The Urban Growth Question

PHILIP E GRAVES, University of Colorado at Boulder

Available at: https://works.bepress.com/philip_graves/80/
Rural to Urban Migration:
Population Distribution Patterns

Philip E. Graves
Associate Professor
Department of Economics
University of Colorado

Marion Clawson
Consultant in Residence
Resources for the Future, Inc.
Washington, D.C.
In preparing this paper we have been greatly assisted by Donald Haurin, who compiled an initial list of references, and by George Tolley, Dean Jansma, Pat Madden, Lee Martin, and Clark Edwards, not only for their helpful suggestions as we began our task, but also for an extensive initial list of references. Errors of interpretation are our sole responsibility.

P. E. G.

M. C.
Why do people live, work, and play where they do? How much, why, and between what points do people migrate, both temporarily and permanently? These are rather simple questions, but the answers are much less clear and do not fall neatly into one field of human knowledge. Because choices are involved, choices having costs and benefits, the matter of migration and settlement patterns involves economics and the economist. But the sociologist, the demographer, the urban or regional planner, the geographer, and perhaps other specialists are, and traditionally have been, also involved. The field of migration shades off into many other considerations and disciplines. If one interprets the subject broadly, there has been and continues to be a vast literature dealing with migration. This is not a full review of all related literature but represents the authors' ideas of the major divisions of the general subject, with many illustrative references, from which the interested person may go further.

Some of the professional writing is primarily analytical, being either theoretical or empirical. It seeks to present or develop a theoretical or analytical framework or to provide some empirical data, reaching conclusions of a more or less quantitative nature. Other literature deals primarily with a problem or a situation for the purpose of developing a policy or a program to solve that problem or to improve the situation. There is obviously not a clean line between these two major categories: some pieces may be primarily analytical but lead up to an evaluation of an ongoing program, for instance. In the present survey, we are primarily interested in the analytical pieces.
The first section presents a background discussion of migration and settlement, designed to provide a historical context for the discussion of analytical trends. The next section introduces two migration "triads" as a conceptually useful device for understanding the diverse analytical and empirical threads weaving in and out of the migration literature. A prediction of the direction that will be taken in future migration work emerges. The final section summarizes the review of population distribution patterns.

*Migration and Settlement Design*

Migration has built and rebuilt the United States, from the days of the first permanent settlements of Europeans along the Atlantic Coast until today; for many decades people moved westward across the nation, and moved from rural to urban locations in great numbers, more or less continuously. Rural to urban net migration now appears to be largely over—in recent years nonmetropolitan counties have been growing more rapidly than their more urban counterparts.

Migration of human beings is a complex personal, economic, social, demographic, psychological, and political process. The personal characteristics of individuals greatly influence their decision to move or to stay, and affect the nature of the migration process when they decide to move. But the migration process in turn affects individuals, both those who move and those who stay; migrants, in particular, may experience substantial personal change as a result of moving. Loss of some members of the local society has an impact on the area of origin in numerous and varying ways; gains of migrants similarly alter the areas of destination in equally varied ways.

Migration flows may usefully be classified in several ways. There is gross flow in one direction: total numbers of persons leaving a particular area, or total numbers entering another area, during some defined time period. There is generally a reverse flow: for instance, even in the years of heavy migration from some rural areas to some urban area, there were considerable numbers of persons moving in the opposite direction. Some of these may well be described as return migrants; in an earlier period they moved from rural to urban areas, and now they return to their former location. There could of course be reverse flow without return flow. The difference between gross and reverse flows is a net flow; in many instances, one knows only about these net changes. The Census, for instance, may show numbers of people who live in a city, county, or state and whose birthplace was elsewhere; changes in such numbers between two Census dates show net migration for a given area during the Census interval. In more recent years, the Census has obtained information on numbers of persons living in each location who moved there during the preceding five
years. This measure is one of gross migration, since migrants in the opposite direction are not deducted.

A number of recent studies have dealt with individual survey data. These newer sources of data are leading to important new theoretical and econometric approaches to migration, as will become clear in the following section.

