UC Berkeley
IURD Working Paper Series

Title
A Theory of Movements: Introduction

Permalink
https://escholarship.org/uc/item/2n9386pp

Author
Alonso, William

Publication Date
1976-06-01

eScholarship.org Powered by the California Digital Library
University of California

A THEORY OF MOVEMENTS: (I) INTRODUCTION

William Alonso

June 1976

Working Paper No. 266

This work is being carried out under grant #SOC74-24115 from the National Science Foundation, through the Institute of Urban and Regional Development, University of California, Berkeley. My special thanks go to the many people who have listened to me with patience over the past few years. Lyna Rossi deserves special mention for her help.

INTRODUCTION

This paper is the first of a series which will present a more general theory of movement systems. In the past few decades a great many models of movement have been developed in the social sciences, including input-output models of the economy, models of urban traffic, of migration, of social mobility, of the changes in status of employees in organizations, and many more. The theory presented in these papers offers a common logical and mathematical framework, of which various particular models are special cases.

It will be important to keep in mind what it is the theory does and what it does not do. It does not speak to the empirical correctness of particular models, and it does not offer, in itself, recommended alternatives. It does, I hope, offer a more general framework for thinking about these diverse models, a common notation, and thus a way of comparing one model to another in their logic and suppositions, many of which are usually implicit and unrecognized. Thus, divergent empirical findings and theoretical conclusions among models may stem not from the perversity of nature, but rather from the various portraits of nature which differ in the implicit logic of particular models at their inception. This more general framework may help to trace unrecognized implications of the forms of particular models and thus sometimes help to open new ideas or to discard old ones.

At this writing I am still developing and interpreting this more general theory, and I expect to develop it more fully in later papers, together with a reinterpretation of particular models current in social science literature. In this paper I present a capsule version of the theory itself.

The theory is quite brief in its formal expression, but it is highly involuted, in that many variables are implicit functions of each other. This makes certain questions difficult to analyze. Further, the general theory is highly circular. Certain suppositions and relations may be used as a point of departure in the construction of certain models, while other relations are derived or implied from these first ones. In other models, on the other hand, first and derived relations may be quite the other way around. But the circularity and tautological nature of the theory is its strength: wherever particular models hop onto the circle, if logical consistency is applied, it will carry them all the way around. The exposition of the theory which follows, therefore, is only one way of making this circle.

THE THEORY IN BRIEF*

Movements Originating from i

There is a number n of classes of origins and destinations. These classes may be regions, populations, occupations, and so forth according to the particular model. Movements out of class i depend

*Earlier versions of this theory were presented in two of my papers: "National Interregional Demographic Accounts: A Prototype," Monograph #17, Berkeley: Institute of Urban and Regional Development, University of California, February 1973; "Policy-Oriented Interregional Demographic Accounting and a Generalization of Population Flow Models," WP-247, Berkeley: Institute of Urban and Regional Development, University of California, March 1975. The notation used there was somewhat different, and the analysis less advanced.

on certain characteristics of this class and its population, and on the pull or draw of the system upon the class and its population. We may write this as

(1) $$M_{ix} = v_i D_i^{\alpha_i}$$

where M_{ij} = movement from class i to class j;

$M_{ix} = \Sigma_j M_{ij}$, total movements from i;

v_i = a function of characteristics of class i and/or its population;

D_i = the draw or pull of the system at i, a function of the relation of class i to the rest of the system, per unit of v_i;

α_i = rate (elasticity) or movement response from i to its relation to other classes in the system (D_i).

Several observations are in order:

(a) What v_i may be varies widely from model to model. In some it is an extensive measure, such as population; in others, it is an intensive measure such as an unemployment rate; in others it is a value derived from complex relations of attributes of the class (such as climate in a region) and of its elements (such as age, sex, education of its population); finally, in others, such as input-output models, v_i may simply be zero or unity as a dimensionless variable indicating the existence of this class as a possible origin.

