MICROECONOMIC APPROACHES TO STUDYING MIGRATION DECISIONS

Julie DaVanzo

May 1980

N-1201-NICHD

The National Institute of Child Health and Human Development, National Institutes of Health
This project was supported by Grant No. 5 R01-HD10864-02 awarded by the National Institute of Child Health and Human Development, DHEW.

The Rand Publications Series: The Report is the principal publication documenting and transmitting Rand's major research findings and final research results. The Rand Note reports other outputs of sponsored research for general distribution. Publications of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

Published by The Rand Corporation
MICROECONOMIC APPROACHES TO STUDYING MIGRATION DECISIONS

Julie DaVanzo

May 1980

N-1201-NICHD

The National Institute of Child Health and Human Development, National Institutes of Health
PREFACE

This is a revised version of a paper prepared for the workshop on Micro-Level Approaches to Migration Decisions: Theory and Methodology in the Tenth Summer Seminar in Population at the East-West Population Institute, East-West Center, Honolulu, Hawaii, June-July 1979. It will be published as a chapter in Gordon F. De Jong and Robert W. Gardner, Migration Decisionmaking, Pergamon Press, forthcoming.

This note should be useful to people wishing to learn more about microeconomic theoretical and empirical models of migration, and to people designing surveys to provide data for analyses of migration.

Preparation was supported by grant no. HD10864-02 from the National Institute of Child Health and Human Development, U.S. Department of Health, Education, and Welfare.
SUMMARY

This Note introduces the reader to the main ideas and concepts of microeconomic models for studying migration decisionmaking, summarizes and critiques the main empirical approaches used to test these models, summarizes their principal findings, and highlights the resulting implications for the design of surveys to provide data suitable for estimating microeconomic models of migration.

The basic human capital model is introduced and is extended to incorporate information costs, uncertainty and risk, and (consumption and financing) income effects, and to apply to family migration and return and repeat migration. The Note discusses issues concerning measurement of the migration variable and of the explanatory variables that influence migration decisions.

The models and methods discussed pertain to migration in both developing and developed countries. Modifications are noted that might be necessary to adapt, to a developing country setting, models that are usually used to study migration in developed countries.
ACKNOWLEDGMENTS

The author would like to thank John Antel, William Butz, Gordon De Jong, Sidney Goldstein, Will Harriss, James Hosek, Arleen Leibowitz, and Peter Morrison and participants at the Summer Seminar workshop for their helpful comments on an earlier draft.
CONTENTS

PREFACE ................................................................. iii
SUMMARY ................................................................. v
ACKNOWLEDGMENTS ..................................................... vii

Section
I. INTRODUCTION ...................................................... 1

II. THE MICROECONOMIC PERSPECTIVE ON MIGRATION
DECISIONMAKING ....................................................... 3
The Basic Microeconomic Model ..................................... 3
Variants of the Basic Model .......................................... 7
Family Considerations ................................................ 7
Return and Repeat Migration ........................................... 7
Information Costs ....................................................... 7
Uncertainty and Risk .................................................. 8
Income Effects .......................................................... 9
Comparison of Microeconomic Models with
Other Micro Models of Migration .................................... 10

III. EMPIRICAL IMPLEMENTATION OF THE BASIC
MICROECONOMIC MODEL OF MIGRATION ......................... 12
The Dependent Variable: Migration ................................... 12
Explanatory Variables ............................................... 17
The Conceptual Variables are Utilities ............................. 18
The Conceptual Variable Refers to the
Lifetime Stream of Real Income ................................... 20
We Cannot Observe What an Individual
Would Earn in Other Areas .......................................... 21
The Conceptual Variables are Expected Utilities ............... 26
Ability to Finance an Investment in Migration
Affects the Likelihood of Moving .................................. 29
Distance is a Commonly Used Proxy for
Many Costs of Moving .............................................. 32

IV. FAMILY CONSIDERATIONS ........................................ 34
Theoretical Approaches ............................................. 34
Empirical Approaches ............................................... 36

V. RETURN AND REPEAT MIGRATION: AN ILLUSTRATION OF
LINKING HUMAN CAPITAL ORIENTATION TO NEW
ISSUES OF MIGRATION ............................................... 39
Theoretical Considerations .......................................... 39
Empirical Findings .................................................... 42
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>VI. RECOMMENDATIONS FOR SURVEY DESIGNERS</td>
<td>46</td>
</tr>
<tr>
<td>Survey Design</td>
<td>46</td>
</tr>
<tr>
<td>To Measure Migration</td>
<td>46</td>
</tr>
<tr>
<td>To Delimit Working Samples</td>
<td>47</td>
</tr>
<tr>
<td>To Construct Variables to Explain Who Moves and Why</td>
<td>47</td>
</tr>
<tr>
<td>To Explain Destination Choice</td>
<td>49</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>51</td>
</tr>
</tbody>
</table>
I. INTRODUCTION

Migration is one of the important ways by which people can improve their lives. People migrate for many reasons: to escape distressed areas, to find better jobs, even to satisfy their wanderlust. Rural people are attracted by the bright lights of the city; city people move to the countryside to get away from it all. Retired people move to places that offer a pleasant climate or abundant recreational opportunities. Migration is equally important to government planners and policymakers, because the aggregate movements of migrants often cause important shifts in the political power and economic fortunes of the places they leave and the places they go to. Migration is also a force for equilibrium in the job market. Migrants' motivations and decision-making mechanisms therefore constitute an important subject for study. Since the discipline of economics deals with resource allocation, it is not surprising that economists have long been interested in migration.

Most early economic analyses of migration relied on aggregate data referring to areas and broad segments of the population, and dealt with migration as an equilibrating mechanism that acted to reduce geographic wage differentials. More recently, the center of attention has shifted toward a micro-oriented approach: Many economists who study migration now try to understand why individuals (or families) choose to migrate.

Part of the impetus to this new direction has been theoretical, but a significant facilitating factor has been the growing availability of microdata in the last decade. These rich information sources have allowed rigorous tests of microeconomic models of migration and have motivated many refinements in them. For example, economic-model builders have broadened their scope to embrace topics often considered to be outside the economist's domain, such as family considerations.

Persons who wish to appreciate the direction in which research on migration is evolving, and what its theoretical and empirical foci are likely to be, should understand what this recent but rapidly developing
approach is about. Accordingly, this paper introduces the reader to
the main ideas and concepts of microeconomic models for studying migra-
tion decisionmaking; summarizes and critiques the main empirical ap-
proaches used to test these models; summarizes their principal findings;
and highlights the resulting implications for the design of surveys to
provide data suitable for estimating microeconomic models of migration.

The paper does not pretend to be exhaustive or to review all micro-
economic studies and approaches. Furthermore, it is restricted mainly
to studies that use microdata, on individuals and households, rather
than aggregate data on areas and groups of people. The focus will be
on the determinants of migration rather than its consequences; however,
since the consequences that people anticipate are a key determinant
of whether and where they move, some of the methodological issues con-
sidered here are common to both topics.*

The models and methods discussed here pertain to migration in
developing and developed countries alike. I have given somewhat more
attention to U.S. migration, since relatively more studies on this
topic use microdata and also because I am more familiar with them. I
have pointed out modifications that might be necessary to adapt the
models to a developing country.

Section II describes the main tenets of the microeconomic approach
to studying migration decisionmaking and contrasts it to other approaches
to studying migration. Section III discusses the empirical approaches
that have been used to test this model and summarizes the main findings.
It also discusses issues concerning measurement of the migration variable
and of the various explanatory variables that influence migration de-
cisions. Sections IV and V deal with two recent extensions of the eco-
nomic model: family considerations and return and repeat migration.
Section VI concludes with a list of the paper's implications for the
design of surveys that seek to provide data for estimating microeconomic
models of migration.

*For a review of critique of the literature on micro-level con-
sequences of migration, see Antel (forthcoming).
II. THE MICROECONOMIC PERSPECTIVE ON MIGRATION DECISIONMAKING

THE BASIC MICROECONOMIC MODEL

The basic premise underlying most microeconomic models of migration decisionmaking is that an individual migrates in the expectation of being better off by doing so. Alternatively stated, the economic model posits that a person chooses to migrate if he or she believes the benefits will exceed the costs. Through appropriate definition of the expected benefits and costs, this general model can apply to any type of migration (e.g., military-related, college-related, economically motivated) and, by extension, to seasonal migration and circulation and other forms of territorial mobility found in certain developing countries (Goldstein, 1978). It applies, then, to a range of circumstances in which it is assumed that people are motivated by a desire to do what is best for themselves. As noted at the end of this section, at this most general level the microeconomic model differs little from other disciplinary approaches to studying migration decisionmaking.

The benefits of migration (and perhaps some of the costs) may accrue over some period of time, which renders migration an investment (an activity that entails costs now in the hope of benefits in the future). Because this investment is embodied in the human being, it is an investment in human capital, i.e., an investment increasing the productivity of human resources (Sjaastad, 1962, p. 83), and models that examine migration from this perspective are referred to as human capital models.

One unique feature of the human capital model compared with other disciplines' approaches to studying migration is explicit recognition of the fact that the benefits of migration occur over a period of time. This helps explain why, for example, migration rates decline with age. The length of the payoff period, and hence the present value of the return to migration (assuming positive differences in all years), is inversely related to age.* Furthermore, considering migration as an

* Other factors may also contribute to the negative relation between migration propensities and age: The quantity of location-specific capital—for example, seniority and friendships—that a person builds up in
investment means that it is not unreasonable for a person to move even if he or she does not expect earnings to increase immediately afterwards (perhaps because it takes time to look for and train for a good job), as long as the person expects that in the longer run he or she will be better off by moving.

A common but mistaken impression is that the economic perspective is restricted to "economic" costs and benefits, narrowly defined, and ignores such important considerations as being near friends and relatives.* It is true that most economic studies of migration do not explicitly consider such factors; this is not because the factors cannot be incorporated into the general model, however, but because they have been difficult to operationalize empirically.