Migration flows may also be classified according to the intent of the mover and/or the time period involved in the migration. For instance, some migration is seasonal, with the intent being to return to the place of origin within the year: primitive people followed their herds of livestock; men left the village to do seasonal work elsewhere while the families stayed at home; and in the modern world both rich and retired persons follow the sun in winter. Some persons migrate on a tentative, or trial, basis—to see how they like life in the big city, while retaining the old home in the village as a line of retreat; others migrate with every expectation of never returning; and still others have intermediate intentions. The reality may differ, in either direction, from the expectation. What began as a temporary move may prove permanent, what was intended as permanent may be reversed, and so on. Data on migration, collected annually, as for instance the U.S. Department of Agriculture (USDA) series on persons leaving and persons returning to farms, may reveal rather different numbers than those collected at five- or ten-year intervals, as with Census data. Population flows in the United States are numerous and complex, and to some extent the relationships depend on how and when the data are collected.

Finally, migration affects both the area of origin and the area of destination. Not only does it remove some people, but it takes away part of the labor force, some of the customers, and some of the participants in every economic and social process. Migrants typically share certain personal characteristics, as will be discussed briefly later; hence both the area of origin and the area of destination are changed by migration—that is, both areas are not the same as would be the case if the migrants were a random cross-section of the whole population. It is possible for the same migration stream to impoverish or to enrich both the area of origin and the area of destination. For instance, if the better-educated and relatively higher income blacks move from southern rural areas—as they have—this decreases the average level of education and income among the blacks who remain, but these migrants may be less well educated and poorer than the average person in the areas of their destination, hence lowering the average level of education and income there as well. This particular type of migration may be highly desirable from a national viewpoint, but the policy implications in a regional context are substantial. Similar origin-destination effects are of course present for variables other than income flows, notably stocks of physical, financial, and human capital.
The foregoing general ideas have been developed, elaborated, or illustrated in many books and articles, of which only a few can be briefly mentioned. Long ago Shanahan [1927] discussed these ideas in rather general but perceptive terms, using Britain and western Europe for many of his examples but applying them to North America also. He concluded, pessimistically, that the outlook for prosperity in farming areas was bleak, because outmigration will always lag, owing to personal resistances to movement, below the level necessary to equate rural and urban real income. In more recent times, Sjaastad [1962] considered human capital aspects of migration, in an improved theoretical and empirical analysis, though he concentrated more upon the economics of population movement than upon its sociological or human terms and did not consider the effects of the migration either on areas of origin or on areas of destination. Sjaastad also included a fairly complete listing of the pertinent economics literature, up to the time of his writing. Park [1966], considering depressed areas generally and not specifically agricultural areas (though these are also included), found that outmigration may take the most employable persons from the local labor force, because they are best employable elsewhere, leaving the source area even less capable of competing economically. Vanderkamp [1970, 1971] conducted both theoretical and empirical studies for the Canadian experience. He found that migration was greater in times of prosperity than in times of high unemployment. Moreover, when migration takes out of an area people who were unemployed but who drew welfare payments from a central government, the income of the area is reduced and its unemployment is increased (but to only about 40 percent of the gain due to outmigration of the unemployed).

There has been an extended literature on the amount, the flows (direction, end point, etc.), the timing, the human characteristics, and so on, of migration in the United States, especially since World War II. Only three such studies will be referenced in this section as indicative of the type. Bowman [1965] made an extended analysis of migration from the South to other parts of the country, with particular reference to the age, educational level, and incomes of those who migrated and those who stayed, and with special concern for racial differences. The period of the analysis was the 1950s, and the prime source of data was the 1960 Census. She concluded that migration tends to take the better educated, especially among blacks, but it also takes some of the most poorly educated; average incomes rise with education, until at the highest educational levels southern whites did about as well as similarly educated northerners, but incomes of southern blacks were substantially below those of northern blacks at each education level. She stressed the very low level of education characterizing the South for both blacks and whites, circa 1960.
Fein [1965] presented an analysis of migration in the 1955-60 period, by Census regions within the South. His analysis was much influenced by the migration pattern in Florida, which had a large immigration, dominated by whites, older persons, the very well educated, and those in higher status occupations. Largely because of this heavy immigration to Florida, net migration from the South as a whole was rather low—heavy outmovements elsewhere more than overbalanced the heavy inmovement to Florida. Although gross rates of outmovement were higher for whites than for blacks, the inmovement was largely white, so the rate of net outmovement was higher for blacks. Fein estimated the dollar values lost from the South by migration and concluded that it was not a major part of the South’s total economic problem. Of note is that this study observed a multiplicity of migration motives which figured importantly in recent analyses; although many were moving out of the South seeking higher incomes and less discrimination, others were moving into the South motivated by low price levels and climatic amenities.

