AN EVALUABLE THEORY FOR A CLASS OF MIGRATION PROBLEMS

G. Haag* and W. Weidlich*

October 1983
CP-83-58

* Institut für Theoretische Physik
der Universität Stuttgart,
Pfaffenwaldring 57/III,
D-7000 Stuttgart 80, FRG

Contributions to the Metropolitan Study: 4

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

LIST OF CONTRIBUTIONS TO THE METROPOLITAN STUDY

1. Anas, A., and L.S. Duann (1983) Dynamic Forecasting of Travel Demand. Collaborative Paper, CP-83-45, International Institute for Applied Systems Analysis, Laxenburg, Austria.

2. Casti, J. (1983) Emergent Novelty, Dynamical Systems and the Modeling of Spatial Processes. Research Report, RR-83-27, International Institute for Applied Systems Analysis, Laxenburg, Austria.

3. Lesse, P.F. (1983) The Statistical Dynamics of Socio-Economic Systems. Collaborative Paper CP-83-51, International Institute for Applied Systems Analysis, Laxenburg, Austria.

4. Haag, G., and W. Weidlich (1983) An Evaluable Theory for a Class of Migration Problems. Collaborative Paper, International Institute for Applied Systems Analysis, Laxenburg, Austria.

FOREWORD

Contributions to the Metropolitan Study:4

The Project "Nested Dynamics of Metropolitan Processes and Policies" was initiated by the Regional & Urban Development Group in 1982, and the work on this collaborative study started in 1983. The series of contributions to the study is a means of conveying information between the collaborators in the network of the project.

This paper by Haag and Weidlich presents an approach to modeling migration (population or household relocation). It makes a clear distinction between (i) the decision process or, in other words, the underlying motivation for migration, and (ii) the macrolevel outcome which in model terms is obtained as an aggregate picture of a dynamic stochastic process. The latter is formulated in terms of transition probabilities which are functions of trend parameters which may be related to characteristics of the housing market, transportation system, workplace accessibility of different locations, etc.

It is observed that empirically estimated trend parameters may be compatible with more than one type of microlevel decision process. In this respect the analysis makes one central theoretical issue in the Metropolitan Study apparent: the resolution of microlevel assumptions and macrolevel descriptions of a dynamic process.

Börje Johansson
Acting Leader
Regional & Urban Development Group

November, 1983

ABSTRACT

A master equation formulation for a class of migration problems describing the spatio-temporal dynamics of a system of regions is introduced. The transition probabilities are functions of trend parameters, which characterize preferences, growth pool and saturation effects. The trend parameters can be determined by regression analysis from the empirical migration matrix. The solution of meanvalue equations yields a nonlinear migration prognosis. The relation between trend parameters and motivation factors, e.g., income per capita, infrastructure and transportation costs, is also discussed. Numerical simulations illustrate the influence of the superposition of migration trends on the evolution of the system.

CONTENTS

1. INTRODUCTION

Migration processes are an example of socio-economic dynamics of particular interest for quantitative research because of the following reasons:

On the one hand the *underlying motivations* for a given kind of migration are relatively well defined and specific and thus available for inquiry.

On the other hand all these motivations must always result in *a clear decision* to maintain or to change the location of the unit under consideration in a given interval of time. The number of relocations of a group of units between a set of sites or areas can be counted. And the understanding of the dynamics of these changes is the objective of a quantitative migration theory.

The motivation structure behind migration patterns has been intensely investigated in recent work on the *microlevel* as well as on the *macrolevel:* factors like the housing market, neighborhood quality, distance from working place and transportation costs, preferences for an urban or rural life style, the labor market and the structure of the economy of an area have been considered (see Clark and Smith, 1982; Clark and Burt, 1980; Curry, 1981; Dendrinos and Mullally, 1981). The effect of motivation factors on the dynamics of a migration system was taken into account by different versions of *utility functions* (see for example, Clark and Smith, 1982; Leonardi, 1983). Furthermore, several approaches towards a *general dynamical theory of migration* have recently been developed (see Griffith, 1982; Haag and Dendrinos, 1983; MacKinnon, 1970; Sonis, 1981; Weidlich and Haag, 1980, 1983; and in particular the articles in Griffith and Lea [eds], 1983). The migration model presented here follows and extends the line of argumentation in Weidlich and Haag (1983) and Griffith and Lea (1983, pp24-61).

We begin with the remark that the structure of the process to be described in a migration theory indicates some consequences with respect to the adequate form of the theory:

Firstly, since we cannot expect to describe the individual decisions on a fully deterministic level, their *probabilistic treatment* seems adequate. As a consequence the resulting theory should be a *stochastic* one. This means, we expect that the theory yields as a main result the evolution of a *probability distribution* over the possible configurations (i.e., area population numbers) arising in the migration process. From such a moving distribution it will then be possible to derive deterministic equations of motion for the meanvalues of the numbers of migrating units. Thus the stochastic level of description comprises the evolution with time of meanvalues of the relevant variables.

Secondly, it seems reasonable to make a certain separation between the *structure of motivations* and the *resulting dynamics* of a migration process in the following sense:

Let us assume that the migration dynamics is fully determined by a set of parameters $\{\tau_1,\dots,\tau_n\}$ appearing in the equations of motion for the probability distribution and the meanvalues. (In our model the parameters $\{\nu,\delta_i,\kappa_i\}$ determining the transition probabilities (2.7) are of this type.) And vice versa we assume, that the numerical values of the $\{\tau_1,\dots,\tau_n\}$ can be extracted from empirical knowledge about the migration process by a regression analysis. We shall denote this kind of parameter as "trend parameters".

Since the migration is ultimately generated by certain motivation factors $\mu_1,\ldots,\mu_N$ describing the intensities of different "reasons" 1,2,...,N to change the location, the trend parameters will be functions of these motivation factors μ_j. This functional relationship however is not necessarily an unambiguous one. It can happen, that two different sets $\{\mu_1^{(1)},\ldots,\mu_N^{(1)}\}$ and $\{\mu_1^{(2)},\ldots,\mu_M^{(2)}\}$ of motivation factors under appropriate assumptions give rise to the same set $\{\tau_1,\ldots,\tau_n\}$ of trend parameters and hence to the same migration dynamics. In this case the migration analysis does not distinguish between the motivation sets (1) and (2), since they lead to the same dynamics. The two motivation factor sets $\{\mu_1^{(1)},\ldots,\mu_N^{(1)}\}$ and $\{\mu_1^{(2)},\ldots,\mu_M^{(2)}\}$ then are equivalent with respect to the migration process.

For these reasons we proceed in two steps: firstly we introduce trend parameters, which determine the dynamics, and vice versa are determined by the dynamics of the system; and only secondly and separately we discuss the eventual dependence of these trend parameters on motivation factors.

The paper is organized as follows:

In chapter 2 the migration model is developed on the basis of its master equation and its meanvalue equations. The mathematical tools of the derivation of the master equation and some of its more specific properties, e.g., detailed balance, are summarized in the appendix A. In chapter 3 the trend parameters are determined from empirical migration data by a regression analysis; an evaluation scheme indicates the possible conclusions to be drawn from this analysis. The relationship between trend parameters and motivation factors is also established. Finally, the numerical simulation of chapter 4 shows the practicability of the model.

2. THE MIGRATION MODEL

In the following model for simplicity of notation we restrict ourselves to the migration of human populations, although the transition to a migration of other units can easily be made.

The individual motivations and resulting decisions in the migration process of populations are highly complex. Therefore a reasonable and practicable description of such decisions is formulated in probabilistic terms: for a member of a certain population there exists a certain probability per unit of time to move from one area or region to another. These transition probabilities are assumed to depend on certain trend parameters, whose numerical values fully determine the dynamics of the system. Before going into details, however, we have to draw a general conclusion: if the *individual* decisions are stochastic, the evolution of the *global* system composed of migrating individuals cannot be fully deterministic either. Instead the system must correctly be described by an equation of motion for the evolution of a *probability distribution* over its possible states. This equation is denoted as master equation and some of its general properties are summarized in appendix A. In the next sections the migration model is specified to which the master equation will be applied.

