seem to have affected the interpretation of which indicator is the more important measure of the latent variable.

Dance, drama, and music events has a negative coefficient, as expected. This is consistent with its negative coefficient in regression analysis and the negative sign of the amenities latent variable in the full model. Table IX shows that SMSAs with more dance, drama, and music events had significantly lower out-migration. This contrasts with the results of both the regression model and the full latent variable model. The beta coefficient indicates that an increase of one standard deviation in the number of drama, music, and dance events per year, 25 events, was associated with a decrease of about one out-migrant per 3,000 population.

A Single Latent Variable Model

The above sections have modelled out-migration as a function of the economic, climatic, demographic, amenity, and disamenity determinants. In addition, migration can be modelled as a function of the single unobservable latent variable, expected quality of life. As discussed in Chapter II, out-migration would be expected to be negatively associated with expected quality of life, being higher where expected quality of life is lower.

The question that must be answered is what indicators can represent quality of life. On the one hand, one could expect the multi-dimensional quality of life to be best represented by a large number of indicators.

On the other hand, a simplier model with just a few indicators summarizing the various aspects of quality of life may be more appropriate.

The first approach was tried but did not provide an adequate fit. As in the full model, it may be that latent variables are best represented by two or three indicators only. Using the second approach did provide a model with an adequate fit, but only when the economic and city indicators were excluded. The only simple model that fit the data included as indicators age, education, and cold. This model is shown in Figure 5 and the results are presented in Table X.

The single latent variable model performs very well. The chi square statistic is low and the associated probability of fit of 0.5314 is higher than those of the previous models. The error term of 0.5373 implies an unadjusted R^2 of 0.4627. The latent variable is indicated by mobility propensities plus cold. Taking account of mobility propensities is important. However, after doing so, only cold entered as an indicator. A superior method of taking account of mobility propensities may be to disaggregate the sample by age and education. Doing so may then allow more indicators to enter the single latent variable.

Summary and Conclusions

In this chapter the latent variable approach was used to estimate the relationship between out-migration and the economic, climatic, demographic, amenity, and disamenity categories of equation (4) in Chapter II. Two models were estimated: the full latent variable

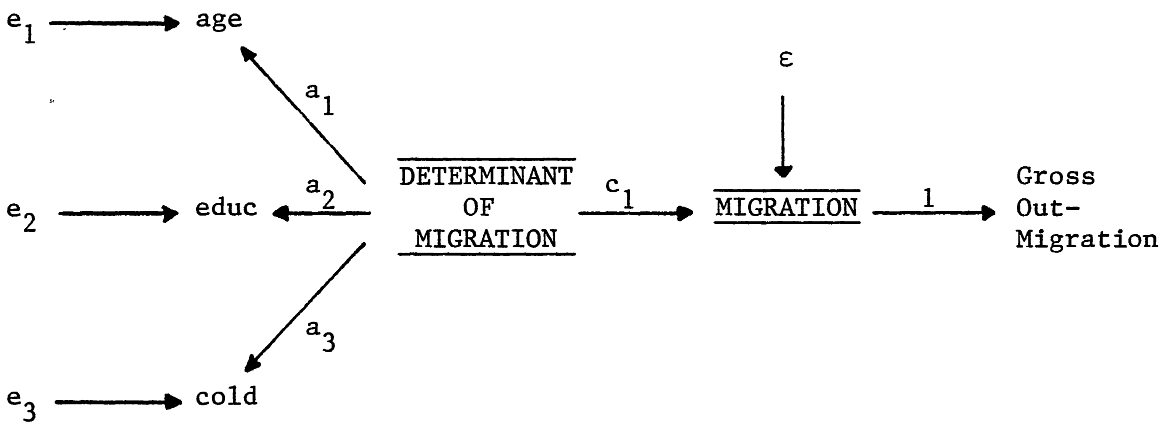

Figure 5. A Single Latent Variable Model

TABLE X

SINGLE LATENT VARIABLE MODEL

	Beta Coefficient	Standard Error	t-Statistic
	0.6802	0.1438	5.3285
Error Term	0.5373	0.1098	4.8935

Chi Square Statistic = 1.2647
Probability of Fit = 0.5314
D.F. = 2
N = 77

Indicator	Beta Coefficient	Standard Error	t-Statistic	Error Term
Single Latent Variable:				
age	0.8877			0.2118
educ	0.7139	0.1464	5.4925	0.4904
cold	-0.2622	0.1398	-2.1127	0.9312

model analogous to the full specification out-migration model of Chapter III, which did not provide a good fit; and a simplier model with only two multiple indicator variables, which did provide a good fit.