The costs and benefits of migration are by no means restricted to narrowly defined "economic" variables. The costs include not only out-of-the-pocket expenses for transportation and moving of belongings ("direct costs"), but also such factors as (1) earnings forgone while moving and looking for a new job ("opportunity costs"); (2) loss in the value of location-specific assets that are more valuable in the current location than somewhere else, such as clientele; (3) the psychic cost of leaving friends, relatives, and familiar surroundings; (4) expenses entailed in keeping up with friends and relatives left behind (e.g., long-distance telephone calls and visits back); and (5) costs of financing the investment in migration (see the subsection on "Income Effects" below). Similarly, the benefits include not only increased earnings or fringe benefits over one's lifetime, but also increased nonwage income (e.g., higher welfare payments or agricultural subsidies).

---

* An area tends to increase with age; one of the costs of leaving an area is the decrease in the value of these location-specific assets. Young people may be less risk-averse and hence apply a lower discount rate to their future earnings streams. And other costs of moving, for example, those related to number of belongings, may increase with age.

* Nevertheless, when survey respondents are asked why they moved, "economic" reasons are usually offered in the majority of cases (e.g., Long and Hansen, 1979).
or improved amenities (e.g., a more pleasant climate, better cultural facilities, more convenient health clinics, better schools or training programs, or nearness to friends and relatives).* (With regard to publicly provided goods, Tiebout (1956) has hypothesized that people "vote with their feet" and choose a community whose local government best satisfies their personal preferences, and hence move in order to express their preferences for publicly provided goods.) Because factors other than expected changes in earnings affect migration decisions, it is not inconsistent with the economic perspective to observe people moving even if they do not improve their earnings; nor is it inconsistent to observe someone choosing not to move in spite of an offer of a higher salary in another location. Man does not live by bread alone.

In the basic human capital model of migration the potential migrant considers the expected net benefits and costs of moving to each potential destination and chooses to move if there is at least one destination for which expected benefits exceed costs. In this case, he or she will choose the destination offering the greatest expected net gain. Conceptually, this net gain is measured as the difference between the present discounted value of the expected lifetime real income, or utility (well-being), stream the migrant would receive if he or she moved to that destination and the present discounted value of the lifetime real income stream the migrant expects to receive if he or she does not move, less the costs of moving. That is, a person living in area $i$ will move if

$$PV_{ij} = \sum_{t=1}^{T} \frac{U_{i}^{t} - U_{j}^{t} - C_{i}^{t}}{(1 + r)^{t}} > 0 \text{ for at least one area } j(\neq i) \quad (1)$$

and will choose the destination $j$ where $PV_{ij}$ is greatest,**

---

* The distinction between costs and benefits is not clear-cut (or necessary). Negative benefits (e.g., not being near friends and relatives) can be viewed as costs, and vice versa.

** An alternative statement of the model is that a person chooses to move if the internal rate of return on the migration investment is
where $PV_{ij}$ = present value of net gain of moving from $i$ to $j$,

$U_k^t$ = expected utility (or real income) in area $k$
(k = $i$ or $j$ at time $t$,*

$j$ = potential destination,

$i$ = origin,

$r$ = discount rate ($0 \leq r \leq 1$),

$T$ = expected length of remaining lifetime, and

$C_{ij}^t$ = cost incurred in time period $t$ of moving from $i$ to $j$.**

Dividing net real income gains by the discount factor, $(1 + r)^t$, adjusts for the fact that people typically attach less value to future occurrences than to the same occurrence today.*** Because uncertainty about the future may be greater in less developed countries (LDCs) than developed ones, discount rates may be higher in LDCs. If so, potential migrants there would attach less value to future gains.

The above equation embodies the essence of the microeconomic approach to migration decisionmaking. Other migration models, such as the Harris-Todaro model, and most sociological models, can be viewed as special cases of the general human capital model just

greater than that available on alternative uses of the funds invested in migration. The internal rate of return is the discount rate, $r$, which makes the present value of the net gain to the migration investment, $PV_{ij}$, equal to zero (i.e., the discount rate that equalizes the present value of benefits and costs).

* We assume that each person has a utility position, $U = f(x_1, x_2, \ldots)$, which translates objective variables (e.g., earnings) and subjective ones (e.g., climate) into a single summary indicator of well-being. Utility cannot be directly measured but can be inferred by observing behavior: People do (or do not do) something because they think, a priori, that choice will make them better off than the alternatives perceived to be available.

** For convenience, most analysts assume that all costs occur at the time of the move and regard any costs that occur later as negative benefits.

*** Having to wait reduces the present utility received from a future amount because of convenience of earlier availability, preference for assured receipt over contingent receipt (e.g., will the person still be alive?), and the productivity of resources over time. The discount, or interest, rate, $r$, is the "price" of earlier availability of rights to use goods rather than later availability.
presented. Specific models vary in which subcomponents of the Us and Cs they emphasize and in how they measure those subcomponents.

I now discuss some variants of this basic model that have been considered in recent economic studies, and conclude this section by contrasting economic theoretical models with those of other disciplines.

**VARIANTS OF THE BASIC MODEL**

**Family Considerations**

Most economists who use the model just presented implicitly assume that the decisionmaker is an individual. Several recent studies have considered the implications of relaxing this assumption and treating the family as the decisionmaking unit. We consider this extension in Sec. IV.

**Return and Repeat Migration**

Most migration analysis implicitly treat migration as a one-time event—a fait accompli. In any given year, however, most moves are not people's first moves, but are repeat moves: either onward to new locations or back to places where they lived before (return moves). Section V deals with return and repeat migration.

**Information Costs**

A common misconception of how an economist views migration (or the world in general) is that the economist assumes that people contemplating migration possess detailed information on all the costs and benefits of moving and then sit down, calculator in hand, and carry out the cost-benefit analysis summarized in Eq. (1) (e.g., Shaw, 1974). This is a mere convenience; economists recognize that real people do not behave this way. It is just that the outcome of their behavior often accords with such a model.

With perfect information and perfect foresight, the potential investor in migration should always correctly weigh the advantages and disadvantages in deciding whether and where to move. But information is not costless, and uncertainty is a fact of life. People weigh the advantages and disadvantages as they perceive them; their perceptions may be faulty, and they may consider only one or a few destinations (perhaps only those where they have friends or relatives). If a person moves, we infer that, ex ante, given the information available, he
or she believed the expected returns from migration would exceed
the costs; but this does not necessarily mean that, \textit{ex post}, things
worked out that way. Like migration, information-gathering about pos-
sible destinations is an investment, an activity that entails costs and
returns; a potential migrant will incur these information costs only if
he or she has reason to believe that moving will be profitable, and
will invest in "search" only as long as the benefits of having more
information are perceived to outweigh the costs of gathering it.*

\textbf{Uncertainty and Risk}

In deciding whether to move and where, the migrant compares the
\textit{expected} utility from living in each alternative location ($U_i$ and $U_j$
in Eq. (1)). The \textit{likelihood} that particular outcomes will occur will
affect the expected utility gain from moving. (Todaro (1969) brought
attention to this consideration \textit{vis-a-vis} the probability that the
migrant would find a job, i.e., the probability of unemployment.)**
Furthermore, in evaluating alternatives, the uncertainty associated
with them may not be a neutral consideration. This depends on the per-
son's attitude toward risk. A person who is \textit{risk averse} attaches dis-
utility to uncertainty and prefers a certain outcome to an uncertain
one with the same expected value in statistical terms. For example,
a migrant who is risk averse might prefer (i.e., gets more utility
from) a certain wage increase of $50 over a 60-40 chance of increasing
wages by $100, despite the fact that the latter has a larger expected
value in mathematical terms.*** Differences in attitudes toward risk,
like other differences in subjective valuations of factors in alterna-
tive locations, can impel two people to evaluate the same prospect dif-
ferently. Other things being equal, people should be more likely to
move the less averse they are to risk.

*See Goodman's (1979) and Haberkorn's (1979) papers prepared for
this workshop for more on \textit{search behavior} and the decision-making process.
** This is discussed further in Sec. III under the heading "The
Conceptual Variables are Expected Utilities."
*** In contrast, a \textit{risk preferer} prefers uncertainty and would choose a
50 percent (or perhaps even lower) chance of a $100 wage increase over a
certain increase of $50. Evidence suggests that most people are risk
averse in most situations, however.
The notion of risk is especially important here because migration can be a risky endeavor. This is partly due to its "lumpiness." Because only one choice of migration destination can be pursued at a time, "investors" in migration are unable to diversify commitments to reduce risk within the amount of resources allocated to migration (Schultz, 1977). One way a family can reduce risk is to have only the family member most likely to benefit from migration (e.g., the most highly educated son or daughter) move, while the rest of the family remains (at least initially) in the origin location—a common practice in developing countries. Additional information can also reduce risk and uncertainty, but can never eliminate it altogether.*

Income Effects

Because most of the costs of an investment, including one in migration, must be incurred before the returns can be reaped, those costs must somehow be financed. Banks are unlikely to lend for a risky investment in human capital such as migration, whose returns cannot be legally offered as collateral in most countries (because of prohibitions against slavery). Therefore, persons who already have some wealth, or who can turn to relatives for loans or other help at the destination (such as room and board) should be better able to finance the direct costs of moving and to afford to forgo wages while looking for a job after moving.

In addition to enabling people to finance a move, a higher income enables them to satisfy their desires for normal goods,* some of which may be only available through relocating—a more pleasant climate, for example (Graves, 1979; Keeley, 1980). This may explain why "non-economic" motives for migration appear to have become increasingly important in higher-income industrialized countries (e.g., Zelinsky, 1974; Long and Hansen, 1979).

* For more on the topic of risk and migration see Goodman (1979) and David (1974).
** A normal good is one that people buy more of as their income increases.
COMPARISON OF MICROECONOMIC MODELS WITH OTHER MICRO MODELS OF MIGRATION

Noneconomists often criticize some of the assumptions underlying microeconomic models. They argue, for example, that utility is not observable or measurable, that people do not actually make cost-benefit calculations, and that people do not always make rational decisions. However, the main assumptions underlying microeconomic models of migration are really no different from those in other disciplinary perspectives on migration, for example, Lee's (1966) model, the place-utility model (see the Brown and Gilliard (1979) paper prepared for this workshop), or the value-expectancy model (see De Jong and Pauw, 1979). The criteria upon which migration decisions depend in these other models—"factors" associated with origin and destination (Lee), place-utility, or value-expectancies—are no more directly measurable or observable than is expected utility in the microeconomic model; nevertheless, all four approaches assume that people behave as if they do reach a summary assessment. Furthermore, all four approaches assume that people make conscious decisions on whether to move and where and that they try to do what is best for themselves and their families. These decisions are "intendedly rational" ex ante but not necessarily actually so ex post. In addition, each of the disciplinary perspectives acknowledges that information is not perfect or costless; consider, for example, the mental or cognitive map or the search or awareness space, in the place-utility model.