2.1 Specification of the Model

We consider one population of N members (the units of migration) migrating between L areas or sites 1,2,...,L. The possible states i of the migration system are then characterized by the *"socioconfiguration"* [1)]

$$i \rightarrow n = \{n_1, n_2, \ldots, n_L\} \tag{2.1}$$

with
$$\sum_{k=1}^{L} n_k = N \tag{2.2}$$

where the integer n_k is the number of units in area k. After this identification the formulas of the appendix can be applied. In particular, the probability distribution function

1) For the general definition of the socioconfiguration see Weidlich and Haag (1983).

$$P(n_1,\ldots,n_L\ ;\ t) \equiv P(n\ ;\ t) \tag{2.3}$$

can be introduced, where n abbreviates the vector $\{n_1,\ldots,n_L\}$ and where $P(n\ ;\ t)$ is the probability that the socioconfiguration $\{n_1,\ldots,n_L\}$ is realized at time t. The general form of the master equation then reads according to (A.9)

$$\frac{dP(n\ ;\ t)}{dt} = \sum_k w(n;n+k)P(n+k;t) - \sum_k w(n+k;n)P(n;t) \tag{2.4}$$

Here, $w(n+k;n)$ is the transition probability per unit of time from the socioconfiguration $n = \{n_1,\ldots,n_L\}$ to a neighboring configuration $n+k = \{n_1+k_1, n_2+k_2,\ldots,n_L+k_L\}$, where k_j are positive or negative integers. The sums on the r.h.s. of (2.4) extend over all k with nonvanishing $w(n;n+k)$ and $w(n+k;n)$, respectively.

2.2 Choice of Transition Probabilities

In order to make the model explicit the transition probabilities $w(n+k;n)$ which govern the dynamics of the system have to be specified.

We start from the individual transition probability p_{ji} per unit of time of one of the n_i members in area i to migrate into area j. In principle p_{ji} can be a function of the situation in all areas and in particular of the total socioconfiguration n, i.e., of the distribution of the whole population over the areas 1,2,...,L. It is highly plausible, however, to assume that p_{ji} is a function of the attractiveness of the new residence area j and the old area i only, leaving aside the situation in areas $l \neq i,j$.

Let the attraction of the area i be characterized by a function $f_i(n_i)$ depending on parameters specific to that area, which include the number n_i of its residents. Increasingly positive (negative) values of $f_i(n_i)$ by definition mean higher (lower) attractiveness to that area. Thus $f_i(n_i)$ may be interpreted as a *"utility function"* with respect to the area i and the migration type under consideration.

The simplest form of this attractivity or utility function is

$$f_i(n_i) = \delta_i + \kappa_i n_i + \rho_i n_i^2 \qquad (2.5)$$

where the parameters δ_i, κ_i, ρ_i can be interpreted as follows:

δ_i = *preference parameter* of area i, since growing δ_i increases the attractivity of area i

κ_i = *cooperation parameter* of area i, since a positive κ_i leads to increasing attractivity of area i with growing density ($\sim n_i$) of the population in i. Hence κ_i describes the growth pool effect.

ρ_i = *saturation parameter* of area i, since for $\rho_i < 0$ the attractivity decreases for sufficiently large population densities.

The problem is now to describe the effect of the utility functions on the migration dynamics. We solve this problem by proposing that the individual transition probabilities *directly depend* on the utility functions. (Then the utility functions directly influence the dynamics, and vice versa, the dynamical process itself measures the utility of the areas.) In this sense we choose the following exponential form of p_{ji}, which--beyond other advantages (see also Leonardi, 1983)--guarantees the positive definiteness of these transition probabilities:

$$p_{ji}(n_j, n_i) = \nu \exp [f_j(n_j+1)] \exp [-f_i(n_i)] \qquad (2.6)$$

with $\quad i,j = 1,2,\dots,L$

where ν is a global mobility parameter determining the time scale on which the migration process takes place. The form (2.6) obviously means, that the transition probability from area i to area j is larger than that from j to i, if

$$[f_j(n_j+1) - f_i(n_i)] > [f_i(n_i+1) - f_j(n_j)]$$

or equivalently, if the attractivity of j exceeds that of i.

Since the parameters $\tau = \{\nu, \delta_i, \kappa_i, \rho_i\}$, $i = 1,2,\dots,L$ *directly* influence the dynamics of the system they are trend parameters τ_ℓ in the sense of the definition given in the introduction.

It is now easy to construct the transition probabilities $w(n+k;n)$ between socioconfigurations. The n_i residents of area i change to area j with individual transition probabilities (2.6) and thus give rise to the configuration transition

$$\{n_i,\dots n_i\dots n_j\dots n_L\} \rightarrow \{n_1,\dots(n_i-1)\dots(n_j+1)\dots n_L\}.$$

Hence they contribute the term

$$w_{ji}(n+k;n) = \begin{cases} n_i p_{ji}(n_j,n_i) \equiv w_{ji}[n] \\ = \nu n_i \exp[f_j(n_j+1)] \exp[-f_i(n_i)] \\ \text{for } k = \{0,\dots(-1)_i\dots(+1)_j\dots 0\} \\ 0 \text{ for all other } k \end{cases} \tag{2.7}$$

to the transition probability $w(n+k;n)$. Since the transitions between all areas take place simultaneously and independently, the total transition probability $w(n+k;n)$ is the sum of all contributions (2.7) so that

$$w(n+k;n) = \sum_{ij=1}^{L} w_{ji}(n+k;n). \tag{2.8}$$

2.3 Explicit Form of the Master Equation

Inserting (2.8) with (2.7) into (2.4) the explicit master equation is obtained which can be cast into a more convenient form, because according to (2.7) only transitions between adjacent socioconfigurations are possible, which differ by $k_i = -1$, $k_j = +1$ at the sites i and j. This final form of the master equation reads

$$\frac{dP(n;t)}{dt} = \sum_{i,j=1}^{L} (E_i^{+1}E_j^{-1}-1)\{w_{ji}[n]P(n;t)\} \tag{2.9}$$

In formulating (2.9) we have used translation operators E_i acting on any function of $\{n_1,\ldots n_L\}$ as

$$E_i^{\pm 1} f(n_1,\ldots n_i,\ldots n_L) = f(n_1,\ldots (n_i \pm 1),\ldots n_L) \tag{2.10}$$

2.4 Detailed Balance

It can now be proved, that the transition probability (2.8) satisfies the condition of detailed balance (A.11) or, equivalently (A.13). For this aim we choose a set of smallest closed chains of states which, however, are sufficient for the proof, since arbitrary closed chains of states can be composed of these smallest ones. A smallest closed chain L_{ijk} connects the following states:

$$\begin{aligned} &\{..n_i..n_j..n_k..\} \equiv 0 \rightarrow \{..(n_i-1)..(n_j+1)..n_k..\} \equiv 1 \\ &\rightarrow \{..(n_i-1)..n_j..(n_k+1)..\} \equiv 2 \rightarrow \{..n_i..n_j..n_k..\} \equiv 3 = 0 \end{aligned} \tag{2.11}$$

and corresponds to the ring migration of a member of the population between areas i→j→k→i. Only one term of the r.h.s. of (2.8) contributes to each of the transitions (2.11). Hence, the formula (A.13) reduces to (read from right to left):

$$\begin{aligned} &\frac{w(0,2)}{w(2,0)} \cdot \frac{w(2,1)}{w(1,2)} \cdot \frac{w(1,0)}{w(0,1)} \\ &= \frac{w_{ik}[n_i-1,n_k+1]}{w_{ki}[n_k,n_i]} \cdot \frac{w_{kj}[n_k,n_j+1]}{w_{jk}[n_j,n_k+1]} \cdot \frac{w_{ji}[n_j,n_i]}{w_{ij}[n_i-1,n_j+1]} = 1 \end{aligned} \tag{2.12}$$

which easily can be checked to be satisfied. This completes the proof that detailed balance holds.

2.5 Stationary Solution of the Master Equation

The most important consequence of detailed balance is, that in the case of time-dependent transition probabilities, e.g., for constant trend parameters $v, \delta_i, \kappa_i, \rho_i$, the stationary solution $P_{st}(n)$ can be constructed by using (A.12).