(b) Some models consider only moves out of the class, while others allow for "moves" within the class; i.e. staying as a form of moving. The practical difference is that in the former M_{ii} is set equal to zero. This difference among models does not matter for our purposes at this time, as it only involves whether a transitional variable t_{ii} (to be introduced below) is set equal to zero or some other constant value. In a later paper I will analyze the rather straightforward consequences of these different types of models.

(c) D_i, the relation of i to the system, remains formally undefined for the moment; it will be derived below. Its intuitive interpretation, however, is the pull which the rest of the system (including or not i = j) exerts for movements from i. D_i are the attractions of the system, the opportunities available per unit of v_i. It may be thought of as a demand or a draw, hence the notation D.

(d) Whereas v_i, however constructed, is the result of the characteristics of class i, D_j is a function of the entire system, a scalar systemic variable evaluated at i. By analogy, D_i is the local value of a magnetic field. I will show later that, although this variable is seldom explicitly set forth, it is always imbedded in the various models by logical necessity.

(e) The responsiveness of moves from i to the attractions or draw D_i is measured by the elasticity α_i. In some models, the number of moves originating in i are unaffected by the attractions D_i. In such models $\alpha_i = 0$, and departures M_{ix} are a function only of local attributes v_i; however, as I have said, the variable D_i will of necessity exist in other relations of the system. In other models, the number of moves is fully proportional to D_i, and in these $\alpha_i = 1$. Although I have not encountered in the literature any models where α_i is greater than unity or smaller than zero, these are logically possible. An instance of α_i being negative might occur when increasing draws upon elements of a class result in greater class cohesiveness to combat such temptations. An instance where α_i is greater than unity is "gold rush" situations. On the other hand, while α_i values intermediate between zero and unity do not appear explicitly in the literature, they make a great deal of sense: the response of moves to attractions may exist but be less than fully proportional. Therefore I shall pay considerable attention to such intermediate values.

<u>Movements Arriving at j</u>

The moves arriving at a particular class j may be similarly set as

(2) $$M_{xj} = w_j C_j^{\beta_j}$$

where M_{ij} = movement from class i to class j;

$M_{xj} = \Sigma_i M_{ij}$, total arrivals at j;

w_j = a function of the characteristics of class j and/or its population;

C_j = competition, crowding, congestion or potential pool of moves into j per unit of w_j; a function of the relation of j to rest of the system, per unit of w_j;

β_j = rate of expansion (elasticity) or arrivals to competition at j.

Again, some remarks are in order:

(a) What w_j may be varies widely among models, from extensive, to intensive, to mixed, to existence. Intuitively, w_j may be thought of as the attractiveness or pull of j. In some models w_j is equal to v_j, but more commonly they are different functions of different variables.

(b) Some models allow i = j in M_{ij}, and some do not. This does not affect our general discussion.

(c) C_j is the relation of j to the rest of the system, and remains formally undefined for the moment. Its intuitive interpretation in the case of migration models, for instance, is a weighted measure of the pool of migrants available to j per unit of w_j. It may be thought of as a measure of crowding, competition, or congestion, hence the notation C.

(d) As with v_i and D_i, w_j is a function of characteristics evaluated totally within class j, whereas C_j is a systemic variable, a function of the whole system evaluated at j.

(e) In some models the number of moves into j is determined exclusively by w_j. In such models the potential pool of moves does not matter and $\beta_j = 0$. In other models, it is fully proportional, and $\beta_j = 1$. For the same reasons as for α_i I shall concentrate on the range $0 \leq \beta_j \leq 1$ in later papers.

Movements from i to j

If we think of the total arrivals and departures to and from each class as the marginals of an n x n matrix, it becomes clear that there is no unique way of filling the cells for class to class movements. The sum of rows and columns consistent with these marginals provides 2n equations, while there are n^2 unknowns. Therefore the particular form of the M_{ij} relation cannot be derived from equations (1) and (2), and some further suppositions are needed.

A possible set of further suppositions is that the share of departures from i going to j will be (a) proportional to the attractiveness of j (w_j); (b) proportional to the probability of a potential arrival's entry into j, to enjoy its attractiveness ($C_j^{\beta_j - 1}$); (c) proportional to any special relation that may obtain between i and j, such as ease of movement or special affinity (t_{ij}); and (d) inversely proportional to the total opportunities or alternative attractions available to a departure from i (D_i^{-1}).