Although the assumptions underlying the general models that various disciplines use to study migration differ little—the differences are mainly of terminology rather than of underlying ideas—economic models have some particular advantageous features. The human capital model explicitly recognizes that the benefits of migration accrue over a period of time, and hence focuses attention on investment-related concerns such as the time horizon of the investment, risk, discount rates, and ability to finance a move. Consideration of other income effects allows noneconomic factors, such as amenities, to be integrated into the model. And, as discussed in Secs. IV and V, family considerations and return and repeat migration can be easily incorporated into the basic microeconomic model.
Empirical implementation of the economic model often differs considerably from that of other disciplines. Economists try to understand actual behavior, rather than to describe survey respondents' reports of motivations, goals, or intentions.* This leads the economist to seek objective measures thought to influence the potential migrant's assessment of expected benefits and costs of migration. One result is that the relatively objective variables the economist considers, e.g., income, wage rates, availability of public services, are likely to be more amenable to direct policy influence than the more subjective variables others use, e.g., residential satisfaction. The next section of this paper discusses some of the issues and problems typically encountered in trying to operationalize microeconomic models for empirical analysis.

*Unlike some sociological models (e.g., Sell and DeJong, 1979), most economic studies do not specifically distinguish discrete stages in the migration decisionmaking process--e.g., desire to move, expectation to move, actual move--nor do they explicitly consider why some people who wish to move do not do so. Such sub-decisions are regarded as implicit in the information-gathering process. The individual is regarded as continually evaluating alternatives to determine whether he or she is doing the most personally advantageous thing. The results of an informal evaluation may motivate the person to gather more information about one or several prospective destinations. That additional information will result in a more informed (but still not perfectly informed) decision regarding whether to move. However, we cannot be sure whether the report of a "wish" to move reflects only vague uncertainty ("grass is greener on the other side") or is the result of a more formal evaluation of alternatives. Most economists regard talk as cheap and do not have much confidence in the reliability of data on people's intentions. We prefer to analyze what people do, not what they say they want to do.
III. EMPIRICAL IMPLEMENTATION OF THE BASIC
MICROECONOMIC MODEL OF MIGRATION

This section discusses some empirical approaches that have been
used to test the basic model just described. It first discusses
measurement of the dependent variable—migration—and then discusses
measurement and the main findings for some of the independent variables
typically considered.

THE DEPENDENT VARIABLE: MIGRATION

Most of the early statistical tests of the general model set
forth in Sec. II were based on aggregate census data in which migra-
tion is measured over a multi-year reference period. In the United
States the reference period is usually five years, e.g., 1965–1970.
In developing countries, the reference period is often the individual's
lifetime (i.e., the presence or absence of migration is determined by
comparing the person's place of residence at the time of the census
with the birthplace). The dependent variable is often a region-to-
region migration rate defined as:

\[
m_{ij} = \frac{M_{ij}}{P_i}
\]

where \( m_{ij} \) = gross migration rate from region \( i \) to region \( j \) between
time \( t_0 \) and time \( t_1 \) (\( t_1 \) = date of census, \( t_0 \) = some earlier
date, e.g., 5 years earlier or birthdate; \( i, j = 1, \ldots, \)
number of regions; \( i \neq j \));

*Some studies (e.g., Bowles (1970) and Fields (1976)) have used
net migration rates. However, net migration rates are not well suited
to testing micro models, because there is no such person as a "net
migrant." Typically, the denominator of a net migration rate is not
defined correctly (technically, all people at origin and destination
are at risk). Also, when net migration rates are analyzed, the effects
of explanatory variables on in- and out-migration cannot be allowed to
be asymmetric.
\[ M_{ij} = \text{number of people who lived in region i in } t_0 \text{ but in } \]
\[ \text{region j in } t_1 \text{ (this derives from information on places of residence at the times } t_0 \text{ and } t_1); \text{ and} \]
\[ P_1 = \text{number of people who lived in region i in } t_0 \text{ and were alive in } t_1. \]
Such a dependent variable can be interpreted as the probability that a representative individual living in region i at the beginning of the reference period will have moved to region j by the end of the period.*

With this dependent variable, one can explain both whether people move and where they choose to go.

There are several problems with measuring migration in this way. First, a comparison of places of residence many years apart underestimates the number of moves that took place during the period; this is especially true for "lifetime" migration rates. Persons who made multiple moves in the period (e.g., \( i \rightarrow k \rightarrow l \rightarrow j \)) are included only once in the numerator if their \( t_0 \) and \( t_1 \) places of residence differ, even though they may never have moved directly from i to j. However, people who made a sequence of moves during the interval and whose last move was a return to the place of residence at the beginning of the interval (e.g., \( i \rightarrow k \rightarrow i \)) are not even counted as migrants with such a measure. (This is discussed in more detail in Sec. V.)

Second, most studies of aggregate data lump together people of different ages, educational levels, and other sociodemographic characteristics. Studies that have stratified by these characteristics to reduce heterogeneity have shown that parameters differ substantially by these characteristics (e.g., Schultz, 1977; DaVanzo, 1972).

Third, most studies of aggregate data do not distinguish different types of migrants (return, other repeat, primary) and therefore mask the differences among them (exceptions include DaVanzo, 1976a; Kau and Sirmans 1976; Miller, 1977; and Long and Hansen, 1977). As we

---

*If the dependent variable is the number of migrants, it cannot be interpreted as a probability (even when the number at risk is an explanatory variable).
note later in Sec. V, people at risk to return or repeat migration appear to base their migration decisions on somewhat different considerations from those of people who never moved before.

Fourth, aggregate data often do not distinguish military personnel, whose migration may figure prominently in certain migration streams. Since military personnel are highly migratory, but for entirely different reasons from those of their civilian counterparts, data that include military moves both overstate the propensity for civilian migration and distort estimates of the importance of the factors influencing it.*

Fifth, the regions considered (in U.S. data, divisions or even states) are large and heterogeneous, and are difficult to characterize empirically. (For example, how does one measure the climate of the Pacific Division, which extends from Alaska to California and includes Hawaii, or its unemployment rate, which will vary widely among localities?) Moreover, all of the intraregional mobility of the population goes unmeasured, even though much of it would qualify as authentic migration (i.e., change of community of residence).

Sixth, with aggregate data one can explain only variations among groups (why one group has a higher migration rate than another) but not variations within each group (why particular members of the group migrate while others do not). Navratil and Doyle (1977) find that variances within labor market areas are greater than variance among them. Using data from the 5% Public Use Sample of the 1970 U.S. Census, they show that a regression where the observations refer to individuals yields coefficient estimates much closer to those predicted by economic theory than a regression whose observations are aggregated to refer to average characteristics of labor markets. When the latter are used,

*Typically, data do not distinguish students and retired persons, either. Because these groups also move for other than immediate job-related reasons, their inclusion in a sample can bias estimates of influences on job-related migration. For estimates of the extent of bias when students, retired persons, and those in military service are included in the sample, see Long (1976) or DaVanzo (1976b).
one cannot separate the influence of the personal characteristics from that of the area characteristics. Navratil and Doyle conclude that "the process of aggregation camouflages some of the personal characteristics which are important determinants of an individual's decision to migrate while it had only a marginal effect on the labor market characteristics of an area (p. 1558)."

Seventh, since the dependent variables are proportions whose values range between 0 and 1, the statistical technique most commonly used, ordinary least squares regression analysis (OLS), may yield predicted migration rates outside this range. A logit transformation can get around this problem. In addition there are several other econometric problems with the typical specification which estimates an ordinary least squares regression on the $m_{ij}$'s for all $i \neq j$: (1) Since the denominators of the rates, $p_i$, vary in size, variances of the dependent variables, and hence of the equation's error terms, are not equal but are inversely related to $p_i$. This problem of heteroskedasticity can be corrected by using weighted regressions in which each observation in the moment matrix is weighted by the size of the group to which it refers ($p_i$). (2) There are multiple observations on each sending area $i$ ($m_{i1}$, $m_{i2}$, $m_{i3}$, ... ) and the errors for these observations may be correlated with one another. For example, there may be some feature of $i$ not included in the equation (e.g., climate, attitudes) that makes all the outmigration rates from that area higher (or lower) than an OLS regression equation would predict. Similarly, there are multiple observations on each destination area $j$ ($m_{1j}$, $m_{2j}$, $m_{3j}$, ... ) and their errors may be correlated also. If the equation's error terms are so correlated, the OLS assumption of independent error terms

---

* Or some multiple thereof if rates have been normalized to, say, refer to migrants per 1000 or 10,000 at risk. The same point applies when number of migrants is the dependent variable and number at risk is an explanatory variable, for predictions of the former can possibly exceed the latter.

** Where the dependent variable is $\ln \left( m_{ij}/1-m_{ij} \right)$. 

*** There is also the issue that the sum of the outmigration rates from a given area cannot exceed one; i.e., $\sum_{j=1}^{n} m_{ij} \leq 1$ each $i$ (actually $\sum_{j=1}^{n} m_{ij} = 1-m_{ii}$ each $i$).
is violated and the resulting coefficient estimates will not be efficient. In this case, variance-components estimation techniques should be used to produce efficient coefficient estimators. The polytomous logit technique, which estimates an equation explaining how people choose among discrete alternatives, is the most appropriate one for handling all the problems discussed herein at one time (see Schultz, 1977). With aggregate data, the polytomous logit model can be estimated by OLS once the data are appropriately transformed.

More recently, migration studies have turned to longitudinal microdata, in which areas of residence are reported at the time of the initial interview and each following interview.* With such information one can create a 0–1 dependent variable** to indicate whether or not the individual moved between a particular pair of surveys.***

With microdata, migration can be defined with respect to any geographic unit as large as or larger than the smallest reported in the data. For example, if county of residence is reported for residents of the United States, one could construct variables measuring interstate or interdivisional migration as well as intercounty migration.