Let us consider a chain of states

$$\{N,0,\dots 0\} \rightarrow \{N-1,1,0,..0\} \rightarrow \{N-2,2,0,..0\}$$

$$.. \rightarrow \{N-n_2,n_2,0,..0\} \rightarrow \{N-n_2-1,n_2,1,0,\dots 0\} \qquad (2.13)$$

$$.. \rightarrow \{N-n_2-n_3,n_2,n_3,0,..0\} \dots \rightarrow \{n_1,n_2,n_3,\dots n_L\}$$

where $\quad N = \sum_{i=1}^{L} n_i \qquad$ (see (2.2))

starting from the reference state $\{N,0,\dots 0\}$ and ending in the general state $\{n_1,n_2,..n_L\}$. They are connected by nonvanishing transition probabilities. Hence, we can use this chain to construct $P_{st}(n_1,\dots n_L)$ from $P_{st}(N,0,\dots 0)$ according to formula (A.12). By inserting into (A.12) the transition probabilities (2.8) with (2.7), which connect these states, we obtain for instance the intermediate result.

$$P_{st}(N-n_2,n_2,0,\dots 0) =$$
$$\frac{N(N-1)\dots(N-n_2+1)}{n_2!} \exp\{2\sum_{v=1}^{n_2} f_2(v) - 2\sum_{v=N-n_2+1}^{N} f_1(v)\} P_{st}(N,0,\dots 0) \qquad (2.14)$$

Continuing the procedure along the chain (2.13) the result

$$P_{st}(n_1,n_2,..n_L) = \frac{Z^{-1}\delta(\sum_{i=1}^{L} n_i - N)}{n_1!n_2!\dots n_L!} \exp\{2\sum_{i=1}^{L} F_i(n_i)\} \qquad (2.15)$$

can be derived. Here we have introduced

$$F_i(n_i) = \sum_{v=1}^{n_i} f_i(v) \;,\; n_i \geq 1$$
$$F_i(0) = 0 \qquad (2.16)$$

and put

$$P_{st}(N,0,\ldots 0) = \frac{\exp\{2F_1(N)\}}{N!} \cdot \frac{1}{Z} \tag{2.17}$$

Furthermore, the constraint (2.2) has been taken into account in (2.15) by the factor

$$\delta\left(\sum_{i=1}^{L} n_i - N\right) = \begin{cases} 1 \text{ for } \sum_{i=1}^{L} n_i = N \\ 0 \text{ for } \sum_{i=1}^{L} n_i \neq N \end{cases} \tag{2.18}$$

The factor Z follows from the normalization of the probability (2.15) and is

$$Z = \sum_{\{n\}} \frac{\exp\{2\sum_{i=1}^{L} F_i(n_i)\}\delta\left(\sum_{i=1}^{L} n_i - N\right)}{n_1!\, n_2! \ldots\ldots n_L!} \tag{2.19}$$

where the sum extends over all socioconfigurations $\{n\}$. In the case of the utility function (2.5) the exponential term of (2.15) can be further evaluated. The use of (2.5) in (2.16) yields

$$\begin{aligned} F_i(n_i) &= \delta_i \sum_{v=1}^{n_i} 1 + \kappa_i \sum_{v=1}^{n_i} v + \rho_i \sum_{v=1}^{n_i} v^2 \\ &= \delta_i n_i + \kappa_i \frac{n_i(n_i+1)}{2} + \rho_i \frac{n_i(n_i+1)(2n_i+1)}{6} \end{aligned} \tag{2.20}$$

which makes the stationary solution (2.15) of the master equation fully explicit. Evidently this solution factorizes into a product of L factors, where the i-th factor only depends on $\{n_i, \delta_i, \kappa_i, \rho_i\}$. The additional factor (2.18) however links these otherwise independent factors of the distribution. Henceforth we shall neglect the saturation term in (2.31) by putting $\rho_i = 0$.

Using Stirling's formula for the factorials, (2.15) can also be written in the form

$$\left.\begin{array}{l} P_{st}(n_1,n_2,\ldots n_L) = \dfrac{\delta\left(\sum\limits_{i=1}^{L} n_i - N\right)}{Z} \exp\left\{\sum\limits_{i=1}^{L} \phi_i(n_i)\right\} \\ \text{with} \\ \phi_i(n_i) = 2\delta_i n_i + \kappa_i n_i(n_i+1) - n_i(\ln(n_i)-1) \end{array}\right\} \quad (2.21)$$

The numbers $\hat{n}_j$ maximizing P_{st} are found by maximizing the exponent in (2.21) under the constraint (2.2). This leads to

$$\begin{aligned} &\delta\left\{\sum_{i=1}^{L} \phi_i(n_i) - \lambda\left(\sum_{i=1}^{L} n_i - N\right)\right\} \\ &= \sum_{i=1}^{L} \delta n_i \left[2\delta_i + 2\kappa_i(\hat{n}_i+\tfrac{1}{2}) - \ln(\hat{n}_i) - \lambda\right] = 0 \end{aligned} \quad (2.22)$$

Solving for $\hat{n}_i$ we obtain

$$\hat{n}_i = \frac{\exp\left[2\delta_i + 2\kappa_i(\hat{n}_i+\tfrac{1}{2})\right]}{\exp(\lambda)} = \frac{N\exp\left[2\delta_i + 2\kappa_i(\hat{n}_i+\tfrac{1}{2})\right]}{\sum\limits_{j=1}^{L} \exp\left[2\delta_j + 2\kappa_j(\hat{n}_j+\tfrac{1}{2})\right]} \quad (2.23)$$

where the Lagrangian parameter λ was determined by using the constraint (2.2). For given trend parameters $\{\delta_j, \kappa_j\}$ the transcendental equations (2.23) can be used to calculate the most probable stationary area population numbers $\hat{n}_i$, $i=1,2,..L$. The trivial case $\delta_i = \kappa_i = 0$ of equal and vanishing trend parameters naturally leads to equal area population numbers

$$\hat{n}_i(\delta_i = \kappa_i = 0) = \frac{N}{L} \quad (2.24)$$

The stationary distribution (2.15) or (2.21) will turn out to be a useful tool in the analysis of the migration process. By definition it is the time-independent solution of the master equation for *constant* trend parameters $\{\nu, \delta_i, \kappa_i\}$. Furthermore, any time dependent solution of this master equation approaches the stationary solution for $t \to \infty$. In general the migration system is, of course, not in this equilibrium state, firstly because for given constant trend parameters the occupation numbers n_i may not yet have reached their equilibrium values $\hat{n}_i$ and secondly because the trend parameters may also be slowly time dependent. Although the master equation (2.9) is still valid in the last case the transition probabilities $w_{ji}[n]$ are now time dependent (via the $\delta_i(t), \kappa_i(t)$, etc.) and the system does not reach a stationary state at all! Nevertheless, the formal "stationary" solution for the given momentary values of the trend parameters (into which the system would relax sooner or later if from now on the trend parameters would remain constant), provides an insight: comparing the equilibrium values $\hat{n}_i$ calculated from the "stationary" solution according to (2.23) with the empirical values n_i at that point in time, the "distance from equilibrium" of the system under the given trend situation can be estimated. This kind of analysis will be implemented in chapters 3 and 4.

2.6 Equations of Motion for the Meanvalues

The time dependent solutions of the master equation (2.9) can be found numerically. In most cases, however, the full information contained in the probability distribution is not exploited in an empirical analysis. Instead, it is sufficient to solve equations for the meanvalues $\bar{n}_i(t)$ of the population numbers n_i, $i=1,2,...L$. These meanvalue equations will now be derived from the master equation (2.9). The meanvalue of a function $f(n)$ of n is defined by

$$\overline{f(n)} = \sum_n f(n)P(n;t) \tag{2.25}$$

In particular the mean occupation numbers are given by

$$\bar{n}_k(t) = \sum_n n_k P(n;t) \tag{2.26}$$

The equation of motion follows from

$$\frac{d\bar{n}_k}{dt} = \sum_n n_k \frac{dP(n;t)}{dt}$$
$$= \sum_n \sum_{i,j=1}^{L} n_k \, (E_i^{+1}E_j^{-1}-1)w_{ji}[n]P(n;t) \tag{2.27}$$

where (2.9) has been inserted on the r.h.s. The r.h.s. can now be transformed making use of

$$\sum_n E_i^{+1}E_j^{-1}F(n) = \sum_n F(n) \tag{2.28}$$

for any function F(n) since the sum extends over all socio-configurations n, and of

$$n_k E_i^{+1}E_j^{-1}F(n) = E_i^{+1}E_j^{-1}(n_k-\delta_{ik}+\delta_{jk})F(n) \tag{2.29}$$

which follows from the definition of $E_i^{+1}E_j^{-1}$.