Using Census data, Ashby [1964] made an easily understood analysis of one aspect of migration during the 1940s and 1950s—the shifts in employment. He utilized the shift-share type of analysis employed by many economists, showing which states gained by having a mix of employment favorable in terms of rapid nationwide growth. Although this type of analysis deals with migration only peripherally, it does help both to explain why some migration took place and to measure some of the results of that which did occur.

There have been numerous analyses of the farm-urban migration in the United States, a rural to urban movement which characterized the years since World War II and which began tapering off only recently (Beale [1975]). Beale [1971] made many analyses of the demographic aspects of this earlier movement—the ages of persons involved, their educational status, the effect of loss of young people on reproductive rates in the areas of origin, and other aspects.

Diehl [1966], using data for economic areas within the Southeast, concluded that migration rate and farm income are negatively correlated—where farm income is higher, the outmigration rate is lower. Chennareddy and Jones [1972] estimated the expected remaining lifetime earnings of young men about to enter farming and of older farmers if both groups were to migrate away from farms for city employment; this analysis lends support to the rationale by which young men hesitate to enter farming.

Hathaway [1960] wrote an excellent review of the outmigration from agriculture as it appeared at that time.

Price et al. [1969], in an exhaustive review of the rural-urban migration literature to the time of his writing, noted that many stereotypes of the
historical rural to urban migrant were not supported by the data. In particular he observed that
1. most rural-urban migrants were white;
2. recent black migrants were as well educated as urban blacks;
3. relatively small proportions of the migrants were on welfare;
4. migrants earned about the same average incomes as urban nonmigrants.

The Price study dealt in chapter-by-chapter annotated detail with the traditional migration issues: the areas left behind, the decision to migrate, characteristics of rural-urban migrants compared with those of nonmigrants in urban areas, adjustment of migrants in urban areas, return migration to rural areas, effects of migration on rural areas, and effects of migration on urban areas. Readers interested in one or more of these topics would do well to begin their study with the Price synthesis of research findings.

The Price work was concerned not only with rural to urban migration but also with poverty. The latter aspect is pursued here only to the extent that poverty and migration are related. For further discussion of rural poverty, see the Bryant, Rawden, and Saupé survey in this volume.

Much recent work in migration (Graves [1976], Graves [1979b], Liu [1975]) has supplemented traditional (and ongoing) concerns of employment and economic opportunity with the concerns of the environment and amenities generally. This reemphasis no doubt stems from the current revival of growth in nonmetropolitan areas, a revival well summarized by Beale:

The vast rural-to-urban migration of people that was the common pattern of U.S. population movement in the decades after World War II has been halted and, on balance, even reversed. During 1970-73, nonmetropolitan areas gained 4.2 percent in population compared to only 2.9 percent for metro areas. In the eyes of many Americans, the appeal of major urban areas has diminished and the attractiveness of rural and small town communities has increased, economically and otherwise. The result is a new trend that is already having an impact, one that modifies much we have taken for granted about population distribution [1975, p. 1].

In the following section an organizing construct, called here the “Two Triads,” will be presented to categorize the modeling approaches observed in the literature. For the present, it seems clear that the dominant type of migration actually observed has had an (expected) profound influence on the analytical models advanced for the explanation of the migration phenomenon.