The latent variable approach was used because of its ability to estimate both the relationship between the general determinants of migration, the economic factor as an example, and migration, and the relationship between the determinants and their indicators taking into account that the indicators, age and education as an example, do not measure the multi-dimensional determinants perfectly, say the demographic factor. However, the full latent variable model did not provide an adequate fit. The goal of estimating the relationship between the economic, climatic, demographic, amenity, and disamenity factors and out-migration, holding city size constant, could not be met because the factors were not amenable to being modelled as multiple indicator latent variables.

A simplified model was estimated in which only the climatic and demographic latent variables were modelled as multiple indicator variables, measured by two indicators each. The economic side was represented by income growth and the amenities side by dance, drama, and music events. Disamenities measures and the logarithm of city size were not included since an adequate fit could not be obtained when they were included.

The simplified model worked well and upheld the findings of the regression analysis that higher income growth was significantly associated with lower out-migration and that mobility propensities were important determinants of out-migration. However, the finding that climate was not an important determinant of out-migration ran counter to the findings of the regression analysis when cold and temperature variance were entered as separate regressors. The model found that more dance, drama, and music events were significantly associated with less out-migration. This was also contrary to the regression result.

Attempts were made to estimate a single latent variable model. However, satisfactory results were not obtained with either a large or a small model.

Overall, the goals of estimating the relationship between out-migration and its general determinants and between out-migration and a single variable, quality of life, were not met.

Comparing the latent variable results of the chapter with the regression results of Chapter III, the conclusion is that regression analysis was more informative and thus superior. The potential of the latent variable model for migration analysis is large, however, and deserves further study.

CHAPTER VI

INDEX NUMBER RESULTS

Introduction

In thie chapter the index number approach is explored as an alternative to latent variable modelling. First, index numbers are constructed analogous to the latent variables used in the previous chapter. The model is then estimated. The results are compared and contrasted with those of the latent variable models. A summary and conclusions end the chapter.

Construction of Index Numbers

The index numbers were constructed according to the following formula (USDA, 1979):

$$I_{ij} = \sum_{k} C_{ik} \tag{1}$$

where I_{ij} = index number j for SMSA i,

k = indicators included in index j,

$C_{ik} = (X_{ik}/X_{MAX\ k})100$, the value for indicator k in SMSA i in percentage terms,

X_{ik} = the unadjusted values for indicator k in SMSA i, and

$X_{MAX\ k}$ = the maximum unadjusted value for indicator k.

The unadjusted data for each indicator were transformed into percentage terms, as a percentage of the maximum value for the indicator, so that each indicator would have an equal weight in the index. This accords with the USDA data (1979) and the Liu data (1975). It also facilitates comparison with other data in the future. The index numbers were formed by summing the adjusted indicator values. The composition of the index numbers is identical to that of the latent variables of the full model. The index numbers can then be used, along with the logarithm of city size, to express in index number from the full latent variable model. Thus, index numbers are an alternative to latent variables as an attempt to incorporate some of the multidimensionality of the economic, climatic, demographic, amenity, and disamenity determinants of out-migration in a summary measure.

The Full Index Number Model

Figure 6 shows the specification of the full index number model. It is analogous to the full latent variable model. However, it does not show as many relationships and does not give as much information as the latent variable models. Indicator coefficients and indicator error terms are not given.

Table XI presents the results of estimating the model for out-migration. The regression explains 65 percent of the variation in out-migration. This is a much better fit than that of the full latent variable model which had a probability of fit of zero and an R^2 of 0.2384. The F-statistic shows that the equation is significant at the one percent level.

The hypotheses tested in the equation are that better economic conditions are associated with less out-migration, that better climate is associated with less out-migration, that higher percentages of