Taking into account (2.18) and (2.19) the r.h.s. of (2.27) is equal to

$$\sum_n \sum_{i,j=1}^{L} (E_i^{+1}E_j^{-1}-1)n_k w_{ji}[n]P(n;t)$$
$$+ \sum_n \sum_{i,j=1}^{L} E_i^{+1}E_j^{-1}(-\delta_{ik}+\delta_{jk})w_{ji}[n]P(n;t)$$
$$= \sum_n \sum_{i=1}^{L} w_{ki}[n]P(n;t) - \sum_n \sum_{j=1}^{L} w_{jk}[n]P(n;t) \tag{2.30}$$
$$\equiv \sum_{i=1}^{L} \overline{w_{ki}[n]} - \sum_{j=1}^{L} \overline{w_{jk}[n]}$$

so that the exact equations of motion for the meanvalues read

$$\frac{d\bar{n}_k}{dt} = \sum_{i=1}^{L} \overline{w_{ki}[n]} - \sum_{j=1}^{L} \overline{w_{jk}[n]} \tag{2.31}$$

for k = 1,2,....L

It is now assumed that the approximation

$$\overline{w_{ki}[n]} = w_{ki}[\bar{n}] \qquad (2.32)$$

is valid, which certainly holds for narrow unimodal probability distributions. The final set of equations (for k = 1,2,...L)

$$\frac{d\bar{n}_k}{dt} = \sum_{i=1}^{L} w_{ki}[\bar{n}] - \sum_{j=1}^{L} w_{jk}[\bar{n}] \qquad (2.33)$$

then is a self-contained set of coupled nonlinear differential equations in time for the occupation numbers $\bar{n}_k(t)$. The system becomes fully explicit by inserting the form (2.7) of $w_{ji}[n]$ whereby we obtain

$$\left.\begin{aligned} \frac{d\bar{n}_k}{dt} &= \sum_{i=1}^{L} \nu\bar{n}_i \exp\,[f_k(\bar{n}_k)]\,\exp\,[-f_i(\bar{n}_i)] \\ &\quad - \sum_{j=1}^{L} \nu\bar{n}_k \exp\,[f_j(\bar{n}_j)]\,\exp\,[-f_k(\bar{n}_k)]\,, \\ &\qquad k = 1,2,\ldots L \end{aligned}\right\} \qquad (2.34)$$

with $f_i(\bar{n}_i)$ according to (2.5). The stationary solution $\bar{n}_k^{st}$ of (2.34) can be read off immediately:

$$\bar{n}_k^{st} = C \exp\,[2f_k(\bar{n}_k^{st})], \quad k = 1,2,\ldots.L \qquad (2.35)$$

with the normalization factor

$$C = N\{\sum_{i=1}^{L} \exp\,[2f_i(\bar{n}_i^{st})]\}^{-1}. \qquad (2.36)$$

The transcendental equations (2.35) for the stationary meanvalues $\bar{n}_k^{st}$ essentially agree with the equations (2.23) for the most probable values $\hat{n}_k$ of the stationary probability distribution. This shows the consistency of the meanvalue approach with the fully stochastic approach. The meanvalue equations (2.34) with (2.5) are the main starting point for the empirical analysis of migration systems. The general procedure of this analysis will be introduced in the next chapter.

3. DETERMINATION OF TREND PARAMETERS FROM EMPIRICAL DATA

In principle the comparison between the theory of migration and empirical data can proceed on two lines: the "forward" procedure consists in calculating the time dependent solution $\bar{n}_i(t_1), \bar{n}_i(t_2), \ldots \bar{n}_i(t_T), i=1,2,\ldots L$ of the meanvalue equations with a *given set* of trend parameters $\tau = \{\nu, \delta_i, \kappa_i\}$ for a sequence of years $t_1, t_2, \ldots t_T$ and in comparing the result with the data. In general, however, the trend parameters, are not known in advance. Therefore, we have to resort to the "backward" procedure, i.e., a regression analysis consisting of the extraction of trend parameters from the comparison of empirical data with the theoretical expressions. This regression analysis is performed in the next section. The backward and forward procedures may also be combined, for instance by extracting trend parameters from the analysis of past migration data and by using them for predictive purposes under the assumption that the trends remain quasi-stable for a reasonable interval of time.

3.1 The Regression Analysis

In a migration system with L areas and one kind of population the following empirical quantities listed in Table 1 can be observed year by year:

area	population size	Number of transitions per year from i to j							
1	n_1		w_{21}	w_{31}	---	---	w_{j1}	---	w_{L1}
2	n_2	w_{12}		w_{32}	---	---	w_{j2}	---	w_{L2}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
i	n_i	w_{1i}	w_{2i}	w_{3i}	---	---	w_{ji}	---	w_{Li}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
L	n_L	w_{1L}	w_{2L}	w_{3L}	---	---	w_{jL}	---	

Table 1. Observed quantities per year--describing the migration process.

On the theoretical side it follows from the meanvalue equation (2.33) written in the form

$$\frac{\bar{n}_k(t+\Delta t) - \bar{n}_k(t)}{\Delta t} \cong \sum_{i=1}^{L} w_{ki}[\bar{n}(t)] - \sum_{j=1}^{L} w_{jk}[\bar{n}(t)] \qquad (3.1)$$

with $\Delta t = 1$ [year] , that

$$w_{ki}[\bar{n}] = \nu\, \bar{n}_i \exp\{(\delta_k - \delta_i) + \kappa_k(\bar{n}_k+1) - \kappa_i \bar{n}_i\} \qquad (3.2)$$

must be identified as the theoretical expression for the mean number of individuals migrating per year, that is per unit of time, from area i to area k.

The *theoretical* migration matrix $w_{ki}[n(t)]$ for given population numbers $n_i(t)$ therefore has to be matched to the *empirical* migration matrix $w_{ki}(t)$ by an optimal fitting of the trend parameters. There are (L^2-L) matrix elements $w_{ki}[n(t)]$ to be matched to the $w_{ki}(t)$ by fitting the $(2L+1)$ trend parameters $\{\nu(t), \delta_i(t), \kappa_i(t)\}$ year by year.

The optimization of the trend parameters amounts to the determination of the least square deviations between theoretical and empirical expressions. This optimization can be reduced to a *linear* regression analysis. For this aim we introduce (for each year, with the time index t omitted) the empirical quantities

$$r^e_{ji} = \ell n\left(\frac{w_{ji}}{\bar{n}_i}\right) \qquad (3.3)$$

and the corresponding theoretical expressions

$$\left.\begin{aligned} r^{th}_{ji} &= \ell n\left(\frac{w_{ji}[n]}{n_i}\right) \\ &= \mu + (\delta_j - \delta_i) + \kappa_j(n_j+1) - \kappa_i n_i \\ &\text{with } \mu = \ell n(\nu) \end{aligned}\right\} \qquad (3.4)$$

The latter are linear in the parameters $\{\mu, \delta_i, \kappa_i\}$ to be determined. We now require that the sum of the square deviations between r^e_{ji} and r^{th}_{ji}, namely

$$F(\mu,\delta,\kappa) = \sum_{i,j=1}^{L} (r_{ji}^{e} - r_{ji}^{th})^2$$
$$= \sum_{\substack{i,j=1 \\ (i\neq j)}}^{L} \{r_{ji}^{e} - [\mu + (\delta_j - \delta_i) + \kappa_j(n_j+1) - \kappa_i n_i]\}^2 \qquad (3.5)$$

be minized by an appropriate choice of the parameter set $\{\mu, \delta_i, \kappa_i\}$. The requirement of finding the minimum of $F(\mu,\delta,\kappa)$ leads to

$$\frac{\partial F}{\partial \mu} = 2 \sum_{\substack{i,j=1 \\ (i\neq j)}}^{L} (-1)\{r_{ji}^{e} - [\mu + (\delta_j - \delta_i) + \kappa_j(n_j+1) - \kappa_i n_i]\} = 0 \qquad (3.6)$$

$$\frac{\partial F}{\partial \delta_k} = 2 \sum_{\substack{i=1 \\ (i\neq k)}} \{r_{ki}^{e} - [\mu + (\delta_k - \delta_i) + \kappa_k(n_k+1) - \kappa_i n_i]\}$$
$$+ 2 \sum_{\substack{j=1 \\ (j\neq k)}}^{L} \{r_{jk}^{e} - [\mu + (\delta_j - \delta_k) + \kappa_j(n_j+1) - \kappa_k n_k]\} = 0 \qquad (3.7)$$

for $k = 1,2,\ldots L$.

and

$$\frac{\partial F}{\partial \kappa_k} = -2 \sum_{\substack{i=1 \\ (i\neq k)}}^{L} (n_k+1)\{r_{ki}^{e} - [\mu + (\delta_k - \delta_i) + \kappa_k(n_k+1) - \kappa_i n_i]\}$$
$$+ 2 \sum_{\substack{j=1 \\ (j\neq k)}}^{L} n_k\{r_{jk}^{e} - [\mu + (\delta_j - \delta_k) + \kappa_j(n_j+1) - \kappa_k n_k]\} = 0 \qquad (3.8)$$

for $k = 1,2,\ldots L$.