When one comes to consider settlement pattern, the first and basic question is: on what geographic scale? An analysis might be made on a national scale; this will show that the American population until recently has been concentrating in metropolitan areas and in certain regions—Florida, Texas, and the
Pacific Southwest, in particular. Or one might make a study on a regional basis, either using regions (rather than metropolitan areas) as units within a national total or studying changes within parts of a region. Likewise, analysis might be based upon states, or upon (Census-defined) economic areas within states, or by counties. Or the focus might be upon rural-urban contrasts, or upon nonmetropolitan-metropolitan contrasts, or upon those for central city and suburb. At the central city metropolitan scale, population has been dispersing since World War II; suburbs have grown much faster than have central areas. Or one might focus upon settlement pattern within neighborhoods and suburbs—lot size, contiguity or lack thereof for subdivision development, floor/area ratio, and other measures. Each of these scales of analysis is not only defensible but valuable for certain purposes. Difficulties arise when one uses a single scale (often because data are available for it and not for another) but applies one's findings as if they related to a different scale. For example, when one uses Census data on "urban counties," one gets one picture of urbanization in the "megalopolis" that stretches from Boston to Washington, whereas one gets a wholly different picture if one uses aerial photographs that show actual development on the land (see Clawson [1971]).

The study of settlement patterns—in the broadest sense of that term—may be concerned with one or more of several aspects. One concern may be efficiency—but whose efficiency and for what activities? There is a fairly clear concern with efficiency in production by producing units—farms, factories, stores, etc.; there is also clearly efficiency in terms of workers—the relation of their residences to their employment; that is efficiency for the housewife or other shopper; there is efficiency in recreation—the location of recreation areas vis-à-vis location of homes; and so on. But there is also the possibility that what is efficient for the person observed or the participant is inefficient when all those affected are considered. We are aware of this possibility because in recent years we have discovered that the underlying externalities are so pervasive. For example, an individual may personally find driving his or her car to work the most efficient means of transport, but when everyone drives, this may no longer be as efficient for the whole body of workers, in view of the congestion and pollution which results.

But settlement patterns may be examined from viewpoints other than efficiency, at least as the economist defines the latter. Planners, whether city, regional, physical, "comprehensive," or some other variety, architects, and others have been much concerned with "design"—a range of considerations and a viewpoint, even a jargon, with which the average economist is entirely unfamiliar.

Economists have conducted many studies of one or another aspect of the American settlement pattern, yet have made few studies examining the settle-
ment pattern as a whole, explicitly, in terms of efficiency, externalities, or other characteristics. There have been, for instance, a number of studies concerned with growth centers, marketplaces, and trade areas, commuting patterns, and the like (Berry, Goheen, and Goldstein [1969], Fox [1968], Hansen [1971]. In these and similar studies by others, central points or cities are identified in various ways; trade or tributary area is defined by actual commuting or trade patterns or in terms of distance to the center as compared with distance to equivalent centers elsewhere, and worker commuting patterns, trade and commodity flows, cultural influences, or other economic or social relations between center and hinterland are measured or described.

All of this may be considered a modern and more sophisticated, or at least more specialized, version of the classic locational theorizing of Von Thünen [1826], Lösch [1954], and A. Weber [1929 (1909)]. That classical tradition has been continued by Hoover [1937, 1948], Isard [1956], Lefebre [1958], and Beckmann [1968]. Neutze [1967] has provided a beautifully clear discussion of location theory, using agricultural and trade examples.

A somewhat different line, or perhaps only a specialized version of general location theory, has concerned itself with the relationship of city center to its periphery—changes in land use intensity gradients, and the like. Clark [1967] wrote extensively on this subject, using data from various cities around the world. In this country, Muth [1961, 1969] made similar analyses. Winship, as a sociologist, treated the same situation using terms that many economists will find strange. Each of these authors used measures of intensity of land use obtained by dividing total population (or total economic activity) by total land area; there was no place in any of these formulations for idle land, which some of us have identified as one-half or more of the area "withdrawn" by the city from other land uses; each utilized a single-centered model of the city, which seems increasingly inappropriate in this day of decentralized metropolitan areas; and each made additional simplifying assumptions. Nevertheless, these approaches clearly included several positive features.

As might be expected when spatial arrangements are under review, geographers have given the subject considerable attention. One recent issue of Economic Geography (J. A. Brown and Moore, eds. [1971]) is devoted entirely to the subject of urban spatial systems, with eight major articles; some of these are primarily mathematical and theoretical, others represent major reviews of pertinent literature, and some provide substantial empirical findings. Although the focus is on urban spatial patterns, the agricultural economist will find much of interest in this symposium issue.