Another advantage of microdata is that the sample can be delimited for the purpose one has in mind. For example, studies of job-related migration can exclude retired people, students, and military personnel from the sample.

* Some surveys are administered only once. The only migration information available derives from retrospective questions on previous residences. However, contemporaneous information on the factors that influenced that previous migration is usually not available.

Some surveys are administered more than once but do not follow people who moved. For example, the U.S. Survey of Economic Opportunity surveyed residents of particular dwelling units in successive years, but did not follow first-year respondents who moved. Even surveys that try to follow movers are not always successful in doing so. Thus they are likely to underrepresent the sample members we are most interested in—migrants.

** In multivariate analysis, variations in the 0–1 migration variable can be explained by ordinary least squares ("the linear probability model") or by nonlinear techniques such as logit, probit, or discriminant analysis, which are preferable statistically. Also see Keeley (1980).

*** As with Census data, multiple moves within the migration interval cannot be detected.
While microdata are often superior to aggregate data for analyzing which individuals choose to move and why, they are often less suitable for studying how migrants choose among destinations. If the geographic areas are small, the migrant's potential destinations are likely to be too numerous to be manageable for a statistical model of individual choice. Aggregation of alternatives then becomes necessary. However, as one reduces the number of possible destination areas (and in doing so increases the size of each one), one also reduces the recorded number of migrants between areas.* Also, unless one is satisfied with simple cross-tabulations (say, average characteristics of selected destinations and of areas not chosen), statistical analysis of destination choice with microdata can be expensive. To explain how the characteristics of the migrant and of the potential destinations affected the migrant's choice among them, one must use a maximum likelihood multivariate technique such as polytomous logit analysis. (See DaVanzo (1976b, pp. 80-93; and 1978a) for an application of this technique to microdata from the United States and Schultz (1977) for an application to aggregate data from a developing country.)**

EXPLANATORY VARIABLES

In the theoretical model sketched in Sec. II, the potential migrant considers the expected utility of net benefits and costs of moving to each potential destination (that he or she is aware of) and chooses to move if there is at least one destination for which expected net benefits exceed costs. In this case, the person will choose the destination where the difference between benefits and costs, the expected net gain to moving, is greatest. Conceptually, this net gain is measured as the difference between the expected utility of the present discounted value of the lifetime real income stream the migrant

* The same problems occur for aggregate data also, but they are mitigated somewhat by the fact that aggregated data usually derive from a census, so that sample size is not a problem as it often is with survey data. The econometric problems in analyzing destination choice are ignored in most analyses of aggregated data.

** See De Jong and Fawcett (1979, pp. 5-9) for a discussion of other considerations in defining migration, and Hugo (1979) for discussion of sampling strategies.
expects to receive if he or she moves to that destination and the expected utility of the present value of the lifetime real income stream the migrant expects to receive if he or she does not move, less the costs of moving. Of course, the desired measure cannot be directly observed. In empirical analyses we seek to use explanatory variables that are likely to be highly correlated with this conceptual variable. The remainder of this section discusses some issues that should be considered in defining the empirical counterparts to the conceptual variable of interest.

The Conceptual Variables Are Utilities

The conceptual variables $U_i$ and $U_j$ in Eq. (1) refer to the individual's utility or well-being in the present and alternative locations. However, as noted in the explanatory footnote to that equation, utility cannot be directly measured, although certain factors are presumed to change utility in predictable ways. Economists assume that people receive utility from income, so that, other things the same, an increase in earnings, say from wage employment or self-employment, or in income from other sources (e.g., interest payments or welfare receipts) is presumed to increase utility. The size of the utility increase derived from a dollar increase in income (marginal utility) can vary among people and, for a given individual, may depend on his or her level of income.*

However, as noneconomists are quick to point out (though not necessarily in these words), the total utility a person derives from living in an area depends not only on his or her income there, but also on nonpecuniary considerations, such as the value of leisure time** or the pleasure received from living near friends or relatives or in an area with a pleasant climate.

---

* Economists frequently assume diminishing marginal utility of income, wherein each additional dollar results in a smaller increase in utility than that yielded by the preceding dollar.

Also for monetary variables, there is the issue of whether or not to adjust for regional cost-of-living differences. See Wertheimer (1970).

** For example, if the migrant receives a higher wage rate after moving but chooses to work proportionately less, so that his earned income is the same as before, he will still be better off because he enjoys more leisure time.
To my knowledge, no empirical analyses of migration have attempted to combine the various factors hypothesized to affect utility into a single indicator of utility, presumably because it is difficult to measure many of the factors and impossible to know what weights to assign to each factor. However, a number of studies have included as separate variables some of the key factors (e.g., wage rates, measures of climate) likely to affect an individual's utility in alternative locations (and hence the likelihood of the person's moving there). In effect, this approach lets the regression equation determine the relative utility weights on the various factors.

Economists' studies often do not include the nonpecuniary factors mentioned above but rather consider mainly "economic" variables, often no more than wages or money income. This may be because it seems safe to assume that, other things the same, any person derives additional utility or satisfaction from an additional dollar received. However, not all individuals may value nonpecuniary factors similarly. A cold climate may be a "plus" to a young adult who likes to ski but a "minus" to an elderly person with arthritis. Areas with good schools may be preferred by families with school-age children, but people whose children are grown up may prefer areas with lower local tax rates.*

Non-work-related considerations aside, it is important that studies of migration, especially in developing countries, begin to go beyond restricting their explanatory variables to mainly monetary wage incomes. This is because a sizeable fraction of the real incomes of many families in developing countries is not received in monetary form. For example, employees are often remunerated in kind (with food or housing) rather than with money wages. Furthermore, much of what the family consumes is produced at home in the nonmarket sector (e.g., homegrown agricultural products, products of cottage industry), rather than by using

* Sometimes there is a family characteristics, e.g., presence of school-age children, likely to distinguish those who prefer more of an areal attribute from those who prefer less. In this case, in empirical analysis we can interact those family characteristics with the indicator of the areal attribute, and hence allow the effect of the attribute to vary among the groups hypothesized to have different "tastes" for it. In this way areal effects can be "personalized."
money received from a wage-paying job to pay for a product or service sold in the market sector. This should be kept in mind for migration analyses attempting to measure a family's "income" in their current location and to characterize their opportunities elsewhere.*

The Conceptual Variable Refers to the Lifetime Stream of Real Income

In the theoretical model presented in Sec. II, the potential migrant compares two present discounted values of lifetime streams of real income: the one the migrant expects to receive at destination, and the one the migrant expects to receive if he or she does not move, plus the costs that moving would incur. Recognition of the fact that the benefits of migration occur over a period of time is, indeed, the distinguishing characteristic of the human capital model. However, the conceptual variables cannot be measured empirically. Even if we could measure real income, we would not observe it for future periods (and neither, of course, would the migrant). Most empirical tests of the human capital model have assumed that current wage differences between areas will persist or that wages in each area will grow at a constant rate.** Using explanatory variables based on these assumptions, early tests of the human capital model (e.g., Schwartz, 1968; Bowles, 1970) demonstrated that present discounted values of income differences between destinations and origins summed over the expected working life of the prospective migrant *** performed

*If the goods this family produces and consumes at home are sold by other families in the market, the market price could be used to value this family's production. If the good is not sold in the market, we could consider how much time family members actually spend on such production and attempt to estimate the value (addition to well-being) of the non-market activities each individual engages in. For an example of this approach see Kusanic and Davanzo (forthcoming).

**However, a migrant may choose a destination area precisely because he or she expects earnings to grow there at a faster rate than elsewhere. It is difficult to incorporate such a notion into empirical analysis because information is usually not available to construct a variable measuring the steepness of the expected wage profile.

***That is, \[ \sum_{t=1}^{T} \frac{y_t - y_i}{(1 + r)^t} \]
better than did the mere current differences \((Y_j - Y_i)\) in explaining variation in migration rates of various age-education groups.

We Cannot Observe What an Individual Would Earn in Other Areas

Not only do we not observe expected income streams over time in all potential locations, but also we do not even observe wages or income for a given individual in more than one place at a given point in time. The conceptual comparison is between what the potential migrant would receive by moving and by staying. However, we observe only one of these, depending on whether the individual moves. If the person moved between time \(t\) and time \(t+1\), we observe post-migration \((t+1)\) income at destination, but we cannot observe what \(t+1\) income would have been in the absence of a move. The latter has been approximated in empirical studies of both aggregate and microdata by the migrant's pre-migration \((t)\) income at origin or by the \(t\) or \(t+1\) income of "otherwise similar" individuals at origin, as discussed later in this section. But the question is how similar these other individuals are, since for an identical set of observed characteristics they chose not to move, while the person in question decided otherwise. The unobserved differences between the two groups that caused one group to move and the other to stay may well also affect the wages the migrants would have received had they stayed. For example, the migrants may have chosen to leave the origin precisely because they expected to receive less if they stayed or more if they moved than did the individuals who ultimately stayed. While their pre-migration income may give some clues to how the migrants expected to fare had they stayed (for at least it controls for some unmeasured characteristics of the individuals), it is by no means a perfect indicator.

\[
\begin{align*}
\text{where } Y &= \text{income} \\
\text{j} &= \text{potential destination} \\
\text{i} &= \text{origin} \\
\text{r} &= \text{discount rate (usually between 5\% and 15\%)} \\
\text{T} &= \text{expected number of years till retirement.}
\end{align*}
\]

* Perhaps one could try asking a survey respondent about income expectations if he or she stays in the present location.
Similarly, we do not observe what a person who stayed could have earned by moving. As a proxy for that quantity, some empirical studies have used the wage or earnings change experienced by "otherwise similar" people who did move. Polachek and Horvath (1977) did so, for example, in a study using microdata. Other analyses have used, as a proxy, the wages or earnings received by "otherwise similar" people at destinations to which the person in question could have moved. Aggregate data were used for that purpose by Schwartz (1968), O'Neill (1970), and DaVanzo (1972), and microdata were used in a study by DaVanzo (1976b).