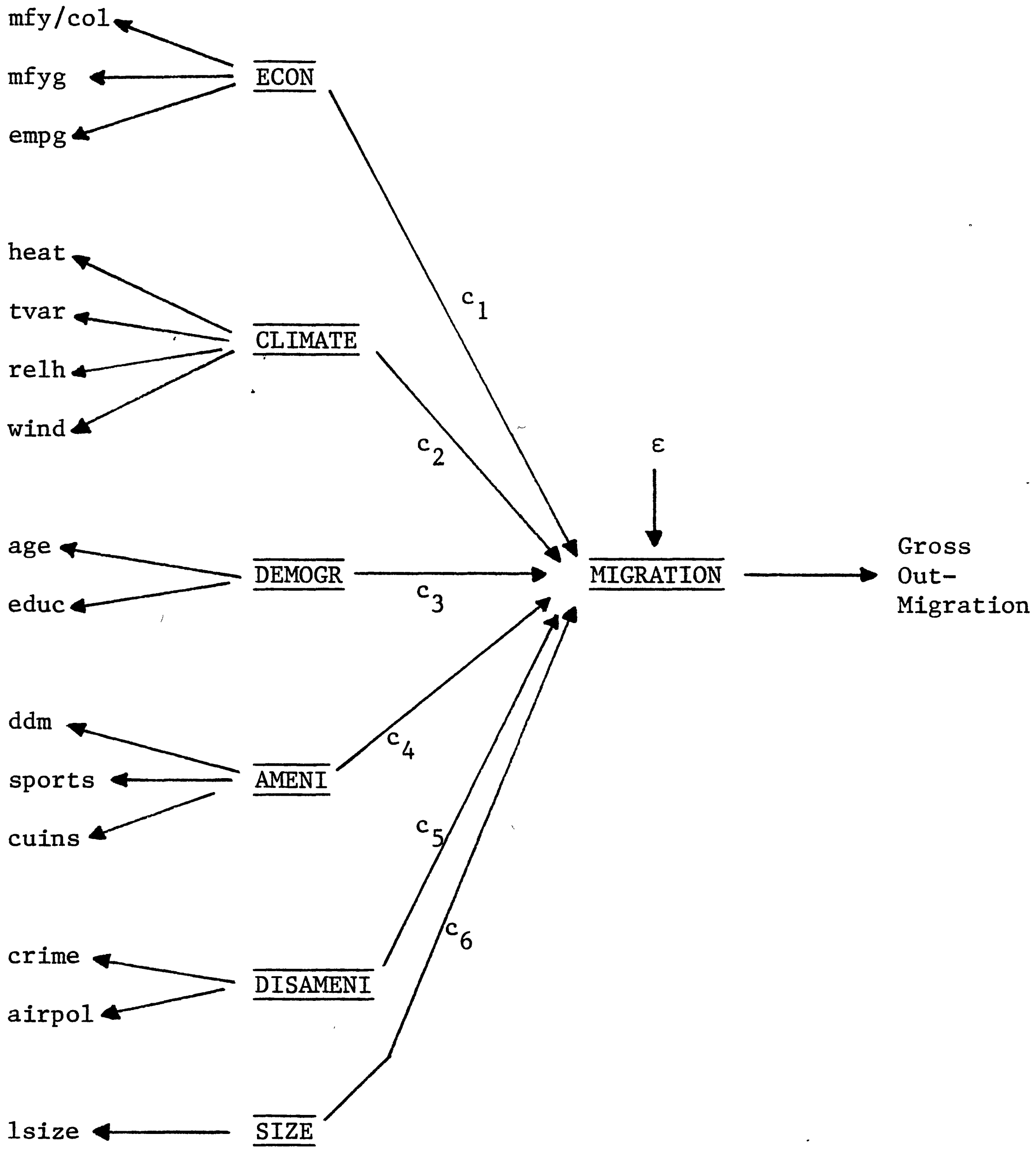

Figure 6. The Full Index Number Model

TABLE XI

FULL INDEX NUMBER MODEL

	Out-Migration			
Variable	Coefficient	Standard Error	t-Statistic	Probability
CONSTANT	1.945091	0.533873	3.64	0.0005
ECON	-0.005030	0.001607	-3.13	0.0026
CLIMATE	-0.000943	0.000869	-1.08	0.2818
DEMOGR	0.014213	0.001774	8.01	0.0001
AMENI	-0.000187	0.000852	-0.22	0.8270
DISAMENI	0.005050	0.001984	2.55	0.0131
LSIZE	-0.000026	0.000005	-4.82	0.0001

$R^2 = 0.6535$

$S^2 = 0.0949$

N = 77

F = 22.01

F-probability = 0.0001

residents in the more mobile age and education groups promote out-migration, that cities with more amenities have less out-migration, and that cities with more disamenities have more out-migration. The relationship between city size and out-migration is indeterminate beforehand.

Economic conditions are negatively associated with out-migration, as expected. That is, SMSAs with better economic conditions had less out-migration. The relationship was significant at the one percent level. This result is quite different from that of the full latent variable model in which the economic latent variable, also measured by real income, income growth, and employment growth, had a positive sign and was insignificant. There is agreement with the simpler latent variable model in which income had a significant negative association with out-migration. The multiple regression results in Table VI also show that income growth was significant with a negative sign, but the other two indicators were not significant. Overall, the importance of economic considerations in the out-migration decision, taking account of quality of life factors, mobility propensities, and city size, is upheld. This is in contrast with the finding of Porell (1982), the only other study holding the other factors constant, that the economic side did not affect out-migration. Porell used index numbers for quality of life factors.