Since only the differences $(\delta_j - \delta_i)$ appear in the expression (3.5), the δ_j are only determined up to an additive constant. Hence we can put

$$\bar{\delta} = \frac{1}{L} \sum_{j=1}^{L} \delta_j = 0 \qquad (3.9)$$

Furthermore, we introduce the abbreviations

$$R_k^s = \sum_{\substack{i=1 \\ (i \neq k)}}^{L} (r_{ki}^e + r_{ik}^e) \; ; \; R^s = \sum_{k=1}^{L} R_k^s \tag{3.10}$$

$$R_k^{as} = \sum_{\substack{i=1 \\ (i \neq k)}}^{L} (r_{ki}^e - r_{ik}^e) \; ; \; R^{as} = \sum_{k=1}^{L} R_k^{as} \equiv 0 \tag{3.11}$$

and

$$\bar{\kappa} = \frac{1}{L} \sum_{j=1}^{L} \kappa_j \tag{3.12}$$

The evaluation of (3.6) and (3.7) then yields

$$\mu = \frac{1}{2L(L-1)} R^s - \bar{\kappa} \tag{3.13}$$

and

$$\delta_k = \frac{1}{2L} R_k^{as} - \frac{1}{2}(2n_k+1)\kappa_k + \frac{1}{2L} \sum_{i=1}^{L} \kappa_i (2n_i+1) \tag{3.14}$$

for k = 1,2,...L

The results (3.13) and (3.14) can be inserted into (3.8) to obtain equations for the parameters κ_k alone. The straight forward calculation leads to

$$(L-2)(\kappa_k - \bar{\kappa}) = B_k^s \qquad \text{for } k = 1,2,....L \tag{3.15}$$

where

$$B_k^s = R_k^s - \frac{1}{L} R^s \; , \qquad \text{with} \; \sum_{k=1}^{L} B_k^s = 0 \tag{3.16}$$

Since the L equations (3.15) are linearly dependent (the sum of the l.h.s. as well as the sum of r.h.s. of the equations vanish), the parameters κ are only determined up to an additive constant. This constant can be fixed by putting

$$\bar{\kappa} = 0 \tag{3.17}$$

Reinserting (3.17) into (3.13...3.15) the final results are obtained:

$$\mu = \frac{1}{2L(L-1)} R^s \tag{3.18}$$

$$\delta_k = \frac{1}{2L} R_k^{as} - \frac{1}{2}(2n_k+1)\kappa_k + \frac{1}{2L}\sum_{i=1}^{L}(2n_i+1)\kappa_i \tag{3.19}$$

for $k = 1,2,...L$, with $\bar{\delta} = 0$

and

$$\kappa_k = \frac{1}{(L-2)} B_k \tag{3.20}$$

for $k = 1,2,...L$, with $\bar{\kappa} = 0$

By the system (3.18, 19, 20) all trend parameters are uniquely determined and expressed by the known empirically r_{ji}^e , i,j,=1,...L, if $L \geq 3$. It is remarkable that the cooperation parameters κ_k and the mobility μ depend on R_k^s only, i.e., on symmetrical expressions in the r_{ki}^e, r_{ik}^e, $i \neq k$, while the preference parameters δ_k also depend on the quantities R_k^{as} which are asymmetric in the r_{ki}^e, r_{ik}^e, $i \neq k$.

3.2 The Evaluation Scheme and Conclusions

The linear regression analysis of the foregoing section leads to the optimal determination of the trend parameters $\mu(t)$, $\delta_i(t)$, $\kappa_i(t)$ for the years $t = 1,2,...T$, if the empirical migration matrix $w_{ij}(t)$ and population sizes $n_i(t)$ of Table 1 are known for these years. Which conclusions can be drawn from this analysis?

The first step consists in comparing the actual empirical situation with the accompanying "virtual" equilibrium situation: For each set of trend parameters $\{\mu, \delta_i, \kappa_i\}$ we may make use of (2.35) and (2.21) to determine the theoretical stationary population sizes $\{\bar{n}_1^{st}, \bar{n}_2^{st}, \ldots \bar{n}_L^{st}\}$ and even the theoretical stationary distribution $P_{st}(n_1, n_2, \ldots n_L)$ corresponding to these trend parameters. These stationary quantities describe the equilibrium situation into which the system would evolve, if the trend parameters would remain constant from this point of time. The actual system, however, in general is not in that equilibrium! Comparing the actual population sizes $\{n_1, n_2, \ldots n_L\}$ with these virtual equilibrium populations $\{\bar{n}_1^{st}, \bar{n}_2^{st}, \ldots \bar{n}_L^{st}\}$ we therefore obtain a measure for the momentary deviation of the migration system from its equilibrium state. This "distance from equilibrium" can be seen as a measure for the *migratory stress* in the population. The most compact formulation of the deviation between the actual and the equilibrium population distribution is given by the correlation coefficient

$$r(n, n^{st}) = \frac{\sum_{i=1}^{L} (n_i - \bar{n})(\bar{n}_i^{st} - \bar{n}^{st})}{\sqrt{\sum_{i=1}^{L} (n_i - \bar{n})^2 \sum_{j=1}^{L} (\bar{n}_j^{st} - \bar{n}^{st})^2}} \tag{3.21}$$

with

$$\bar{n} = \bar{n}^{st} = N/L \tag{3.22}$$

Obviously $|r|<1$, and r approaches 1 for $\{n_1 .. n_L\} \to \{\bar{n}_1^{st} \ldots \bar{n}_L^{st}\}$.

As a second step the results of the trend parameter determination for the *past* of a migration system can be used for *prognostic purposes*. Let us assume that the trend parameters $\{\mu(t), \delta_i(t), \kappa_i(t)\}$ have been found by regression analysis as above for a sequence of past years and that at most a slow time dependence was found, which can be approximated by

$$\mu(t) = \mu_0 + \mu_1 t \; ; \; \delta_i = \delta_{i0} + \delta_{i1} t \; ; \; \kappa_i = \kappa_{i0} + \kappa_{i1} t \; . \tag{3.23}$$

Apart from newly arising interfering factors it would be a plausible assumption to extrapolate the slow trend evolution into the future. The theoretically predicted values $n_1(t)$,

$n_2(t),\ldots n_L(t)$ of the population sizes in the areas then follow by solving the meanvalue evolution equations (2.33) using the trend parameters (3.23). (In a realistic prognosis of course also birth-death processes in each area have to be taken into account.)

Thirdly, we consider the case in which the empirical analysis exhibits a more pronounced and nonlinear time dependence of the trend parameters. Then it may be promising to analyze the correlation between a (representative) trend parameter τ and possible motivation factors μ_s, $s = 1,\ldots m$ creating the dynamic trends. The following standard method then can be applied: let $\tau^e(t)$ be an empirically determined trend parameter with linear trend in time subtracted and let $\mu_s^e(t)$, $s = 1,\ldots m$, be empirical properly standardized socioeconomic motivation factors net of linear trend with time, in the time interval $0<t<T$ under consideration! A tentative theoretical linear relation between the trend parameter and the motivation factors can then be assumed:

$$\tau^{th}(t) = \sum_{s=1}^{m} a_s \mu_s^e(t) \quad , \; t = 1,2,\ldots T \tag{3.24}$$

Between the empirical trend parameter $\tau^e(t)$ and its theoretical expression (3.24) there will exist a random deviation

$$(\tau^e(t) - \tau^{th}(t)) = \eta(t) \quad , \; t = 1,2,\ldots.T \tag{3.25}$$

The $\eta(t)$ have to be minimized by an optimal choice of the coefficients a_s in (3.24). The principle of the least sum of squares then yields

$$\frac{\partial F(a_1\ldots a_s)}{\partial a_s} = 0 \quad , \; s = 1,2,\ldots m \tag{3.26}$$

with

$$F(a_1\ldots a_s) = \sum_{t=1}^{T} [\tau^e(t) - \sum_{s=1}^{m} a_s \mu_s^e(t)]^2 \tag{3.27}$$

The evaluation of (3.26) leads to the set of linear equations for the a_s:

$$\sum_{r=1}^{m} C_{sr} a_r = D_s \ , \ s = 1,2,\dots,m \tag{3.28}$$

with

$$C_{sr} = C_{rs} = \sum_{t=1}^{T} \mu_r^e(t)\mu_s^e(t) \tag{3.29}$$

and

$$D_s = \sum_{t=1}^{T} \tau^e(t)\mu_s^e(t) \tag{3.30}$$

A measure for the agreement between $\tau^{th}(t)$ and $\tau^e(t)$ is the correlation coefficient $r(\tau^e,\tau^{th})$ defined as

$$r(\tau^e,\tau^{th}) = \frac{\sum_{t=1}^{T} (\tau^{th}(t) - \bar{\tau}^{th})(\tau^{r}(t) - \bar{\tau}^{e})}{\sqrt{\sum_{t=1}^{T} (\tau^{th}(t) - \bar{\tau}^{th})^2 \sum_{t=1}^{T} (\tau^{e}(t) - \bar{\tau}^{e})^2}} \tag{3.31}$$