Such proxies, however, are subject to the same potential biases noted in the preceding paragraph. We call these selectivity biases. They arise because the outcomes we observe depend on the choices made, which in turn were based on expectations about those outcomes. That is, we observe post-migration income at destination only for individuals who chose to move; and they chose to move precisely because they expected to receive more real income. We observe post-nonmigration income at origin only for nonmovers. They have chosen not to move precisely because they believed they were earning more real income than they could anywhere else. Therefore, their experiences may not be indicative of what migrants would have received had they not moved, just as the post-migration earnings of migrants may not be indicative of what nonmigrants would have experienced had they moved. Two recent studies (Nakosteen and Zimmer, 1980; DaVanzo and Hosek, 1980) suggest that the extent of selectivity bias in procedures usually used to estimate the actual or potential income gains from migration is not negligible.

Let us now look in more detail at the main empirical approaches that have been used to estimate what potential migrants could earn at origin and in alternative potential destinations.

Most studies of aggregate Census data have assumed that potential migrants earn the same income at origin as does the average or median individual there, and that if these people were to move to another

*See Lewis (1974), Heckman (1979), or Maddala (1978).
area, they would earn the same income in this new location as the average or median individual currently residing there. Using migration and income data on relatively homogeneous demographic subgroups helps make the income variables somewhat better measures of the income an individual in the subgroup might receive in each possible area (e.g., DaVanzo, 1972; Schultz, 1977), but it by no means avoids the problem of selectivity bias. For example, the average income of a destination's already-present residents may include returns to location-specific assets (e.g., an established business with a steady clientele); new migrants are unlikely to possess such capital.*

Kaluzny's (1975) analysis of microdata from the Panel Study of Income Dynamics (PSID) uses an explanatory variable measuring "expected gain to the head from leaving the current location," $EG$, that is based on aggregate data. This variable, which is used to explain 1968-69 and 1969-70 migration rates by race and by poverty status, is defined as

$$EG = \sum_{j=1}^{n} \left( P_{ij} \left[ Y_j - Y_i \right] \right) / Y_i$$

where $j$ indexes metropolitan and nonmetropolitan areas of the nine census divisions ($n = 18$);

* Analysis using census income data suffer from other problems as well. Aggregated income data are often not reported for as detailed groupings as migration data. For example, the 1960 U.S. Census reported interdivisional migration for subgroups of the population stratified by age, education, sex, and race. However, it reported median divisional incomes only for groups classified by sex, race, and education simultaneously, but not by age also. The four-way classification was available only for the four Census regions.

Furthermore, the income data in the U.S. Census refer to income over the calendar year preceding the Census. For most migrants this will be the income they received after moving (possibly as many as three or four years after moving, since the U.S. Census uses a five-year migration interval). To the extent that migration affects an individual's or an area's income, post-migration income may be the effect of migration, rather than a determinant. Thus the coefficient of such an income variable may be subject to a simultaneity bias in an equation explaining migration.
\( p_{ij} \) = the probability the \( j \)th area will be selected; \( p_{ij} \) is defined as the proportion of migrants, by race, from a given census division who moved to each of the 18 potential destinations during the 1955-60 period;

\( Y_k \) = income prevailing in area \( k \) (\( k = i \) [origin] or \( j \) [potential destination]); \( Y_k \) is based on 1967 Social Security estimates of mean earnings by sex, race, and age.

This EG variable was highly significant in explaining which sample members moved from a metropolitan (nonmetropolitan) area of a state to a nonmetropolitan (metropolitan) area of that same state or to a different state. However, such a variable suffers from the same problems of selectivity bias as discussed in the preceding paragraph.

In my 1976(b) analyses of microdata, also from the PSID, I estimated separate wage equations for the nine U.S. Census divisions and used them to impute what each individual could earn by moving there, given his or her personal characteristics. Separate wage equations were estimated for males and females. Explanatory variables included such personal characteristics as age, education, work experience, occupation, and industry. Estimating separate equations for each geographic area allowed the "values" of these personal attributes to vary among labor markets. The inclusion of a fairly rich set of explanatory variables allowed the wage to be tailored, to the extent possible, to the particular individual's characteristics.

After making certain adjustments to these imputed wages (for example, to adjust for the fact that migrants would likely start out in a new location with a new job), they were used to calculate the present values of earnings differences for each potential destination area and were found to help explain how migrants choose among destinations.

To explain whether or not a person moves (rather than the person's choice of destination), these present values of earnings differences for eight different areas were reduced into a single indicator of whether the person could gain by moving somewhere else. The resultant variable played a significant role in explaining who chooses to move.
Polachek and Horvath (1977), who also use the PSID data to explain why people move (but not choice of destination), use a different approach to estimate what the potential migrant could receive if he or she moved. They estimate separate equations for husbands and wives explaining the difference between the individual's wage at the end of the migration period and at the beginning, and include as one of the explanatory variables a dummy indicating whether the person migrated during the period.* This dummy is also interacted with several explanatory variables measuring characteristics of the person (e.g., age) to allow the effect of migration on earnings-change to vary according to those characteristics. Polachek and Horvath then use this equation to estimate the earnings-change that nonmigrants with particular characteristics would have experienced had they moved, and what migrants' earnings-change would have been had they not moved. Polachek and Horvath include both the difference between the husband's with-migration and without-migration estimated earnings-changes and a comparable variable for the wife, and find that both the husband's and wife's expected gain variables have significant power in explaining whether the couple moved.

Both the Polachek-Horvath and DaVanzo approaches to estimating the potential gains to migration have their pros and cons. My study

---

* One problem with data sets such as the PSID (the data set used by Polachek and Horvath, Kaluzny, and DaVanzo, 1976b) is that the annual surveys take place in the spring of each year (and thus migration refers to a spring-to-spring interval), but respondents report wages, hours worked, income, and so forth for the calendar year preceding the survey. Thus calendar-year economic measures do not always unambiguously precede or follow recorded moves. For example, for a 1971-72 move, 1970 (or earlier) income as reported in the 1971 (or earlier) survey would unambiguously precede the move, while only income for 1973 (or later) as reported in the 1974 (or later) survey would unambiguously follow the move. Polachek and Horvath consider the effects of 1971-72 moves on "Change in earnings 1971-72," but do not specify whether the latter 1971 and 1972 refer to the years of the surveys or are the calendar years to which the earnings variables pertain. (Regardless of which it is, the variable does not compare earnings unambiguously before and after migration.)
ignores information on actual wage gains by migrants, but allows for regional differences in wages. The Polachek and Horvath approach assumes that the potential wage gain to migration depends only on the potential migrant's personal characteristics and is independent of pre-migration geographic location, but it incorporates information on individuals' actual post-migration (or post-nonmigration) experiences. Some combination of the two approaches would probably be best, although, as noted earlier, both are likely to suffer from selectivity bias. Nonetheless, in allowing estimates of potential gains to migration to be a function of the potential migrants' personal and family characteristics, both the Polachek-Horvath and DaVanzo approaches are far superior to earlier approaches, which assumed that all potential migrants would earn the average area (modern sector) wage, for example, as all empirical tests of the Todaro (1969) model have.

The Conceptual Variables Are Expected Utilities

The relevant conceptual variables are expected utilities of expected real income at origin and destination. The expected values will reflect the probability that the potential migrant will actually receive the various components of real income; expected utility adjusts for attitudes toward risk.

Since wage income is the largest component of real income for most people, and possibly the component most likely to change if they move, an important consideration is the likelihood that the person will have a job—i.e., the probability of employment. Todaro (1969) was one of the first to emphasize that potential migrants should base their decisions on differences in expected earnings. He deals mainly with rural-urban migration in developing countries and points out that in this context urban earnings should be multiplied by the probability that the potential rural-urban migrant would find an urban job if he or she moved. Such a formulation is an improvement over those used previously, for it acknowledges that not all migrants will be able to obtain a high-paying urban job; nonetheless, the Todaro model has
several shortcomings. For one, it ignores attitudes toward risk. It is assumed that the migrant is indifferent between a 20 percent chance of earning $99 or a 99 percent chance of earning $20. As noted in Sec. II, unless the person is neutral to risk, he or she may not receive the same *expected utility* from these two prospects, even though their expected value is the same in statistical terms. In this case, it may be better to include the two components of the expected value—the probability of obtaining a job and the wage that would be received if that job were obtained—separately in empirical analysis.

Another problem with the typical empirical formulation of the Todaro model is that the data used for the probability of employment and the average employment wage are based on the average experience of residents of the urban area, often only those with modern sector jobs. These statistics may not be indicative of what a poor, uneducated rural peasant would experience in the city. A preferable approach is to estimate a wage appropriate to the particular potential migrant under consideration, given his or her personal characteristics, as described in the previous subsection, and to do likewise, in estimating the probability that this particular person will actually receive that wage; that is, in addition to imputing a wage tailored to the potential migrant's individual characteristics, we would impute to each individual a probability of unemployment for each potential destination area based on the individual's personal characteristics and the way they affect personal unemployment probabilities in alternative geographic areas. This should be vastly superior to using the area unemployment rate, for that would implicitly assume that every individual has the *same* probability of unemployment in a particular place.

Despite these shortcomings, empirical tests of the Todaro model have found that it predicts better than a formulation that ignores the probability of finding a job (see Todaro, 1976, pp. 69-71, for a summary of evidence). In fact, Fields's (1976) analysis of U.S. migration suggests that availability of jobs may be *more* important in influencing migration decisions than is the income paid once migrants are settled in those jobs. However, except for Hay's (1974) analysis
of Tunisian migration, which used proxy variables in lieu of data on actual urban income and employment rates, to my knowledge there have been no tests of the Todaro model with microdata.

The Todaro model deals with the possibility of unemployment at destination, but is not concerned with the likelihood of unemployment at origin (this is assumed to be irrelevant in an LDC setting). U.S. studies that used aggregate data to assess the relationship between origin unemployment and out-migration have obtained mixed results, but unemployment is often measured at the end of the migration period and hence may have been affected by the intervening migration (simultaneity bias).

Aside from these simultaneity (and definitional) problems, aggregate data cannot tell us whether an estimated positive relationship between out-migration and the origin unemployment rate is due to the unemployed being more likely to migrate, or to people being more likely to leave areas with high unemployment rates, regardless of their own employment status.