In formal terms, this amounts to:

(3) $$\frac{M_{ij}}{M_{ix}} = w_j C_j^{\beta_j - 1} t_{ij} D_i^{-1}$$

where $\frac{M_{ij}}{M_{ix}}$ = share of i's departures going to j;

w_j = a value representing the attractiveness of j determined by the characteristics of this class or its population;

$C_j^{\beta_j - 1} = C_j^{\beta_j}/C_j$ = entry rate or probability of a potential arrival's entry into j; this is the ratio of actual to potential arrivals at j; by equation (1), $C_j^{\beta_j}$ are the actual arrivals per unit of w_j, and C_j is the pool of moves available to j per unit of w_j;

t_{ij} = a special relation obtaining between class i and class j, such as ease of movement in a traffic or migration model, a transitional probability in a Markov model, or a technical coefficient in an input-output model;

D_i = total draw, attractions or opportunities offered by the system to a departure from i.

C_j: Competition, Congestion, Potential Pool of Moves

By combining equations (1) and (3), and rearranging terms, we obtain the equations for the number of moves from i to j.

(4) $$M_{ij} = v_i D_i^{\alpha_i - 1} \; w_j C_j^{\beta_j - 1} \; t_{ij}.$$

By summing equation (4) over i, we obtain

(5) $$M_{xj} = \Sigma_i M_{ij} = w_j C_j^{\beta_j - 1} \; (\Sigma_i v_i D_i^{\alpha_i - 1} \; t_{ij}),$$

Combining equations (2) and (5), we have

$$M_{xj} = w_j C_j^{\beta_j} = w_j C_j^{\beta_j - 1} \; (\Sigma_i v_i D_i^{\alpha_i - 1} \; t_{ij}),$$

which simplifies to

(6) $$C_j = \Sigma_i v_i D_i^{\alpha_i - 1} \; t_{ij}.$$

We have thus arrived at an operational definition of C_j, derived from equations (1), (2), and (3).

Its interpretation may be visualized more easily if we rewrite equation (6) by using equation (1) as

$$C_j = \Sigma_i (M_{ix}/D_i) t_{ij}.$$

C_j, the pool of potential arrivals per unit of w_j, is equal to the sum over all possible origins of the total departures from each origin (M_{ix}), relative to the total opportunities or draws open in the system to each mover from each origin (D_i), and weighted in each case by the affinity or ease of transition between each origin and destination j (t_{ij}). In other words, potential arrivals C_j per unit of attraction at j are simply the system's departures per opportunity, weighted from j's perspective.

D_i: Demand, Draw, or Opportunities

We can derive D_i similarly by summing equation (4) over all destinations j,

$$\Sigma_j M_{ij} = v_i D_i^{\alpha_i - 1} \Sigma_j (w_j C_j^{\beta_j - 1} t_{ij}).$$

Combining it with equation (1) $M_{ix} = v_i D_i^{\alpha_i}$, we obtain

$$M_{ix} = v_i D_i^{\alpha_i} = v_i D_i^{\alpha_i - 1} \Sigma(w_j C_j^{\beta_j - 1} t_{ij})$$

which reduces to

$$(7) \qquad D_i = \Sigma_j w_j C_j^{\beta_j - 1} t_{ij}.$$

The total opportunities, draws, or attractions from the perspective of a unit of v_i are all of the attractions w_j, weighted by the entry rate ($C_j^{\beta_j}/C_j$) of actual to potential arrivals per unit of w_j and by the ease of movements from i to each j (t_{ij}).

In addition, we could write

$$D_i = \Sigma_j (M_{xj}/C_j) t_{ij}$$

This may be interpreted as stating that the system's opportunities available to a mover from i consist of the actual entries (M_{xj}) at each destination j, negatively weighted by the degree of competition (C_j) per unit of w_j (e.g., jobs), weighted by the ease of access (t_{ij}). In other words, opportunities per prospective mover from i are the system's successful entries per try, weighted by ease of access, from i's perspective.