Climate is insignificant and has an unexpected negative sign. Climate was insignificant in the first two latent variable models with a negative sign in the simpler model. On the other hand, multiple regression analysis showed that cold, temperature variance, and wind were significantly related to out-migration while only relative humidity was not. Graves (1980) also found cold and temperature variance significant, but Porell (1982) did not find two climate indexes significant for out-migration. That is, when the climate indicators are entered as separate regressors, the evidence indicates that climate is an important determinant of out-migration. But when they are grouped together, climate does not seem to be a significant factor.

The demographic variable is highly significant and positively associated with out-migration, as expected. Indeed, it was significant throughout the study with the exception of the full latent variable model. However, the full latent variable model did not have any significant variables. This points out the importance of holding constant the mobility propensities of the population when evaluating the role of the economic and other determinants of out-migration, something that is not always done.

Although SMSAs with relatively more amenities did have less out-migration, as expected, the relationship is insignificant. This agrees with the full latent variable model. The simplier latent variable model used dance, drama, and music events as the amenities measure and also found a negative relationship, as did multiple regression analysis. However, the simpler model was the only one to find evidence that the amenities side was significantly related to out-migration. It must be remembered, that this model did not include city size since its inclusion did not result in an adequate fit for the model. An index number version of the simple latent variable model gave the same results as the latent variable model, but of course, did not provide information on the relationship between the demographic and climate variables and their indicators. Overall, city amenities by themselves, holding other factors constant, especially city size, do not seem to play an important role in determining out-migration. This result agrees with that of Porell (1982) who found that an index of amenity indicators was not significantly related to out-migration.

Disamenities are positively related to out-migration, as expected, and as they have been in all models. However, unlike the other models, the full index number model shows disamenities to be significantly related to out-migration. The only other study that included disamenities, Porell (1982), found a crime index and air pollution index to be insignificant in explaining out-migration. Since the full index number model does include city size, the finding that disamenities play an important role in determining out-migration is strengthened. In general, the results imply a more important role for economic and quality of life factors than those of Porell (1982), since he did not find these factors significant determinants of out-migration.

City size has a negative and significant coefficient in the full index number model. This is in contrast to the full latent variable model but in agreement with the multiple regression results. Support is found for Miller's (1973) argument that areas with larger job markets have less out-migration. As with mobility propensities of the population, city size should be included in the model when evaluating the role of the determinants of out-migration.

Summary and Conclusions

This chapter has explored the index number approach as an alternative to latent variable modelling. An index number version of the full latent variable model of the previous chapter was estimated and the results compared and contrasted with those of the latent variable models and of the regression analysis of Chapter III.

The results showed that index number models can provide a good fit when all the variables are multiple indicator variables. The full latent variable model, using the same data, did not provide an adequate fit. However, had the full latent variable model been able to provide a good fit, it would have yielded more information than the index number model. The index number model appears to be superior in regard to large models that contain several multiple indicator variables, say three or four indicators per variable. For simpler models, though, for which latent variable modelling can provide a good fit, the latent variable technique supplies the user with more information than index number modelling. Latent variable models show how well the indicators measure the latent variable and also show the error term associated with each indicator.

CHAPTER VII

SUMMARY AND CONCLUSIONS

Introduction

This chapter summarizes the results of the study. It draws conclusions concerning the determinants of migration. Conclusions are also reached concerning the roles and uses of the regression, latent variable, and index number approaches, and concerning the implications of the results of the study for policy action. Recommendations for future research end the chapter.

Determinants of Migration

The multiple regression analysis of Chapter III showed that in-migration had a significant positive association with both real income and employment growth, but was more responsive to higher levels of real income than to higher rates of growth in employment. Growth in income, however, did not significantly affect in-migration. The conclusion is that economic considerations are an important determinant of in-migration.

The climate measures were negatively related to in-migration. That is, SMSAs with better climate had more in-migration, but only heating degree days was significant of the four climate measures. Temperature variance, relative humidity, and wind speed did not play an important role in determining in-migration. The conclusion is that, other things being equal, in-migration was responsive to cold and warmth.

The educational attainment of the resident population in an SMSA served as a proxy for the educational requirements of the industries in the SMSA. Education had a highly significant and positive relationship with in-migration. The computed elasticity of in-migration with respect to education was relatively high. The conclusion is that a higher percentage of industries in an SMSA requiring better educated workers promotes more migration into the SMSA.