Here

$$\bar{\tau}^{th} = \bar{\tau}^{e} = 0 \tag{3.32}$$

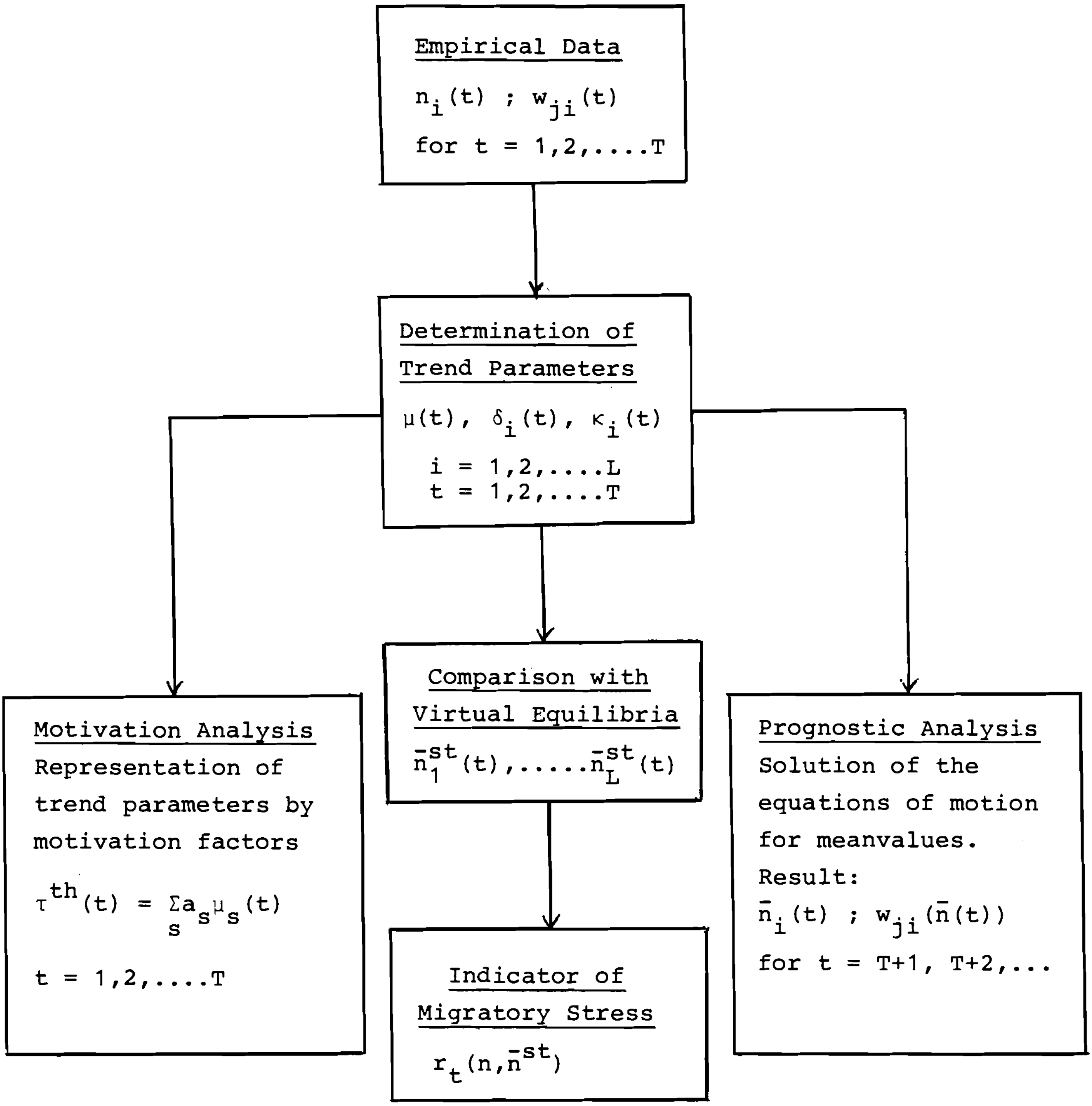
holds, since all time series of parameters are taken net of linear trend. As already mentioned in the introduction, the "explanation" of trend parameters by motivation factors can be ambiguous. An adequate definition of "equivalence" between different sets of motivation factors in the framework of linear regression analysis could be the following: two sets of motivation factors $\{\mu_1^{(1)},\dots\mu_s^{(1)}$ and $\{\mu_1^{(2)},\dots\mu_r^{(2)}\}$ are equivalent by definition, if

$$r(\tau^e,\tau_1^{th}) = r(\tau^e,\tau_2^{th}) \tag{3.33}$$

holds. Here τ_1^{th} and τ_2^{th} are linear combinations of $\{\mu_1^{(1)},...\mu_s^{(1)}\}$ and $\{\mu_1^{(2)},...\mu_r^{(2)}\}$, respectively, with optimal coefficients determined by linear regression analysis (see (3.26...30)).

Finally, we summarize the proposals of this section for the evaluation of the empirical data of a migration system in an "evaluation scheme".

Evaluation Scheme for a Migration System



4. A NUMERICAL SIMULATION

In this final section we demonstrate some aspects of the numerical evaluation of the model. We consider a fictitious migration system consisting of $L = 10$ areas, for which at time $t = 0$ the trend parameters ν, δ_k and κ_k and the initial area population numbers $n_k(t=0)$ for $k = 1,2,\ldots,10$ are known. Furthermore we assume that the trend parameters remain constant during the further evolution. Under such circumstances the meanvalue equations (2.34) can be used for predictive purposes. Their solution yields the expected area population numbers $n_k(1)$, $n_k(2),\ldots n_k(t),\ldots n_k(\infty) = n_k^{st}$ and, as a consequence, the expected migration matrices $w_{jk}(\underset{\sim}{n}(t))$. The population numbers $n_k(t)$ due to the structure of the meanvalue equations approach their stationary values $n_k(\infty) = n_k^{st}$ belonging to the (constant) trend parameters δ_k, κ_k, Hence, the correlation coefficient $r(n,n^{st})$ introduced in (3.21) can be expected to approach 1 for $t \to \infty$.

The choice of the trend parameters δ_k and κ_k for $k = 1,2,\ldots 10$ is represented in Figure 1. Positive preference parameters δ_k are assumed for areas $k = 6,\ldots 10$, and negative δ_k for areas $k = 1,\ldots 5$. Simultaneously, areas $k = 1,7,8$ have positive cooperation parameters κ_k (growth pool effects) while areas $k = 3,4,10$ have negative κ_k (repulsion effects) and areas $k = 2,3,6,9$ are neutral, i.e., $\kappa_k = 0$, with respect to density effects. The equations (3.9) and (3.17) are satisfied by this choice. Furthermore, the time scaling factor is chosen as $\nu = 0.01$.

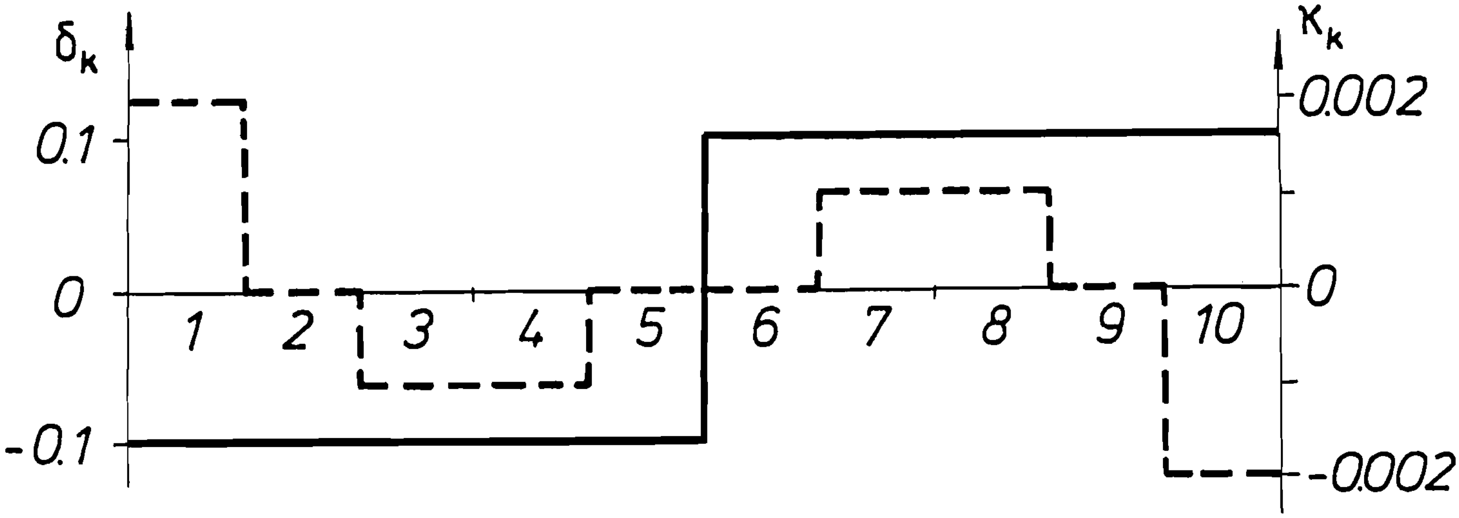


Figure 1. Choice of trend parameters δ_k (————) & κ_k (--------) for the 10 areas $k=1,2,\ldots,10$.