Total Moves

It can be readily seen that total moves in the system are

$$\Sigma_i \Sigma_j M_{ij} = \Sigma_i v_i D_i^{\alpha_i} = \Sigma_j w_j C_j^{\beta_j}.$$

This simply says that in the matrix of moves the sum of all the cells is equal to the sum of the row marginals and to the sum of the column marginals. Or most simply, that total departures equal total arrivals in a closed system.

The Full Set of Relations and Alternative Derivations

Equations (1), (2), (4), (6), and (7) describe the fundamental relations of this theory of movements. This full set of relations is redundant, as we have seen by deriving equations (6) and (7) from the primary relations (1), (2), and (4). We could equally have picked another set of relations as primary and derived the remainder.

For instance, I may start by postulating that the moves from i to j depend on the local characteristics at i (v_i) and at j (w_j), and to their degree of connectedness (t_{ij}). I also postulate that the moves from i to j are proportional in some degree to the alternative demands on moves from i, where D_i are these demands and Ψ_i is the degree of response or elasticity, resulting in $D_i^{\Psi_i}$. Similarly I postulate that the moves will be proportional to the competition or congestion at j, C_j, modified by some elasticity or rate of adjustment η_j of j's capacity to this congestion, resulting in $C_j^{\eta_j}$. From these postulations I obtain the equivalent of equation (4):

$$M_{ij} = v_i D_i^{\Psi_i} w_j C_j^{\eta_j} t_{ij}.$$

A prospective mover from i views the universe of his opportunities (D_i) as the attractions or characteristics of each destination (w_j), weighted by their accessibility, (t_{ij}), by their congestion and by their elasticity of adjustment ($C_j^{\eta_j}$). This results in an equation equivalent to (7):

$$D_i = \Sigma_j w_j C_j^{\eta_j} t_{ij}.$$

Similarly from a destination j, the level of congestion (C_j) may be viewed by the aggregate of the characteristics of the sources of moves (v_i), weighted by their accessibility (t_{ij}) and by their response to opportunities ($D_i^{\Psi_i}$). This provides an equation equivalent to (6):

$$C_j = \Sigma_i v_i D_i^{\Psi_i} t_{ij}.$$

Now we may derive the equivalents of equations (1) and (2) for M_{ix} and M_{xj}. Summing the equation for M_{ij} over j we obtain

$$\Sigma_j M_{ij} = M_{ix} = v_i D_i^{\Psi_i} (\Sigma_j w_j C_j^{\eta_j} t_{ij}).$$

And since the expression within the parenthesis is the definition of D_i, we can say

$$M_{ix} = v_i D_i^{\Psi_i + 1}$$

Similarly, by summing over the i's we obtain

$$\Sigma_i M_{ij} = M_{xj} = w_j C_j^{\eta_j} (\Sigma_i v_i D_i^{\Psi_i} t_{ij})$$

and since the expression within the parenthesis is C_j, we have that

$$M_{xj} = w_j C_j^{\eta_j + 1} .$$

Thus we have all five relations again, having derived the equivalent of equations (1) and (2) from the equivalent of equations (4), (6), and (7). The only difference is that here we have exponents Ψ_i and η_j instead of α_i and β_j. These translate readily as

$$\Psi_i = \alpha_i - 1$$
$$\eta_j = \beta_j - 1.$$

However, the difference is more than one of notation and it is instructive to examine it in greater detail. In this last derivation Ψ_i was called the rate of response or elasticity of an origin from the point of view of a prospective destination; η_j was characterized as the rate of response or elasticity of a destination from the point of view of a particular origin. In fact, these rates of response are the result of two distinct effects. Thus, an increase in available opportunities will elicit more moves from a source if α_i is greater than zero. At the same time, this increase in alternative opportunities will reduce the share

$$\frac{M_{ij}}{M_{ix}} = w_j C_j^{\beta_j - 1} t_{ij} D_i^{-1} = w_j C_j^{\beta_j - 1} t_{ij} / \Sigma_j w_j C_j^{\beta_j - 1} t_{ij}$$

that goes to an invariant destination unless the elasticity of outward moves is equal to or greater than unity (i.e., $\alpha_i \geq 1$). Thus the total effect $(D_i^{\Psi_i})$ in place to place flows (M_{ij}) is derived from both the effects on total flows $(D_i^{\alpha_i})$ and the effect on a particular flow (D_i^{-1}). The equivalent analysis shows that $C_j^{\eta_j}$ is the composite of local expansion $(C_j^{\beta_j})$ and reduced shares from invariant sources (C_j^{-1}). For this reason the use of α_i and β_j is preferable to that of Ψ_i and η_j.