City amenities were measured by dance, drama, and music events; sports events; and cultural institutions. None was significant in the multiple regression analysis. The computed elasticities were close to zero. The conclusion is that city amenities did not influence in-migration appreciably during 1965 to 1970.

City disamenities, measured by crime and air pollution, were also not significant. As with city amenities, the conclusion is that city disamenities were not an important determinant of in-migration.

The logarithm of population showed a highly significant and negative relationship with in-migration. It had the highest computed elasticity among the independent variables. In-migration was relatively very responsive to differences in city size. The results agree with both the argument that smaller cities are preferred and with the argument that smaller job markets induce more in-migration. The conclusion is that city size should be controlled for when specifying the in-migration equation.

The multiple regression results showed that income growth had a significant negative relationship with out-migration, but that real income growth was upheld by the simpler latent variable model. The results of the index number model showed that the economic variable had a highly significant and negative association with out-migration. The conclusion is that economic considerations are important influences in the out-migration decision.

With regard to climate, on the one hand, when the climate measures are entered as independent variables in multiple regression, the results indicate that climate is an important determinant of out-migration. On the other hand, when the climate indicators are grouped together, climate does not significantly affect out-migration.

Age and education had a highly significant positive association with out-migration in multiple regression analysis. Age had the largest computed elasticity of the independent variables. The same significant association was found for the demographic variable in both the simple latent variable model and in the index number model. The conclusion is that the mobility propensities of the resident population should be held constant when evaluating the importance of the determinants of migration.

City amenities were not significantly related to out-migration according to the multiple regression results. This was also found in the index number model. On the other hand, the simple latent variable model, which did not include city size, found dance, drama, and music events to have a significant negative relationship with out-migration. The results are somewhat mixed with regard to city amenities. Overall, they do not appear to have been very important in affecting out-migration.

City disamenities showed a significant association with out-migration only in the index number model. Those results showed that higher levels of disamenities were related to higher levels of out-migration.

City size had a negative and significant coefficient in both the multiple regression analysis of Chapter III and in the index number model of Chapter VI. Support was found for the argument that larger job markets tend to have, other things being equal, less out-migration. The conclusion is that, as with the demographic variable, city size should be controlled for when analyzing the determinants of migration.

In general, the results of the study disagree with those of Graves (1980) and agree with those of Porell (1982) in concluding that both economic and quality of life factors affect migration. The results of the present study provide stronger support, however, for the importance of economic considerations in the migration decision than do Porell's since they show a significant relationship between growth in income and out-migration whereas Porell did not find the economic side to be an important determinant of out-migration. In addition, the results go beyond Graves and Porell to indicate that, in predicting migration, the educational attainments of the resident population are very important and that the size of the city should be controlled for in the migration model.

For policy makers, the questions posed at the introduction to the study have been answered in the affirmative. Regional economic development policies are rendered less effective because government has little influence over some determinants of migration such as climate. However, the results imply a more effective role for economic development policies than do the results of Graves' (1980) study by showing, like Porell (1982), that better economic conditions are associated with more in-migration and less out-migration. Also, for predicting migration and evaluating the possible impact of regional policies on migration, the life cycle and education aspects, which are highly related to migration, must be taken into account.

Regression, Latent Variable, and Index Number Models

Each model has its own uses. For simple, straightforward systems with one indicator per variable, the multiple regression model is the best choice. However, when the model, even though small, has more than one indicator per variable, the latent variable model can be useful. By incorporating the multi-dimensional nature of the variable into the model, the researcher has a more flexible model to work with.

The multiple regression model is the tool to use when investigating the relationship between a specific variable and migration. Direct comparisons can be made between the effects of different variables on migration. The latent variable approach can provide additional insight by being able to model migration at a more general level as a function of overall determinants. For example, the relationship between migration and mobility propensities, as indicated by age and education, can be estimated. The latent variable model can also add to the knowledge gained in multiple regression by incorporating measurement error. Thus, an evaluation can be made as to how well, relatively, each indicator measures the latent variable. For example, latent variable analysis showed that age is a better measure of mobility propensities than education.

Two or more highly correlated variables can cause the problem of multicollinearity in the multiple regression model. Their separate influences on the dependent variable cannot be disentangled and both appear as insignificant. Also, the researcher cannot evaluate the relationship between the general determinants that they represent and the dependent variable. The latent variable model can be useful when multicollinearity is present because the variables can be specified as indicators of a more general variable. The relationship between this latent variable and the dependent variable can then be estimated. Also, the indicators can be compared to see how well they measure the latent variable. The other variables without multicollinearity problems enter the latent variable model or single indicator variables, just as they enter the multiple regression model.