The main results of the solution of the meanvalue equation (2.34) are shown in Figures 2 and 3. In Figure 2 the area population numbers n_k, k = 1,...10 are represented for the initial time t = 0--where an equipartition was assumed--for an intermediate time t = 10, and for the final time $t \to \infty$ when the stationary distribution belonging to the trend parameters δ_k, κ_k has been reached. It can be seen, that strong growth effects in areas 7,8 or evacuation effects in areas 3,4 evolve, if preference and cooperation parameters have the same (positive or negative, respectively) sign and hence act in parallel. On the other hand, in area 1 the negative preference parameter δ_1 is over compensated by the positive cooperation parameter κ_1 so that a net growth effect prevails. The same holds, with inverted signs, for area 10. This demonstrates the important role of cooperative effects in migration theory.

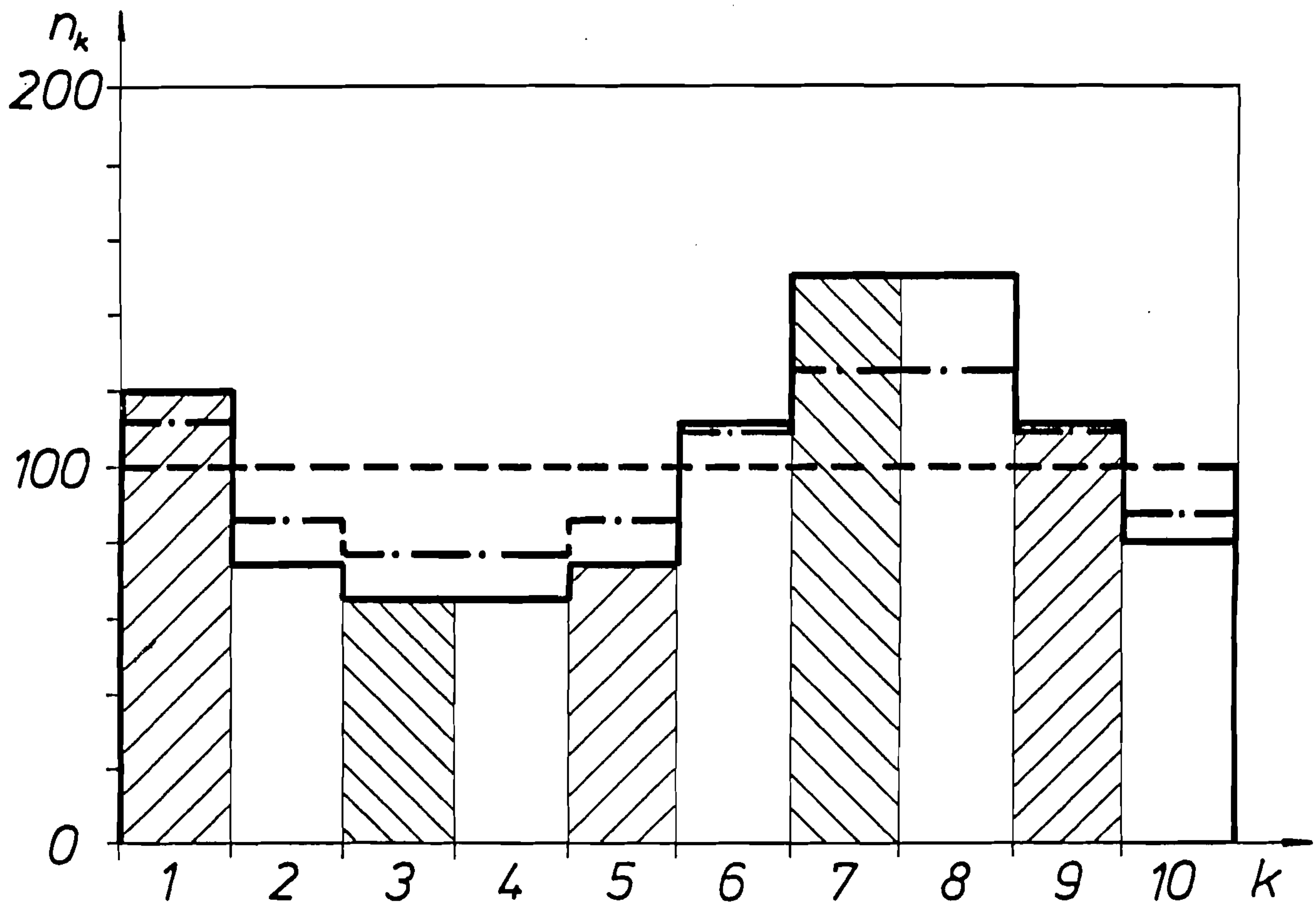


Figure 2. Evolution of area population numbers n_k for k=1,2,...,10: Initial distribution $n_k(t=0)$ (--------) ; intermediate distribution $n_k(t=10)$ (-·-·-·-·) ; final stationary distribution $n_k(t\ \ ) = n_k^{st}$ (————).