There are obviously many ways of arriving at this redundant set of consistent relations. From the point of view of elegance, a parsimonious set would depend on which are the initial and which are the derived relations. Similarly, in applied problems, the count of unknowns and equations will depend on the data and the questions at hand. Various models and applied theories enter this logical circle from different angles, and seldom close it. Similarly, explanation, simulation, and optimization models will start differently and go after different things.

OBSERVATIONS AND CONCLUSIONS

The Inescapability of C and D, and the Values of α_i and β_j

The most interesting feature of the theory consists of the systemic terms C and D. Most models appear to function without them,* and these terms may seem an unnecessary complication. But they are a logical necessity in all flow models. To illustrate, for instance, if an economic study forecasts a labor force w_j at some location without concerning itself with the potential supply of labor C_j, it is in effect setting $\beta_j = 0$ and $M_{xj} = w_j C_j^0$ (allowing for $i = j$). Virtually no such economic base models concern themselves explicitly with migratory patterns, but it follows of necessity that they imply a set of gross migratory flows as in equation (4), where C_j will appear with an exponent $\beta_j-1 = -1$. Conversely, migration flow models, usually econometric in their approach, usually estimate some equivalent of equation (4) as $M_{ij} = v_i w_j t_{ij}$, without concerning themselves with systemic aspects. But if the logical chain is followed to the implied total arrivals and departures at each locality, it becomes clear that they imply that $M_{ix} = v_i D_i^1$ and $M_{xj} = w_j C_j^1$, and that the complete specification of the gross flows is $M_{ij} = v_i w_j t_{ij} D_i^0 C_j^0$.**

---

*Alan Wilson's "family" of spatial interaction models (see A.G. Wilson [1970] Entropy in Urban and Regional Modelling, Pion, London) incorporate systemic variables comparable to D and C (A^{-1} and B^{-1} in his notation). His consideration of "unconstrained," "production-constrained," "attraction-constrained," and "production-attraction-constrained" models correspond implicitly to α and β values of (1,1), (0,1),(1,0), and (0,0) respectively. He also derives these from entropy maximization on flows subject to diverse constraints on total cost, departures and arrivals. In a later paper I will discuss the concept of entropy as behavior or as a statistical construct.

**These arguments will be presented in detail in later papers dealing with the translation of various models to the form of the theory.

In short, the systemic variables C and D remind one of economic rent: wherever it is suppressed in one form, it pops up in another. If the systemic variables are ignored in the cells of the matrix (M_{ij}), they will appear in the marginals, (M_{ix}, M_{xj}) and vice versa. In other words, like Moliere's bon homme, these models have been speaking prose without knowing it.

Indeed, the neglect of these systemic effects has led to a situation where, in all the models which have come to my attention, α_i and β_j have values of zero or one, according to whether the systemic variables were implicitly assigned zero exponents in the cells or in the marginals. Once the systemic variables and their exponents are explicitly recognized, interesting empirical and theoretical questions arise as to the values of α_i and β_j. Unit elasticity and total inelasticity are strong assumptions. It seems likely that values intermediate between zero and unity will be frequent in reality, and there may be circumstances in which they assume values outside this range.

Conclusions

As stated in these brief pages, the theory is quite general, thus abstract and somewhat slippery. Important and unresolved questions remain as to its generality. For instance, the use of power functions and multiplicative relations may seem somewhat arbitrary, and other functional forms may suggest themselves. But a surprising number of models, including input-output, Markov processes, gravity and entropy models, economic base models, and others, are special cases of this more general formulation, and this is encouraging as to its generality.