Indeed, the multiple regression model is a special case of the latent variable model since it is simply the case where all the variables are single indicator variables. The results of estimating a multiple regression equation with the latent variable model or with ordinary least squares are identical, as the appendix shows. The latent variable approach can also be used to model a simultaneous equations system. In Chapter V it was pointed out that there can be more than one dependent variable. The B matrix shows the relationships between the dependent variables. Such a flexible model merits application to the study of migration to see if it can allow additional insight into migration.

In spite of the potential of the latent variable model, the results of this study show regression analysis to be superior at the present to latent variable analysis. For models with many multiple indicator variables, the index number approach seems more useful than the latent variable approach. The appropriate specification of a gross migration latent variable model needs further research. It may be that a fruitful area for latent variable modelling of migration is a simultaneous equations study of migration. The conclusion is that latent variable modelling holds promise for migration studies and so needs further research.

Recommendations

The following are this researcher's recommendations:

1. There should be more research concerning the applicability of latent variable models to the study of migration. This research should include investigation into the process of indicator selection for the latent variables and the process of model building.

2. Research is needed at the level of the individual. Studies concerning the determinants of individual migration moves would be a direct approach to the questions of what the determinants of migration are and what the relative important of each is.

3. Related to the second recommendation is the recommendation that, in order to study the relationship between quality of life and migration more directly, research should be conducted into the generation of quality of life data for individual migrants. Thus, instead of studying the relationship between inputs into quality of life and migration, the relationship between the output, quality of

life, and migration could be studied. Ideally, this is the relationship that should be studied. It should be studied at the level of the individual. More research should be directed, then, at the methodology for obtaining comparative individual data for migrants which allows a comparison of levels of quality of life.

4. Valuable insight into gross migration and its determinants can be gained by the use of simultaneous equation models of analysis.

A SELECTED BIBLIOGRAPHY

Aaker, D. A. and Bagozzi, R. P. "Unobservable Variables in Structural Equation Models with an Application in Industrial Selling." Journal of Marketing Research 16 (May, 1979), 147-58.

Alonso, W. "Policy Implications of Intermetropolitan Migration Flows." Working Paper No. 177. Berkeley: Institute of Urban and Regional Development, University of California, Berkeley, 1972.

Alperovich, G., Bergsman, J. and Ehemann, C. "An Econometric Model of Migration between U.S. Metropolitan Areas." Working Paper. Washington: The Urban Institute, 1975.

Amos, O. "Preliminary Construct of Quality of Life Theory: A Needs Satisfaction Framework." Stillwater, Oklahoma, 1980. (Mimeographed.)

Bagozzi, R. P. Causal Models in Marketing. New York: John Wiley and Sons, 1980.

__________. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error: A Comment." Journal of Marketing Research 18 (August, 1981), 375-81.

Bentler, P. M. "Linear Systems of Multiple Levels and Types of Latent Variables." In Jöreskog, K. G. and H. Wold (eds.), Systems Under Indirect Observation: Causality, Structure, Prediction. Cartigny, Switzerland: Conference on Systems Under Indirect Observation, 1981.

Cebula, R. J. The Determinants of Human Migration. Lexington, Mass.: Lexington Books, 1979.

Cebula, R. J. and Davis, M. N. "A Note on Determinants of Inter-Metropolitan Migration." The Review of Regional Studies 8 (Fall, 1978), 48-51.

__________. "Migration, Economic Opportunity, and the Quality of Life: Reply and Extension." Journal of Regional Science, Vol. 16, No. 1 (1976), 112-115.

Cebula, R. J. and Vedder, R. K. "A Note on Migration, Economic Opportunity, and the Quality of Life." Journal of Regional Science 14 (August, 1973), 205-11.

Darden, W. R. "Review of Behavioral Modelling in Marketing." Review of Marketing (1981), 21-30.

Eldridge, H. T. "Primary, Secondary, and Return Migration in the United States, 1955-1960." Demography 2 (1965), 444-55.

Fields, G. S. "Labor Force Migration, Unemployment, and Job Turnover." The Review of Economics and Statistics 58 (November, 1976), 407-15.

__________. "Place-to-Place Migration: Some New Evidence." The Review of Economics and Statistics 1 (1979), 21-32.

Geraci, V. J. "Simultaneous Equation Models with Measurement Error." (Unpublished Ph.D. dissertation, University of Wisconsin, 1974.)

Graves, P. "A Reexamination of Migration, Economic Opportunity and the Quality of Life." Journal of Regional Science, Vol. 16, No. 1 (1976), 107-12.