* For example, Lowry (1966) found out-migration to be insensitive to local labor market conditions. Lansing and Mueller (1967, p. 91) report data showing that people are slightly less likely to leave depressed areas than nondepressed areas. Some analysts (e.g., Blanco, 1964; Mazek, 1969; Olvey, 1970; and Fields, 1976) have argued that unemployment "push" at origin is operative but that its role is masked by improper measures of true unemployment conditions. Blanco, Mazek and Olvey propose and use synthetic measures such as "prospective" or "potential" unemployment, which estimate what unemployment rates would be if there were no (net) out-migration, and find them positively related to (net) out-migration. Fields (1976) has suggested that the inability of previous studies to find a significant relationship between migration and unemployment rates may be due to the way the latter are measured. The unemployment rate, which pertains to the entire stock of workers and hence includes employed workers who are secure in their current jobs, may be an imperfect index of the tightness or looseness of a labor market. Fields shows that measures of turnover in a labor market (accessions and separations) perform better empirically than beginning-of-period unemployment rates in explaining net migration into 20 SMSAs for the period 1965-70. However, because Fields' dependent variable refers to net migration, one cannot distinguish the separate influences of these turnover variables on in- and out-migration.
Studies using survey data (e.g., Saben, 1964; Lansing and Mueller, 1967; Hay, 1974; Navratil and Doyle, 1977; DaVanzo, 1976b or 1978b) have confirmed that unemployed workers are more likely to migrate than employed workers. Furthermore, I found that local economic conditions (origin unemployment rates) do affect out-migration, but only within the subset of people most seriously affected by them—the unemployed. I also found that the unemployed and other persons looking for work are more responsive to the other economic determinants of migration (family income, origin wage rates, and expected earnings increases) than are persons apparently satisfied with their jobs. Bartel (1979) shows that, in addition, there are important differences according to the reason why the husband left his last job (quit, layoff, or transfer).

Studies using aggregate data to assess the effect of destination unemployment rates on destination choice often suffer from the simultaneity bias and definitional problems noted previously. As noted earlier, to my knowledge, no studies of microdata have attempted to assess the effect of destination unemployment rates on destination choice.

Another factor that affects a potential migrant's expectation about the real income he or she could receive in alternative localities is the quantity and quality of information about opportunities in those places. Many economists hypothesize that educated people possess more (and better) information because of both their superior ability to process information and their tendency to compete for jobs in labor markets that are more national in scope (e.g., Schwartz, 1973). This may be one reason that migration rates increase with education.*

Ability to Finance an Investment in Migration Affects the Likelihood of Moving

As noted in Sec. II, income at origin may not only measure what a person would forgo by moving, but also may be correlated with the

* For more on the role of information, see Goodman (1979).
person's ability to finance an investment in migration (and may have "consumption effects" as well. (See earlier discussion in Sec. II under the heading "Income Effects.")

The size and sign of the independent influence of income on the propensity to migrate is a subject of potential policy interest because, if properly estimated, it can indicate how an income supplement (e.g., from an income maintenance program, tax rebate, increase in Social Security benefits, or relocation bonus) might affect people's likelihood of moving. Will it make them more likely to move because they can better afford the costs of moving? Or will it make them less likely to move because they can better afford, or prefer, to stay where they are? If the former is the case, a policy designed to retard out-migration by giving people more money where they are could backfire if they used the money to finance an investment in migration.

However, it is difficult to measure the independent effect of income on migration, because a person's income in the current location also reflects how well off he or she is there vis-a-vis other places the person could live (i.e., $U_i$ in Eq. (1)). To estimate the independent effect of an increase or decrease in income due to, say, an income supplement or relocation bonus, one should use an income measure likely to induce the same behavioral response as the exogenous change in income that would be brought about by the supplement or bonus. The measure should exclude income components that have their own independent influences on migration decisions. The appropriate income variable thus should exclude wage income, which affects the earnings-return available to migrating and the opportunity cost (wages forgone) while moving and looking for a new job; the variable should also exclude returns on location-specific assets (e.g., farm, business, and professional income), which also have their own independent (negative) effects on migration.* It is not possible to construct

*Alternatively, one can leave these components in the income measure but also include them as separate explanatory variables. The coefficient of the desired (exogenous) part of income will be the same either way.
the desired measure with available aggregate data on average or median incomes in an area, but should be possible with microdata that report the appropriate components of income.*

*Studies that have attempted to estimate the size and sign of the income effect on migration have had mixed results. O'Neill's (1970) study of aggregate Census data found a positive coefficient on origin income entered alone when the difference between destination and origin income was held constant. She attributed this to migration being a "positive consumption good"—people like changes. She noted that a positive independent effect on income at origin can explain why previous studies (e.g., Lowry, 1966) found the push of origin conditions to be weaker than the pull of destination conditions. My 1972 study of Census data found the relationship between income and migration to depend on the distance of the move: The income effect is generally positive and becomes larger in algebraic value as distance increases (presumably because financing is a more important consideration the costlier the move); however, the income effect was estimated to be negative for short moves (perhaps because people, on balance, dislike moving).

Kaluzyn's (1975) analysis of the micro-level PSID data estimates a positive effect of family income (from all sources) on the likelihood that the family will move; the size of his estimated income effect is inversely related to the age of the household head and is larger for whites than non-whites.

In my 1976b study, I used nonwage, non-location-specific income as my income variable and allowed its coefficient to differ depending on the household head's pre-migration employment status. I found the coefficient to be negative and significant for the unemployed (and insignificant for others) and conjectured that since many unemployed are laid off, those with higher nonwage income may be better able to afford to wait for their jobs to become available again.

Analyses of data from the Seattle, Denver, and Gary Income Maintenance Experiments (e.g., Keeley, 1980) suggest that receipt of such income maintenance payments eases the financial constraints on moving and enables families to consider more alternative destinations than they might otherwise. In the Gary Income Maintenance Experiment, families receiving experimental payments were somewhat more likely than control families (those not receiving experimental payments) to move farther from Gary, to move to large urban areas, and to move to places where they had not lived before or to places where they had no friends or relatives. However, Keeley's analysis of data from the Seattle-Denver experiment suggests that this may be due to price effects of the negative income tax rate, which results in a reduced shadow price of consuming "environment." Keeley estimates a negative pure income effect for white males.
Other family or community members may help the migrant finance his or her move; this practice is especially prevalent in developing countries. Those at origin may provide money to help cover costs of moving. Friends or relatives at destination may provide food and lodging during the adjustment period. All these considerations should be kept in mind when designing surveys to provide data to explain migration decisions where persons other than the migrant may help finance the move.

**Distance is a Commonly Used Proxy for Many Costs of Moving**

In estimating models explaining place-to-place migration rates, the distance between a pair of areas is the most typically used proxy for the cost of moving between the two \( C_{ij} \). This is because many moving costs are likely to be greater the longer the distance moved: expenses for transportation and moving of belongings, wages forgone while moving, costs of gathering information about potential destinations, and actual and psychic costs of visiting and keeping up with friends and relatives left behind. (Also, the greater the distance between a pair of places, the more likely there are to be attractive "intervening opportunities." ) Indeed, one of the most consistent findings in empirical migration analyses for both developed and developing countries is a strong negative relationship between the distance separating two places and the relative size of the migration flow between them (other things being the same). *

* An exception to this distance-decay relationship is discussed in Sec. V.

For Census data that include military personnel, the estimated negative relationship between migration and distance becomes even stronger when explanatory variables are included to control for the relative number of military personnel in the origin and destination areas (and hence for the likelihood that some of the moves in question are military moves, which are not as likely to be deterred by distance as are civilian moves) (DaVanzo, 1972) or when military personnel are omitted from the sample (Long, 1976). The former study also shows, through an income-distance interaction, that distance becomes less of a deterrent to migration as income increases, presumably because higher-income people can better afford the costs of moving.
Nevertheless, it is questionable whether each of the costs mentioned above is truly monotonically related to distance. For example, the amount of information a potential migrant has about opportunities in alternative areas may be related to the size of those places and to the number of friends and relatives he or she has there. For example, an economist at The Rand Corporation in Santa Monica, California, probably knows more about job opportunities in Washington, D.C. (because of travel there and acquaintances with people working there) than he or she does about opportunities in Las Vegas, Nevada, only 250 miles away. Just as communications and information networks may be unevenly developed geographically, so may transportation networks. A Thai farmer living in a village 400 miles north of Bangkok may know more about opportunities in Bangkok than in nearby cities that are more difficult to travel to (perhaps because of intervening mountains). Similarly, the "social distance" between two places may depend on cultural and language differences or similarities. Because distance is not always monotonically related to the costs it is meant to proxy, future studies should attempt, whenever possible, to include alternative measures for these costs.

The effect of distance on migration is rarely investigated with microdata because of the aforementioned problems in using them to study destination choice.
IV. FAMILY CONSIDERATIONS

This section considers the implications of viewing the family, rather than the individual, as the decisionmaking unit. Other family-related considerations—e.g., family members as sources of information, utility or disutility (affinity), or financing of investments in migration—are discussed briefly in other parts of this paper and are dealt with in much more detail in Sarah Harbison's (1979) paper prepared for the workshop.

THEORETICAL APPROACHES*

Since members of a family often move together, the family is the appropriate unit of analysis for studying the migration of people who live in families. This means that a family member's migration decision is motivated by expected returns and costs to the whole family, instead of to the member, who might otherwise arrive at a different decision.**

Take the example of a two-person, husband-wife family in which the present value of the husband's expected returns from migrating (monetary and nonmonetary) exceed his expected costs, but the opposite is true for the wife. The couple would decide to move if their joint returns exceeded their costs. (The husband's net gain from migration exceeds in absolute value the wife's net loss.) In this case the wife would be a "tied mover," her migration decision being tied to her husband's. If she had been single, she would have decided to stay. If, for this same example, the absolute

---

*This subsection draws heavily on a recent article by Jacob Mincer (1978).

**The discussion in this and the next three paragraphs in this section assumes that all members of a family move together. Below, I discuss the implications of relaxing this assumption.

Throughout this section we implicitly assume the existence of a family utility function (the U's in Eq. (1) would be family utilities) and ignore the issue of interpersonal utility comparisons. (For example, how much weight does the husband put on his wife's unhappiness about leaving her mother, whom he will enthusiastically leave behind?)
value of the wife's net loss exceeded that of the husband's net gain, they would choose not to move, and the husband would be a "tied stayer."