A particular feature of the mathematical structure of this theory also suggests its generality. The theory is one of flow models, and these are sometimes called "accounting" models, as in the case of input-output and demographic accounts. They are so called because every move must be accounted for at its origin and destination. When they are displayed as a matrix, the row of cells adds up to total departures and the column to total arrivals. Because investigators wish to be able to derive expressions for row and column sums, they have often supposed that the expressions for cells must be linear since it is easy to sum linear functions and to express these sums compactly if there are common variables in the cells. Conversely, it is often believed that the addition of non-linear forms cannot be reduced to compact and manageable forms. But the mathematical form of this theory has the property that cells in which the variables are multiplicative, and raised to powers, do add by rows and columns to forms that reduce to very economical and interpretable expressions. This is the consequence of the systemic variables, which in effect doubly normalize the cells.

It seems to me unlikely that there can be many functional forms which permit such parsimonious and interpretable expressions for the summation of complex non-linear equations. But I must admit to uncertainty on three counts. The first is that indeed there may be many other forms which also reduce elegantly and meaningfully. If so, the theory presented here is less general. The second source of uncertainty is whether this theory encompasses purely linear models of forms such as $M_{ij} = v_i + w_j$. Is there some transform to make this a special case of the expressions I have been discussing?

The third question has to do with some developments in the field of human migration. As I have reviewed elsewhere,* a debate raged as to the effect of push factors in economically distressed regions. Some researchers applying regression analysis to forms of $M_{ij} = v_i D_j^0 w_j C_j^0 t_{ij}$ found no push. Others, performing regressions on forms of $M_{ix} = v_i D_i^0$ claimed, but not convincingly, to find push factors. Neither side seemed aware of the basic difference in their models, namely that the first group implicitly set $\alpha_i = \beta_j = 1$, while the second group implicitly set $\alpha_i = 0$, $\beta_j = 1$. The matter was considerably advanced by Michael Greenwood,** who estimated simultaneous regressions for structural equations equivalent to $M_{ix} = v_i D_i^0$ and $M_{xi} = w_i C_i^0$, and obtained coefficients consistent with theory and common sense. My purpose is not to rehearse that debate once again, but to raise the question of whether the theory represented here is consistent with simultaneous equations approaches to empirical estimation. It seems to me it is, since one can easily go on to specify the implicit missing relations in Greenwood's model ($M_{ij} = v_i D_i^{-1} w_j C_j^{-1} t_{ij}$; $D_i = \Sigma_j w_j C_j^{-1} t_{ij}$; $C_j = \Sigma_i v_i D_i^{-1} t_{ij}$), but I am not sure. On the other hand, it has seemed to me for some time that the set of equations of this theory have the look of the solution to some set of simultaneous equations, but I have been unable to make progress in this regard.

In later papers I will take up various extensions of the theory. These will include the analysis and comparison of many models

*See my paper, "National Interregional Demographic Accounts: A Prototype," Monograph #17, Berkeley: Institute of Urban and Regional Development, University of California, February 1973.

**M.J. Greenwood, "Research on Internal Migration in the United States: A Survey," Journal of Economic Literature, June 1975, 8 (2), pp. 397-433.

expressed in the more general notation, analysis of the dimensionality of the theory and certain problems of aggregation, consideration of stable state equilibria, interpretation of the model in terms of economic concepts, and some other matters. As will become evident, many questions are still unresolved. I wish I could solve them and present a polished product, but it becomes increasingly clear that I cannot. Therefore I am choosing to present these ideas as they are, trying to set forth clearly the matters I have thought through and trying to pose candidly and usefully questions I have been unable to resolve. Many of these, I feel quite certain, will be easily tackled by the many people whose mathematical armament is more extensive and less rusty than my own.

In summary, my purpose in exposing this unfinished theoretical edifice is two-fold. In the first instance, I hope that the greater generality of this theory will further understanding of implicit logical structure of movement systems and thereby improve their appropriateness. The second purpose is to catch the interest of a few students, particularly those with mathematical strength, so that they can help to complete the edifice -- or to redesign it or even knock it down if it is unsound.