Numerous attempts have been made (U.S. Advisory Commission [1968]) to determine economies of scale for cities or metropolitan areas: how do costs (total, or for governmental functions) or returns (by some
measure) or satisfactions vary with size of the urban settlement, and is there an optimum size? Tolley, Graves, and Gardner [1979] presented the most comprehensive treatment of issues pertaining to city size. However, it is extremely difficult to define, and more so to measure, “output” in unambiguous terms—if real governmental costs go up, this may mean either that services are better or that productivity has fallen. Data on “costs” even for governmental activities are seldom uniform; there are few satisfactory measures of quality of life, or even some segments of it; and factors other than size may dominate the results. There is a widely held, but not empirically sound, belief that economies of scale exist among smaller settlements (up to 50,000, or to 250,000, depending upon which writer you read) and an equally firmly held belief that diseconomies exist and can become large above some size (250,000 for some authors, 500,000 or more for others); and most of those holding either or both of these beliefs argue strongly that a population “balance” is needed, by which they mean fewer people in large metropolitan areas and more people in smaller cities or open countryside.

Attempts to measure economies of specific aspects of urban life have been somewhat more decisive. Hirsch [1968], for example, measured the factors affecting costs of garbage collection. Downing [1969a, 1969b] showed the relationship of suburban settlement pattern and location to costs of sewage collection, transport, and treatment. In Clawson [1971] all studies of this type were utilized—they are not numerous.

Two other types of studies relating to efficiency of settlement pattern may be mentioned briefly. The relationships between costs of travel from home to recreation area, and the number of persons making such trips, have now been analyzed in perhaps 100 different studies. With outdoor recreation now an “industry” with a gross value of “output” of the same general magnitude as that for agriculture, and with the travel peaks for recreation often exceeding those for the journey to work, it is obvious that studies of the location of recreational opportunity vis-à-vis the location of the home of the recreationist are likely to increase in number and in sophistication.

The possibility of studying the efficiency of rural—largely farm—settlement patterns had been suggested, but few such studies have been conducted thus far (see Clawson and Knetsch [1966] for references). It is difficult to believe that the rural settlement pattern that evolved when travel speed along rural roads behind a team of horses was four miles an hour and when average farm size was about 130 acres (as it was for several decades) can also be the most desirable settlement pattern now, when travel speed in rural areas is 50 miles an hour and when farms are three or more times as large. What is needed is some research to test the economic and social efficiency of settlement patterns not yet in existence—a research for the future, not one of the past or present.
This brief discussion of both migration patterns and settlement patterns suggests that they are much more closely related than has been emphasized in the past. The following section presents an organizational framework within which they can be jointly understood.

The Two Migration Triads

The large and growing literature on issues surrounding the human migration decision, as reflected in this and other recent literature reviews (Greenwood [1975], Price et al. [1969], Mangalam and Morgan [1968]), covers many hundreds of articles. Yet only bits and pieces of a comprehensive model exist to provide a context for judging the relative merits of individual contributions and for determining how they fit together as a coherent whole. In presenting here what we call the "two migration triads," it is hoped that a conceptual foundation useful to researchers in the area will emerge.

The approach takes as its starting point a very nonrestrictive model of human migration. That is, all one needs to assume is that:
1. people want to locate so as to make themselves as "well off" as possible;
2. people can order their preferences pairwise between locations at a point in time, given what information they possess;
3. if location A is preferred to location B and location B is preferred to location C, location A will be preferred to location C.

An individual following these axioms is said to be "rational" (given what he or she knows of the characteristics of the various locations of the world). The easily derived and intuitively plausible optimality condition resulting from such a model is that a rational individual will relocate any time the present value of benefits from that relocation exceed the associated costs.

The Two Triads

The two migration triads that follow from this approach are depicted in Figure 1. Consider first Triad I—a perfect information, frictionless world in which the characteristics of all possible locations are assumed to be known for each period of the individual's life. The only uncertainties involved are personal (e.g., unexpected death of a spouse, job loss as one's firm goes out of business, and so on). Aside from such unexpected changes, the rational individual should be able to schedule his or her location throughout life. Thus, in this scenario, a series of what may be described as "life-cycle" changes is what leads to migration. Examples of such move-causing changes might be graduation from college (the best job is unlikely to be in the same location as the educational facility); getting married (larger home or apartment demanded), having children (still larger home, perhaps with more land), retirement (greater spare time might result in demands for better weather in which to enjoy it).
In all these cases a change (expected) occurs as a result of which the present value of benefits from migration exceed the costs.