The above examples assume that there is only one destination. In that case, if returns exceeded costs for both the husband and wife considered separately, they of course would move there. If two or more destinations are promising, one may be optimal for the husband and another for the wife. In this case, they will choose the one that offers the maximum net gains for the family. Following this rule may result in (1) their not moving at all, because their reciprocal gains and losses cancel each other out; (2) moving to either the husband's or wife's optimal location; or perhaps (3) moving to a third location, where neither one's personal gains are maximized, but the sum of the two is greatest.

The model can also apply to families with more than two members. For example, other adults in an extended family would have their own costs and returns to migration that could affect the family's migration decisions. (This may be a relevant consideration in developing countries, where many people live in extended families.) Children may increase not only the family's direct costs for transportation and moving belongings, but also the potential value of location-specific capital lost (because of friendships and schooling arrangements). If family costs of moving increase more than returns as family size increases, migration propensities should be inversely related to family size.

So far, we have assumed that members of a family always move together, sometimes sacrificing their own potential personal gains for those of other family members. However, if that sacrifice exceeds the gain they perceive from living with the family, they may choose to leave the family, either permanently or temporarily. Husband-wife couples may resort to divorce or separation, for example. A common manifestation of this choice in many developing countries is seasonal migration, whereby one spouse moves to another part of the country during part of the year in response to attractive employment opportunities there, but returns to the family for the
rest of the year. (It may be too costly to move the entire family, either in terms of direct costs or in terms of origin opportunities that would be forgone.) Similarly, other family members—children or members of extended families—may choose to leave those families (temporarily or permanently) if the potential personal gains from migration outweigh the perceived advantages of remaining with the family. In fact, it may be a deliberate strategy for a family to "diversify its portfolio" by financing an investment in the migration of the family member most likely to benefit from moving. The expected "returns" on that investment include remittances back to the family members remaining at origin.

Since a husband or wife may have been forgoing a private gain or incurring a private loss from migration as a tradeoff against the gains from marriage, marital dissolution should restore or create new private incentives. This explains why we observe, for a time after dissolution, higher migration rates for separated or divorced men or women than for people who are currently married or were never married.

EMPIRICAL APPROACHES

Testing the model just described requires fairly detailed data on all members of a family.* With such data, we can estimate how a particular individual's migration decision is affected not only by his or her own characteristics and opportunities, but also by those of other family members. Ideally, the data should attempt to follow all family members who move.

*Most early economic studies ignored the family context of migration decisionmaking because their data did not enable them to identify families. In my 1972 econometric analysis of aggregate census data, I used weighted averages of male and female incomes as "family" measures of income. Although a number of fairly strong assumptions were used in constructing these family measures, they often performed better in explaining the migration of males and females than did measures that referred only to the sex whose migration behavior was being studied, or only to the opposite sex.
Several recently available longitudinal data sets in the United States provide the desired data.* Some analysts using these data have restricted their working samples to intact husband-wife couples, and have included both the husband's and wife's characteristics (either separately or in combined "family" measures) as explanatory variables. Other studies include all sampled individuals of a particular sex in their working sample and attempt to assess the influence of marital status.

The main findings of these U.S. studies include the following:

1. Other things being the same, a family's propensity to move is generally inversely related to the wife's labor-force commitment. On average, families of women who work are less likely to move than are families of other married women. The likelihood of migration decreases the longer the wife has worked for her pre-migration employer (Sandell, 1977).

2. However, the size of the effect of wife's work on whether the family moves varies with the wife's age, hours of work, her contribution to family earnings, and the husband's reason for leaving his job, and may in certain cases be positive. In particular, the following types of families with working wives appear to be more likely to move than those with nonworking wives: families in which the wife is fairly young and earns less than she

---

*For example, the University of Michigan Panel Study of Income Dynamics (PSID) collects detailed information on the household head, and somewhat less (though still more than most data sets) on the spouse. Some information is also available on other family members. The PSID attempts to follow all members of survey households in subsequent rounds, including those who "split off" from sampled households. The National Longitudinal Surveys (NLS) also provide data on husbands and their wives. The NLS surveys that focus on women report relatively little on the characteristics of husbands, while the surveys of men report relatively little about the characteristics of wives (less than the PSID).

**Long (1974) also found this in earlier tabular analyses of U.S. Census data.
could elsewhere (DaVanzo, 1976b); and families in which the husband is relatively young and is transferred to a new location by his employer (Bartel, 1979).

- The effect of the wife's wage rate or earnings on the likelihood that the family will move depends on the husband's age and whether he quit, was laid off, or transferred from his premigration job (Bartel, 1979).

- Families with working wives do not move as great a distance as families in which the wife was not working before moving. Migration distance (in miles) is negatively related to the wife's premigration wage rate (Bartel, 1975).

- Wife's work status, earnings, and education, and the presence of school-age children are the major reasons for the difference in migration behavior between married and other persons. Marital status has a statistically significant negative effect on migration propensities in a simple correlation, but loses significance with the inclusion of the aforementioned factors (Mincer, 1978).

- In choosing among destinations, families appear to select those where husband's and wife's potential earnings gains are both large (DaVanzo, 1976b). In fact, that study estimated that chosen destinations have relatively higher wage gains (compared with destinations not chosen) for wives than for husbands.*

*I have not mentioned here findings regarding the differential effects of migration on the earnings of husbands and wives. For these, see Sandell (1977), Long (1974), Polachek and Horvath (1977), and Mincer (1978), and references therein.
V. RETURN AND REPEAT MIGRATION: AN ILLUSTRATION OF LINKING HUMAN CAPITAL ORIENTATION TO NEW ISSUES OF MIGRATION

THEORETICAL CONSIDERATIONS

Most economic studies of migration implicitly treat migration as a one-time event, and make no effort to distinguish repeat moves from first-time moves. Demographers (e.g., Goldstein, 1964; Morrison, 1971) have long since demonstrated that most moves are made by people who have moved at least once before; and in developing countries, such alternative forms of mobility as commuting, circulation, and seasonal migration appear to be prevalent (e.g., Hugo, 1979). A few economic studies have recognized this (e.g., Bowman and Myers, 1967; Vanderkamp, 1971 and 1972; DaVanzo 1976a; Kau and Sirmans, 1975; Yezer and Thurston, 1976; Allen, 1979) and have begun to investigate the differences among people moving for the first time, those returning to a place where they lived before, and those moving again but to a new location.

For example, Bowman and Myers (1967) hypothesized that if some degree of learning-by-doing is associated with moving, people who have moved before should find it easier (less costly) to move again. This may explain why repeat migration rates exceed primary migration rates.

In a recent study at Rand (DaVanzo and Morrison, 1978 and forthcoming; DaVanzo, 1978a and forthcoming), we have examined why some migrants choose to return to a location they have recently left, others choose to move on to a new location, and still others choose not to move again. We view the prospective migrant as a person making an investment in human capital, but recognize that people weigh the advantages and disadvantages of moving as they perceive them, and that in some instances their knowledge of potential destinations may be very limited, for example, to places where they have friends and relatives.
Challenged to explain why anyone should move back to a place he previously decided to leave, this analytical framework can offer several guiding hypotheses, derived from two concepts that are central to the analysis: location-specific capital and imperfect information.

"Location-specific capital" is a generic term denoting any or all of the diverse factors that "tie" a person to a particular place. It refers to both concrete and intangible assets whose value would be lost or would steadily diminish if the person moved somewhere else: for example, job seniority, an existing clientele (as in the case of a well-regarded doctor or carpenter), a license to practice a particular profession in a certain geographic area, property ownership, personal knowledge of the area, and community ties and close friendships. The concept of location-specific capital suggests two propositions to explain return migration:

1. When a person who has migrated before moves again, he or she should favor some former place of residence as the destination because the person has location-specific capital there. Other things being the same (including length of absence), the more location-specific capital that is left behind, the greater should be the propensity to return.

2. The longer the interval of absence, however, the weaker should be the propensity to return, because most location-specific capital depreciates in value. For example, the carpenter's clientele cannot wait indefinitely for the migrant to return; old friends may die or become migrants themselves; and of course, the value of information

*Sociologists and psychologists have used such terms as "vested interests," "neighborhood or community or psychological integration," and "social and economic ties" for this concept. The amount of location-specific capital in one's current location is an important influence on the decision whether to leave that location. For example, this explains why self-employed persons are less likely to move than otherwise similar salaried workers. In fact, location-specific capital at origin usually is the main reason people give for not moving (e.g., see DeJong and Fawcett, 1979, pp. 27-29).
about an area depreciates as conditions there change. However, some people who leave an area may intend all along to return, and may continue to invest in their location-specific capital there to keep it from depreciating. In developing countries, for example, people who leave their home village sometimes attempt to maintain the value of their location-specific visits back or by retaining ownership of property. Indeed, the person may "circulate" between his home village and a city with better employment opportunities as a way of gaining some or all of the benefits at the latter location while retaining those at the former (Hugo, 1979). (Hugo notes that such a strategy has the advantage of spreading the risk of failure between two places.)

The concept of imperfect information allows for the fact that moves sometimes turn out to be unwise investments in human capital. The migrant may have overestimated the net benefits of living in the new place, or the anticipated benefits may not have materialized at all. (Perhaps the migrant did not improve his earnings, or failed to find a job.) If so, the person is likely to see the light reasonably soon—within a year to two, not ten or twenty—and "reinvest" in migration soon thereafter. A third explanatory proposition, therefore, is:

3. The concept of imperfect information accounts for return migration as a "corrective" act—one that may often be guided by the migrant's superior information about a familiar area. A migrant who is unable to process information effectively should be more likely to return than one who based his or her initial move on more accurate information.

*For example, Speare (1971) found that only 18 percent of Taiwanese migrants in his sample who had been homeowners before moving sold their homes when they moved. Also see Nelson (1976).
(Also see Yezer and Thurston, 1976; and Allen, 1979.)

A final proposition is:

4. Each move entails some "learning by doing." A person who migrates from A to B and judges the outcome to have been favorable has reasons to stay there; this one success may even embolden the person to venture yet another move. The person who judges the outcome as unfavorable may be less inclined to stay but also may be less venturesome in the future. Consequently, this unsuccessful migrant, if he or she decides to move at all, may do so in the hope of regaining an earlier equilibrium of life by returning instead of braving the unknown (and risking failure) once more, and after returning might be less venturesome about "trying" migration again in the future.