__________. "Income and Migration Revisited." The Journal of Human Resources, Vol. 14, No. 1 (1979a), 112-19.

__________. "A Life-Cycle Empirical Analysis of Migration and Climate, by Race." Journal of Urban Economics 6 (April, 1979b), 135-47.

__________. "Migration and Climate." Journal of Regional Science, Vol. 20, No. 2 (1980), 227-37.

Graves, P. and Linneman, P. "Household Migration: Theoretical and Empirical Results." Journal of Urban Economics 6 (July, 1979), 383-404.

Greenwood, M. "Research on Internal Migration in the United States: A Survey." Journal of Economic Literature 13 (June, 1975), 397-433.

__________. Migration and Economic Growth in the United States. New York: Academic Press, Inc., 1981.

Hall, O. P., Jr. and Licari, J. A. "Forecasting Migration in a Regional General Equilibrium Context." In Brown, A. A. and E. Newberger (eds.), Internal Migration: A Comparative Perspective. New York: Academic Press, 1977.

Hsiao, C. "Identification for a Linear Dynamic Simultaneous Error-Shock Model." Working Paper. Berkeley: Department of Economics, University of California, 1974.

__________. "Identification and Estimation of Simultaneous Equation Models with Measurement Error." International Economic Review (1976).

Jöreskog, K. G. and Goldberger, A. S. "Estimation of Model with Multiple Indicators and Multiple Causes of a Single Latent Variable." Journal of American Statistical Association 70 (September, 1975), 631-39.

Jöreskog, K. G. and Sorbom, D. "Statistical Methods for Analysis of Longitudinal Data." In Aigner, D. J. and A. S. Goldberger (eds.), Latent Variables in Socioeconomic Models. Amsterdam: North Holland, 1977.

Jöreskog, K. G. and van Thillo, M. "LISREL: A General Computer Program for Estimating a Linear Structural Equation System Involving Multiple Indicators of Unmeasured Variables." Educational Testing Service, Princeton, N.J.: Research Bulletin #72-56, 1973.

Kau, J. B. and Sirmans, C. F. "Migration and the Quality of Life." Review of Regional Studies 6 (Winter, 1976), 76-85.

Lancaster, K. "A New Approach to Consumer Theory." Journal of Political Economy 74 (1966), 132-57.

Lawley, D. N. and Maxwell, A. E. Factor Analysis as a Statistical Method, 2nd edition. London: Butterworth, 1971.

Lesourne, J. "A Theory of the Individual for Economic Analysis." In Jorgenson, D. W. and J. Waelbroek (eds.), Vol. 114 of Contributions to Economic Analysis. New York: North-Holland, 1977.

Liu, B. "Variations in the Quality of Life in the United States, 1970." Review of Social Economy 32 (October, 1974), 131-47.

__________. "Quality of Life Indicators in U.S. Metropolitan Areas, 1970: A Comprehensive Assessment." Washington Environmental Research Center, U.S. Enviromental Protection Agency, Grant #R803049-01-0, May, 1975.

__________. "Differential Net Migration Rates and the Quality of Life." Review of Economics and Statistics 57 (August, 1975a), 329-37.

Long, L. and Boertlein, C. G. "The Geographical Mobility of Americans: An International Comparison." Current Population Reports, Bureau of Census, Special Studies, Series P-23, #64.

Long, L. and Hansen, K. "Reasons for Interstate Migration." Current Population Reports, Bureau of Census, Series P-23, #51, March, 1979.

Lowry, I. S. Migration and Metropolitan Growth: Two Analytical Models. San Francisco: Chandler, 1965.

Miller, A. R. "The Migration of Employed Persons To and From Metropolitan Areas of the United States." Journal of the American Statistical Association 62 (December, 1967), 1418-32.

Miller, E. "Is Out-Migration Affected by Economic Conditions?" *Southern Economic Journal* (January, 1973a), 396-405.

__________. "Return and Nonreturn In-Migration." *Growth and Change* (January, 1973), 3-9.

Pindyck, R. and Rubinfeld, D. *Econometric Models and Economic Forecasts*, 2nd edition. New York: McGraw-Hill, 1981.

Porell, F. W. "Intermetropolitan Migration and Quality of Life." *Journal of Regional Science*, Vol. 22, No. 2 (1982), 137-50.

Ravenstein, E. G. "The Laws of Migration." *Journal of the Royal Statistical Society*, 48 (June, 1885), 167-227.

Ritchey, P. N. "Explanations of Migration." *Annual Review of Sociology* 2 (1976), 363-404.