It should be noted that the "present value" part of the previous sentence is important apart from the obvious discounting of more distant benefits. If an old individual faces the same change a young individual faces, the former would be less likely to move for two reasons: the benefits of the new location would be received for fewer periods since the older individual may not expect to live as long, and the costs of movement will be higher owing to loss of job seniority, greater importance of friends and other community ties, and so on. Thus we would expect, and indeed the literature substantiates, that the probability of moving should be, ceteris paribus, negatively related to age.

Another type of movement that would occur in a world of perfect information would be due to changes in the amounts of location-specific goods available. That is, whereas a certain professor might plan on living in the immediate area of his university, changing levels of, say, crime or pollution could well cause a move. Thus changes in supply of location-specific characteristics as well as changes in demand can lead to benefits from relocation exceeding costs.

These illustrations of movement in the presence of high levels of information and mobility suggest that equilibrium models are appropriate. That is, migrants represent the arbitrage activity that keeps utility levels constant across locations. This view of the migration phenomenon has an important
implication, already intimated, which is useful in interpreting existing empirical migration analyses: only changes in variables important in the migration decision should matter, with migration being unaffected by the levels of these variables if an equilibrium view is appropriate. For example, an individual having a certain allergy or sinus problem may, *ceteris paribus*, locate in Arizona, whereas another otherwise identical individual not having this problem might locate in Indiana. If national statistics on the number of people having such difficulties were included in a migration equation among regions or states, the expected coefficient of this variable would be zero in an equilibrium context. Only changes in the level of this variable will result in migration. Although this example may seem trivial, the interpretation of the income and unemployment variables that appear in virtually all net or gross migration studies is seriously affected. If one accepts the equilibrium modeling approach, all expected income differentials across locations represent compensation required to offset amenity and other location-specific differences so that utility is constant across locations. Hence a high urban median income reflects compensation for the lower level of amenities received there as compared with a rural setting, and these income differentials should not affect average migration flows.

This equilibrium modeling approach appears to be the conceptual view of the world that underlies recent emphasis on the effect of amenities on migration (Liu [1975], Graves [1976, 1979a, 1979b, 1980]). In these articles the implicit (not explicit) rationale for migrating to places with high amenity levels becomes, under Triad I, that amenities have high income elasticities of demand. Thus rising incomes over a period of time lead to changed demands for amenities that can only be exercised by moving.

It is readily seen that the approaches to migration represented by the first triad fall squarely into the usual microeconomic framework. That is, perfect mobility and information are incorporated in an equilibrium model. Further, migration may be thought of as resulting from changed demands for goods whose nature is location-specific. Hence the determinants of migration are changes in variables entering the demand or supply schedules for these location-specific goods, in much the same way as in ordinary goods markets. Viewed this way, it is clear that both net and gross migration regressions found in the literature represent reduced forms lacking any direct structural interpretation without additional information.

The second triad in the taxonomy proposed here is probably more important from the standpoint of sweeping regional or world migrations. In this scenario, with imperfect information and mobility, the determining factors in migration are the ever-changing perceptions of level of utility obtained in different locations. These changing perceptions may be of two basic types: "discoveries" of
new information about areas, and slow information transfer and reaction to available knowledge.

So elaborating on the first type: the sporadic discovery of new information would alter the relative attractiveness of areas in such a way that the benefits from moving might suddenly exceed the costs for many people. An example in which a whole new area could be compared with existing feasible locations would be the discovery of the New World. Similarly, although California was known to exist long before the discovery of gold, this new information about California led to rapid immigration as the perceived benefits of moving there were suddenly seen to exceed the costs of that move for many rational individuals. The long history of labor-saving technological advance in agriculture can be seen, in the context of the second triad, to have facilitated the mass (but selective) migration from rural to urban areas prevalent until recently in the United States. This effect, when combined with the low-income elasticity of demand for farm (relative to manufactured) products, has led to real utility differences which have only recently begun to be arbitrated away.