A systematic process of self-selection is implied by this last proposition: Migrants who return should differ from those who do not in their motives, characteristics, and circumstances both before and after the initial move and any subsequent moves. The success or failure of the initial move, then, may serve as one important determinant of the subsequent propensity to migrate and the type of repeat move that will be made.

**EMPIRICAL FINDINGS**

The four propositions above suggest how the propensity to return should change as the interval of absence lengthens, and why some people return while others do not. To examine these propositions empirically, we have used longitudinal data on internal migration in the United States from the University of Michigan Panel Study of Income Dynamics. These data enable us to examine sequences of moves (within which individual moves can be interpreted) and to measure migration at one-year intervals. We are thus able to detect many more moves than can the Census, with its five-year migration
interval.* Over 70 percent of the annual moves in our sample are repeat moves (by persons who moved at least once before). Around 30 percent of these repeat moves are returns back to places where people lived before. We estimate that a migration measure that compares residences at the beginning and end of a five-year period misses roughly one-third of annually measured moves of people who are observed over the five successive years. Half of those undetected moves are self-cancelling initial and return moves. (The other half are moves in nonreturn repeat sequences.) This sizable degree of camouflaged movement arises in part because return moves are disproportionately concentrated in the first few years after the initial move in a sequence.

We find that the propensity to return varies in two noteworthy ways that are consistent with the theoretical expectation that the attractions of a previous locale derive from the potential returnee's ever-depreciating location-specific capital there. First, the propensity to return declines sharply as the interval of absence lengthens. Second, it depends on the person's degree of past attachments to a previous location. At each interval of absence, the

![Graph](graph.png)

Fig. 1 --Average annual rates of short-interval return migration, by interval of absence and type of return (reproduced from DaVanzo and Morrison (1978), p. 12)

*Since the U.S. Census data report areas of residence only at birth, five years before the Census (t - 5), and at the time of enumeration (t),
probability of a return move is always at least twice as high if the potential return destination is the area where the person grew up (which we call his or her "origin"), where the migrant presumably has more location-specific capital than in other areas. These patterns are illustrated in Fig. 1.

We have also explored the differential influence of various explanatory variables on the likelihood of primary, return, and nonreturn repeat migration. The main findings include the following:

- The empirical results are consistent with the hypothesis that persons who are least capable of processing information effectively—the less educated, the less skilled, the less careful planners—are the most prone to return. Presumably, their initial moves, often pressured by an unemployment problem, were more likely to have been based on limited or faulty information. Only those migrants who return promptly, however, conform to this "failure" stereotype of a person whose initial move was disappointing. Those who delay moving back more closely resemble migrants generally and, like them, are by and large positively selected (e.g., they are better educated than others at risk). Nonreturn repeat migrants are also positively selected, especially those who move again the year following their initial move.

- Location-specific capital (especially home-ownership) in the new location deters subsequent migration, whereas location-specific capital in an old location tends to draw the previous resident back.

- Of people who made two or three previous moves, those whose last move was a return are less likely to migrate again than are people who were not returning. Return migrants appear to conform to the proverb, "Once burned, twice cautious."

- One of the most noteworthy findings is a consistently they define return migration as a move between years $t - 5$ and $t$ that is back to one's place of birth.
positive relationship between the distance of the initial move and the likelihood that it is followed by a return or onward move. This finding is consistent with a theoretical framework that views distance as a proxy for information costs, and posits that the poorer the information on which the initial move was based, the less likely it is to prove successful, and the more likely a subsequent "corrective" move will be required. This exception to the usual "distance-decay" relationship, which we find with individual-level data on moves measured over a one-year migration interval back to a place lived in one to six years previously, has also been found by DaVanzo (1976a) and Yezer and Thurston (1976) for aggregate Census data which measure returns as moves in a five-year interval back to one's birthplace. Furthermore, we find with our annual microdata that the likelihood of an onward move is even more positively related to the distance of the initial move than is the likelihood of a return move, a relationship also found by Allen (1979) in his analysis of aggregate Census data.

Except for the positive distance effect, which we find for all types of repeat migration (return or onward, interval of absence \( IA=1 \) or \( IA=2-6 \)), the person who returns after a one-year interval of absence emerges as the main exception to the migration relationships most typically encountered; other return and onward migrants appear not only to conform, but to be "doubly selected." This latter result is consistent with recent studies of U.S. Census data (e.g., Miller, 1977). However, the negative selectivity of short-interval return migration has gone undisclosed in Census data because self-cancelling short-interval returns are disproportionately underdetected in Census data.
VI. RECOMMENDATIONS FOR SURVEY DESIGNERS*

Sections III through V have summarized and critiqued the approaches used in a number of recent economic studies of migration decisionmaking. Because many of the shortcomings of these studies stemmed directly from data limitations, I have compiled the following list of recommendations for researchers who are designing surveys to collect data for estimating microeconomic models of migration.

Survey Design:

1. Where possible use a longitudinal survey design. Longitudinal data enable measurement of the migration determinants before a move is made. Interviews at one- or two-year intervals should suffice. In follow-up interviews, retrospective questions can be asked about moves within these migration intervals. Every effort should be made to follow individuals who move. **

To Measure Migration:

2. Collect information on areas of residence at as fine a level of geographic detail as possible. If the areas are too small, they can always be aggregated, but if they are too large and heterogeneous (e.g., a few regions of a country), it is usually difficult or impossible to break them down later on.

*The discussion in this section implicitly assumes a national probability sample. See Hugo (1979) for a discussion of alternative sampling strategies.

**If one is merely interested in explaining the causes of migration, not the consequences, it is necessary to obtain information only on where the migrant went; this information may be available from friends, neighbors, or relatives left behind. However, losing such individuals from the sample for subsequent interviews will result in a sample bias and will preclude constructing explanatory variables that use information on actual post-migration experiences to estimate expected income returns to migration.
3. Whether longitudinal or retrospective, questionnaires should ask about all moves, not merely residences at specific time points. The latter type of question will miss intervening moves; if it must be used, the points in time considered should be as close together as possible (e.g., one year apart). Special care should be taken to collect information on seasonal migration, circulation, and commuting in settings where they are common.

4. Obtain a retrospective residence history wherever possible. Previous migration is closely interrelated with subsequent migration. Moreover, migration relationships depend on the type of move (primary, return, or other repeat). Accordingly, it is important to know something of the individual's past history, in order to see current or forthcoming moves in the context of larger migration sequences. Also, with retrospective data, determinants of those past moves can be investigated (see 11 below), increasing the number of moves that can be analyzed (and perhaps reducing the number of respondents).

To Delimit Working Samples:

5. Collect sufficiently detailed information on personal characteristics so that the sample can be delimitated as desired. For example, you may wish to exclude retired people, students, and military personnel from the empirical analysis.

To Construct Variables to Explain Who Moves and Why (See also 13):

6. Income data should include separate data on income components likely to have different effects on migration: wage income, returns on location-specific capital, and other income.

7. Information should be also collected on types of location-specific capital that yield nonmonetary returns, e.g., friends and relatives. If at all possible, this information
should be collected not only for the current location but also for potential future destinations. Friends and relatives living in other places are both a source of information about those places and a potential source of support to the migrant during a transition phase. Thus, they may be a powerful influence on destination choice.

8. In settings where a nonnegligible fraction of some families' income is not received in monetary form, data should be collected on in-kind receipts, consumption of own production, and family members' time use so that we can better characterize the families' "real incomes."

9. Wages and other income data should be collected for reference periods that permit the analyst to distinguish what was received before moving from what was received afterward. It may be best to key wage and income questions to moves.

10. Information on moves and personal characteristics that affect migration decisions should be collected for all members of a family. Since members of a family usually move together, the characteristics of other family members may affect the migration decision of a particular member. Data should be collected on actual and expected remittances, since these may be important considerations in migration decisions.

11. Where retrospective migration data are being collected, retrospective information on migration determinants should be gathered also. Retrospective data are cheaper to collect than longitudinal data (which require follow-up interviews) and provide the necessary information sooner (one does not have to wait for the follow-up interview). However, past moves should be explained by the characteristics of the individual at the time of those moves, not at the time of the survey. Characteristics at the time of the survey may have been affected by those previous moves. A life history matrix (see Hugo 1979, pp. 14-16) is
a good way to collect this retrospective information.

To Explain Destination Choice (See also 7):

12. If you wish to explain how migrants choose among destinations, information should be collected on personal characteristics that might affect the individual's wage or probability of finding a job in other places. Wage and probability-of-employment equations can then be estimated for different areas (if sample size permits) and used to impute what the potential migrant might experience if he or she moved there.

13. Try to collect community-level data on the characteristics of the origin and potential destination areas. Availability of social services, labor market characteristics, cost of living, and other areal characteristics may affect the attractiveness of the current location and of potential destinations.* Considering community characteristics is especially important because they represent factors amenable to influence through national or local policy initiatives. Indeed, any policy that does not affect all geographic areas of a country equally will cause those areas to have varying attractiveness to potential migrants. For example, by including the appropriate community level variables, we could determine whether the structural factors emphasized in the Todaro (1969) and Harris-Todaro (1970) models, e.g., union power and civil service regulations, appear to influence migrants' choice of destination. In my opinion, the neglect of areal characteristics is one of the most serious shortcomings of empirical analysis of migration using microdata. It is ironic that in the move from aggregate to microdata, we have gone from an over-emphasis on areal characteristics to an under-emphasis.**

*For more on this topic, see Gardner (1979).

**For an example of a survey that satisfies most of these 13 requirements, see our Malaysian Family Life Survey, especially questionnaires MF2, MF3, and MF10 (Butz and DaVanzo, 1978, and Butz et al., 1978.)
It may be extremely expensive, of course, or even impossible, to implement all of these suggestions in a single survey, especially a small-scale one. It will be useful, however, to consider the above checklist carefully in light of one's plans; it may be possible to pursue some of the suggestions at little added cost. At the very least, a review of the checklist may enable the survey designer to avoid some of the pitfalls that have marred past studies.
BIBLIOGRAPHY


Repeat Migration in the U.S.: Who Moves Back and Who Moves On?
The Rand Corporation, P-5961, June 1978(a).

"Does Unemployment Affect Migration?--Evidence from Micro Data," 

"Repeat Migration, Information Costs, and Location-Specific Capital," Population and Environment (forthcoming).