Saben, S. "Geographic Mobility and Employment Status, March, 1962-March, 1963." *Monthly Labor Review* 87 (August, 1964), 873-81.

Salvatore, D. "A Simultaneous Equations Model of Internal Migration with Dynamic Policy Simulations and Forecasting." *Journal of Development Economics* 7 (1980), 231-46.

Schuessler, R. "Migration--A Multi-Regional Approach: A Critique of Net Migration Studies." *Discussion Paper #76*. Cambridge: Program on Regional and Urban Economics, Harvard University, 1972.

Schwartz, A. "Interpreting the Effect of Distance on Migration." *Journal of Political Economy* 81 (September/October, 1973), 1153-69.

Sjaastad, L. A. "The Cost and Returns of Human Migration." *Journal of Political Economy* 70 (October, 1962), 80-93.

Theil, H. *Principles of Econometrics*. New York: John Wiley and Sons, 1971.

U.S. Department of Agriculture. Economic Development Division. Economics, Statistics, Cooperatives Service. "Indicators of Social Well-Being for U.S. Countries." Rural Development Research Report No. 10, 1979.

U.S. Department of Commerce. Bureau of Census. "Estimates of the Population Standard Metropolitan Statistical Areas: July 1, 1965." *Current Population Reports*, Series P-25, #371, 1967.

__________. *County and City Data Book*. Washington, D.C.: Government Printing Office, 1963, 1973.

__________. "Gross Migration by County: 1965 to 1970." *Current Population Reports*, Series P-25, #701, 1977.

__________. "Estimates of the Population of Counties and Metropolitan Areas: July 1, 1974 and 1975." Current Population Reports, Series P-25, #709, 1977.

__________. United States Census of Population: 1970. Migration Between State Economic Areas.

__________. National Climatic Center. Climatological Data. National Summary, Annual Summary 31 (13), 1980.

U.S. Department of Health, Education, and Welfare. Air Quality Data, 1966 edition. National Air Pollution Control Administration, Publication No. APTD, 68-9, 1968.

Wright, S. "Path Coefficients and Path Regressions: Alternatives or Complementary Concepts?" Biometrics 16 (1960), 189-202.

Zapf, W. "Systems of Social Indicators: Current Approaches and Problems." International Social Science Journal 27 (1975), 479-98.

Zuiches, J. and Fuguitt, G. "Residential Preferences: Implications for Population Redistribution in Non-Metropolitan Areas." U.S. Commission on Population Growth and the American Future, Population, Distribution, and Policy, Washington, D.C., 1972.

APPENDIX

MULTIPLE REGRESSION AS A SPECIAL CASE OF THE LATENT VARIABLE MODEL

This appendix shows that the latent variable model gives the same results in estimating the single equation multiple regression model as ordinary least squares and that the standardized error term, the one that was reported in the study, is equal to $1-R^2$ where R^2 is the unadjusted coefficient of determination of ordinary least squares estimation. To summarize the comparison of the results, the information concerning the coefficients and standard errors will be implicitly compared in the explicit comparison of the t-statistics from the two estimation techniques for out-migration.

TABLE XII

COMPARISON OF LATENT VARIABLE ERROR TERMS WITH REGRESSION R^2s

Variable	Ordinary Least Squares t-Statistic	Latent Variable t-Statistic
MFY/COL	0.27	0.27
MFYG	-4.98	-4.98
EMPG	-0.29	-0.29
HEAT	-3.44	-3.44
TVAR	2.33	2.33
WIND	2.08	2.08
AGE	3.77	3.77
EDUC	3.05	3.05
DDM	-1.05	-1.05
LSIZE	-3.17	-3.17
AIRPOL	-0.01	-0.01
	$R^2 = 0.75645$	Error Term = 0.24355

VITA 2

Rafael Escutia Villagomez

Candidate for the Degree of

Doctor of Philosophy

Thesis: A LATENT VARIABLE APPROACH TO MIGRATION

Major Field: Economics

Biographical:

Personal Data: Born in Cuitzeo, Mexico, August 5, 1948, the son of Raúl Villagomez and Consuelo Escutia.

Education: Graduated from St. Ignatius High School, Chicago, Illinois, in 1967; received the Bachelor of Arts degree in Economics from the University of Illinois, Chicago, Illinois, 1971; received the Master of Arts degree from DePaul University, Chicago, Illinois, 1973; completed requirements for the Doctor of Philosophy degree at Oklahoma State University, May, 1985.

Professional Experience: Instructor, Drake University, Des Moines, Iowa, 1982-1983; Instructor, University of Puerto Rico, San Juan, Puerto Rico, 1983 to present.