Slow information transfer and irregular disturbances leaving real utility differentials account not only for rural-urban migration but also for a great deal of the migration from the East to the West which has been observed over most of the history of the United States. The notion here is that the West might have spectacular scenery, warmer weather, lower humidity, etc., which make it a more attractive location for many people—but unlike the approach of Triad I, this relative attractiveness is only perceived slowly. Thus, even if hypothetically one-half of the population of the East at some point in time were predicted to have greater benefits than costs associated with moving westward, the slow recognition of differentials in locational advantage can account for the long historical process we observe. An individual might move, some years later a friend from the East may visit, who might then after a lag move, and so on. Movement to equilibrium in such a system may well take many decades.

This view of the migration decision suggests disequilibrium model building, which incorporates lags into the analysis and interprets variables in a different way than was done in Triad I. In particular, the income and unemployment differentials are not viewed as compensatorial but as representing real utility differences which should be expected to result in migration flows tending to eliminate them. This is the view most commonly implied in studies of rural to urban migration, particularly until recently.

The job search literature, associated most prominently with labor economists (Phelps et al. [1970]), best represents the emphasis of the second triad. However, labor movement away from locations dominated by extractive
industries (until recently) and away from agricultural areas, the subject here, has been extensively analyzed in the migration literature and is certainly well characterized as a persistent disequilibrium phenomenon. In other discussions, return migration is commonly explained as being due to imperfect information having resulted in a "wrong" initial move.

A Synthesis

The two triads presented here clearly dovetail and jointly underlie the observed migration phenomenon, as implied in Figure 1. Although examples of both predominantly Triad II-motivated migration (e.g., ongoing black migration to the North since the Civil War) and predominantly Triad I-motivated migration exist, in general the motivations are mixed. Hence, while incomes may compensate for amenities at a point in time, the lower nominal wages in the high amenity areas will lure mobile industry in the longer run, raising the demand for labor, which will in turn facilitate additional immigration to desirable areas (which may also become less desirable through this process, as exemplified by Los Angeles smog). Further, since industries and commerce vary in the degree to which they are footloose (location-specific production function shifting variables, such as harbors, are of varying importance according to firm type), this process may be expected to take a very long time. To illustrate: the movement of the steel industry from first Pittsburgh to Chicago, now Chicago to Houston (and points west), depends on prior and concurrent movement by metal fabricators.

The simultaneity of the two triads in most observed migration suggests the direction future developments of the field will take: supply-of-population equations, which will depend on production function shifters (harbors, local wage rates, agglomeration benefits of city size, technological change in agriculture, etc.) and final market demand will be derived, based to a far greater extent than in the past on individual rather than aggregate (usually Census) data. The simultaneous system may then be solved and migration can be predicted as the flow necessary to maintain supply-demand equilibrium in the presence of predicted exogenous shifts in supply (owing to, say, projected income increases or air quality changes resulting from environmental legislation) or in demand (e.g., changes in highway, or harbor, or other infrastructure-altering accessibility to inputs or markets). If reactions were instantaneous (the Triad I case), this would be the whole story. But considerations of the Triad II kind will be incorporated through lags in movement to equilibrium, so that several periods are required to eliminate excess population supplies or demands.

The procedure, outlined here, through which the two triads are combined will take some time for data development, proper specification, identification, choice of functional form, and the like. However, the process of moving in
this direction will be easier if the two-triad formulation presented here is borne in mind when drawing on the existing literature for guidance in formulating the simultaneous system described above.

The importance of a common theoretical framework in which to view migration is heightened by the range of social science disciplines that consider migration, as is seen in the accompanying references. The variables that shift demand and supply and that introduce lags vary considerably across disciplines. That is, a sociologist may stress group pressures (as, e.g., change in racial or religious persecution), whereas an economist might emphasize income and unemployment differentials, and a psychologist might study age effects, and so on. But the fundamental model must inevitably be that:

1. an individual has an initial location;
2. if the individual is in his or her best location, only changes in variables that affect the costs and benefits of relocation will cause migration (Triad I);
3. if the location is not optimal, after perhaps long lags, movement will occur to the optimal location as information is processed and acted upon (Triad II).

**Summary**

As indicated above, the individual who contemplates the possibility of moving from one location to another weighs advantages and disadvantages of present location, of new possible locations, and both monetary and nonmonetary costs of moving. Each of these factors has an uncertainty or risk component; the situation at home today may be better known than the situation in the possible destination location, but what changes in each will bring? Much of the literature speaks of push from areas of origin and pull to areas of destination; but there is always both push and pull from each in varying proportions. That is, the young man contemplating a move may be repelled by the lack of a job at home, but attracted by friends and familiar surroundings there; attracted by a job and the glamour of city life, but repelled by his conception of the impersonality and coldness of the big city. In a given situation, push may dominate—an area experiencing a massive drought or a major flood or a prolonged but local economic depression may literally drive some people out. In other situations, pull may dominate, as when World War II opened up jobs in shipyards for blacks who had never had well-paying job opportunities previously. But neither force is ever entirely absent.

In evaluating the many empirical studies of migration, the question inevitably arises: “What do these coefficients mean?” The two triads presented in the previous section were advanced as an aid in this interpretation. Yet the pervasive intermingling of equilibrium and disequilibrium facets of actual migration leaves the appropriate interpretations fuzzy at best. Further, conceptualization of how the triads may be combined in the future leads to the realization
that most past empirical work lacks behavioral content—the regressions represent reduced forms, neither demand nor supply. This is not to suggest that such studies are not useful in the policy arena or as input into fuller models. Rather, the implication is that such usage is fraught with very real difficulties which should be recognized.

The post-1970 reverse migration to nonmetropolitan areas perhaps best clarifies the fundamental similarity of all migration, as described in the previous section. However, the text and the accompanying references emphasize the rural to urban migration the United States has witnessed since World War II. The recent stress on amenities (climate, pollution, general "quality of life" indicators, and the like) is not really new and was observed by Dahlke and Stonecipher [1946] at the beginning of the period of concern here. It would seem, though, that such concerns will become increasingly important in the future.

With a continuation of migration highly probable and with major, if not profound, changes under way in settlement patterns, metropolitan and nonmetropolitan alike, there would seem to be many opportunities for the young economist seeking fruitful fields for research. It is customary, in a review of past research, to plead for more future research, and we have no wish to break with that tradition. More research is clearly needed, and we hope this paper may have suggested some of the directions in which the payoff is likely to be greatest.

Notes

\[1\] For a good analytical treatment of population movements and metropolitan growth, including trenchant comments on data sources and discrepancies, see U.S. Advisory Commission on Intergovernmental Relations [1968, chap. 1]. See also Clawson [1971], especially chaps. 2 and 3.

\[2\] See Clawson and Kretzsch [1966] for a discussion of this methodology and for a review of such studies up to 1966. For more recent studies, see issues of the Journal of Leisure Research.

\[3\] This section is a revised version of Graves [1977].

\[4\] See Linneman and Graves [1977] for comparisons of city rankings according to nominal and real incomes where the real income measure corrects primarily for climate differences. Clearly other variables (culture, "excitement," and the like) do matter, but the seriousness of their omission for the conclusions is unclear since they may be uncorrelated with the weather variables. Particularly in the rural-urban migration case, the issue of job-specific human capital becomes relevant and can account for why established farmers do not migrate but their offspring do.

\[5\] Changed relative prices for amenities (as, for example, the lower price of summer comfort in the American Southwest owing to the advent of air conditioning) would also enter in this type of model.

\[6\] Promising recent efforts along these lines include those of Graves and Linneman [1979], Bishop [1976], and Polshek and Horvath [1977]. These authors applied
nonlinear estimation techniques (probit or logit) to the probability of individual or household movement in the face of exogenous changes. In a further effort to separate the amenity and job-search motivations for migration, Linnehan and Graves [1979] employed a multinomial logit analysis in which the dependent variable classified households into those changing jobs but not moving and those moving but not changing jobs as well as those doing both.

Greenwood’s [1975] now-classic review considers a broader range of migration research, giving little specific attention to rural-urban migration.

References


Part 4. East South Central States.
Part 5. West South Central States.
Part 6. Western States.


RURAL TO URBAN MIGRATION