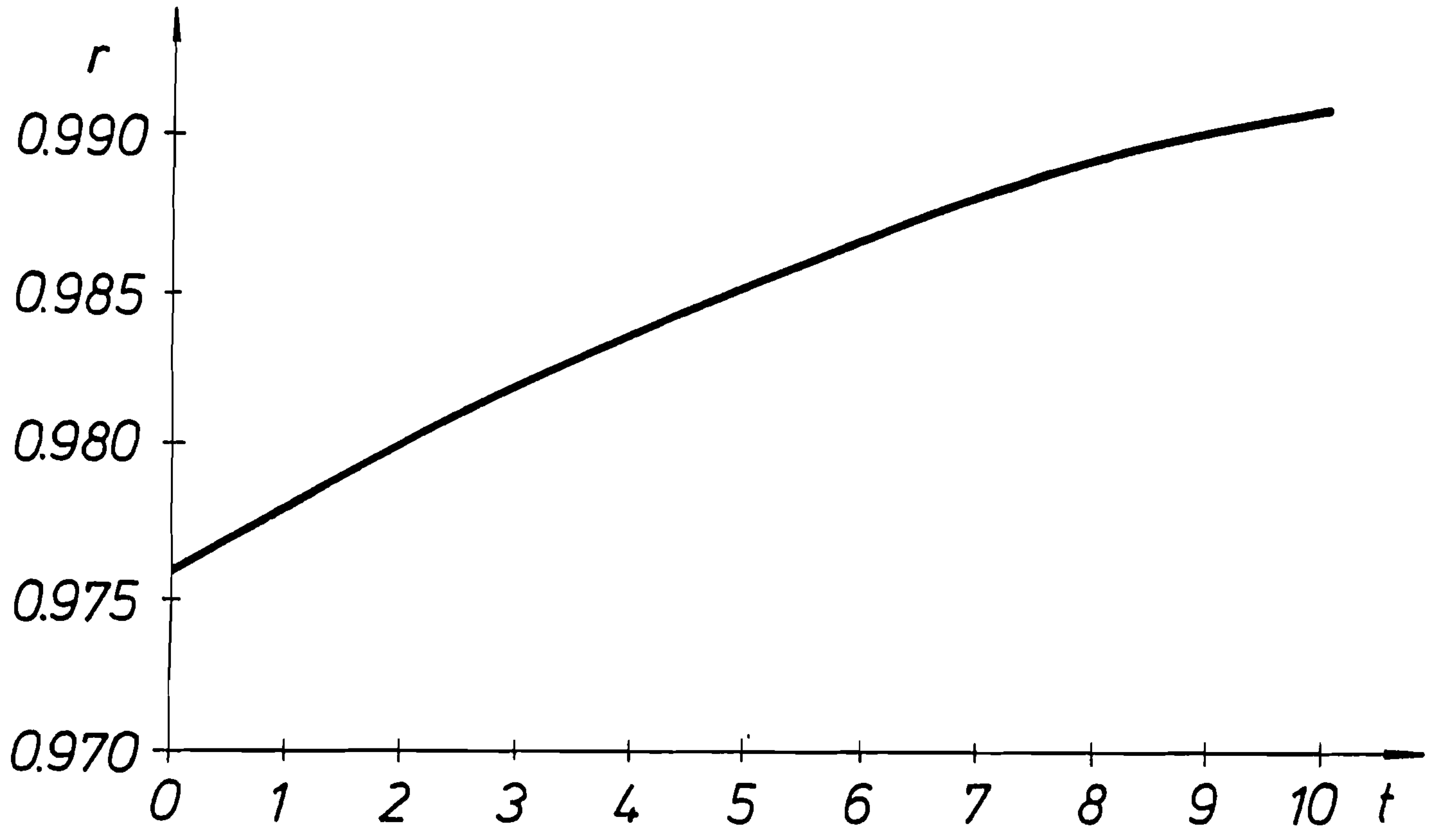


Figure 3. Evolution of the correlation coefficient $r(n,n^{st})$ with time.

Finally, the monotonous evolution of the correlation coefficient $r(n,n^{st})$ towards 1 is represented in Figure 3. It shows, that for constant trend parameters and starting from an initial nonequilibrium situation the system steadily approaches the equilibrium state $\{n_1^{st},\ldots n_{10}^{st}\}$. Since the mean-value equations (2.33) or (2.34) are invariant under the

$$\{n_k,\nu,\delta_k,\kappa_k\} \rightarrow \{an_k,\nu,\delta_k,a^{-1}\kappa_k\}$$

with an arbitrary constant a it is easy to construct migration systems for artitrary N and equivalent results with respect to the mean area population numbers.

APPENDIX: The Master Equation

Consider a system S which at any time is in one of a set of a finite or infinite number of discrete states $\{i\}$. If the system evolves *deterministically*, its state $i(t)$ at any time t is precisely known. If, on the other hand, only *probabilistic* laws about the evolution of S are known, the decisive quantity which then can be calculated is the so-called *conditional probability*

$$P(i,t|j,t_o) \tag{A.1}$$

which is, by definition, the probability to find S in state i at time t, given that S was in state j at the initial time $t_o < t$.

For many systems including migration models it is reasonable to make the *Markow-assumption* that the conditional probability depends on the end state i at time t and the initial state j at time t_o only, but *not* on former states which the system may have traversed prior to t_o. The conditional probability has two obvious properties implied by its definition:

$$P(i,t_o|j,t_o) = \delta_{ij} \tag{A.2}$$

since S is in state $i = j$ at time $t = t_o$ with certainty, and

$$\sum_i P(i,t|j,t_o) = 1 \qquad (A.3)$$

since the l.h.s. of (A.3) is the probability of finding the system in *any one* of the states i at time t, and this probability must be equal to 1.

Furthermore, if the conditional probability is known, we may calculate the further evolution with time of any given probability distribution over the states of the system. Hence, the conditional probability also is denoted as "propagator". Let

$$P(j;t_o) \qquad \text{with} \qquad \sum_j P(j;t_o) = 1 \qquad (A.4)$$

be the (properly normalized) probability to find the system in a state j at time t_o. Then the probability of state i at time $t > t_o$ is given by

$$P(i;t) = \sum_j P(i,t|j,t_o)P(j;t_o) \qquad (A.5)$$

since $P(i;t)$ can be represented as the sum of (conditional) probabilities to reach state i at time t from any of the states j at time t_o, each of them weighted with the probability $P(j;t_o)$ that this state j was realized at time t_o.

The master equation for $P(i;t)$ or for $P(i,t|j,t_o)$ now follows from considering the propagator for short time intervals $(t-t_o) = \tau$. Expanding the propagator in a Taylor series, and taking into account (A.2) and (A.3) we obtain

$$P(i,t_o+\tau|j,t_o) = \tau w_{ij} + 0(\tau^2) \qquad (A.6)$$

$$\text{for } i \neq j$$

with

$$w_{ij} = \left.\frac{\partial P(i,t|j,t)}{\partial t}\right|_{t=t_o} \qquad \text{for } i \neq j \qquad (A.7)$$

and

$$P(i,t_o+\tau|i,t_o) = 1 - \tau \sum_{k \neq i} w_{ki} + 0(\tau^2) \qquad \text{(A.8)}$$

Equation (A.8) follows by using (A.3). The *transition probabilities* w_{ij} are the transition rates of probability per unit of time from state j to state i, where $i \neq j$. In many theories the w_{ij} are known from basic considerations or by plausible assumptions. If the expansions (A.6) and (A.8) are inserted into (A.5) the rearrangement of terms and division by τ yields in the limit of $\tau \to 0$ the *master equation*

$$\lim_{\tau \to 0} \frac{P(i,t+\tau) - P(i,t)}{\tau} \equiv \frac{dP(i;t)}{dt} \qquad \text{(A.9)}$$

$$= \sum_{j \neq i} \{w_{ij}P(j;t) - w_{ij}P(i;t)\}$$

This first order differential equation for the evolution with time of the distribution function P(i;t) can be interpreted as a probability rate equation: the increase per time of the probability of state i is due to the net effect of probability transitions from all states j into state i (first term of the r.h.s.) and on the other hand from state i into all other states j (second term of the r.h.s.). If the transition probabilities w_{ij} do not depend on time t, the master equation thus describes a probability equilibration process starting with an arbitrary initial distribution $P(i;t_o)$ and ending with a final distribution $P(i;\infty) = P_{st}(i)$. The latter is the stationary probability distribution obeying the stationary master equation

$$0 = \sum_j \{w_{ij}P_{st}(j) - w_{ji}P_{st}(i)\} \qquad \text{(A.10)}$$

for all i.

In general it is not easy to obtain a practicable form for $P_{st}(i)$; only the graph-theoretically formulated solution accordingly to the Kirchhoff theorem is known.

In special cases, however, the condition of "detailed balance"

$$w_{ij}P_{st}(j) = w_{ji}P_{st}(i) \quad , \text{ for all } i,j \qquad \text{(A.11)}$$

is fulfilled. It means that not only the global balance (A.10) of all probability fluxes from and to the state i holds, but that the (stationary) probability flux from i to j is equal to that from j to i separately for *each pair* of states i and j. If (A.11) holds, the stationary solution can easily be constructed: take any chain C of states $i_o \equiv 0$, $i_1 \equiv 1, \ldots i_{n-1} \equiv (n-1)$, $i_n \equiv j$ from a reference state 0 to an arbitrary state j, so that all transition probabilities w(1,0),w(0,1),w(2,1),w(1,2),...w(j,n-1), w(n-1,j) are nonvanishing. (At least one such chain has to exist, otherwise no probability flux could reach state j.) The repeated application of (A.11) then yields

$$P_{st}(j) = c \prod_{\nu=0}^{n-1} \frac{w(\nu+1,\nu)}{w(\nu,\nu+1)} P_{st}(0) \qquad \text{(A.12)}$$

Finally, we derive conditions equivalent to the condition of detailed balance (A.11) which do not imply the (as yet unknown) stationary distribution $P_{st}(i)$. For this aim let the chain C be a closed loop L with the end state j equal to the reference state $i_o \equiv 0$. Because of $P_{st}(j) = P_{st}(0)$, it follows from (A.12) that

$$\left.\begin{array}{l} L \prod_{\nu=0}^{n-1} \frac{w(\nu+1,\nu)}{w(\nu,\nu+1)} = 1 \quad , \text{ or} \\[2ex] L \prod_{\nu=0}^{n-1} w(\nu+1,\nu) = \prod_{\nu=n-1}^{0} w(\nu,\nu+1) \\[2ex] \text{for all closed chains L of states.} \end{array}\right\} \qquad \text{(A.13)}$$

Vice versa, if (A.13) holds, it can be seen that $P_{st}(j)$ in (A.12) is, as it should be, independent of the specific choice of chain C from $i_o \equiv 0$ to $i_n \equiv j$, and (A.11) then follows from (A.12). Formula (A.13) will be used in section 2.3 to prove that detailed balance holds for the migration model and (A.12) will be used to construct the stationary solution of this model.

REFERENCES

Clark, W.A. and T.R. Smith (1982) Housing Market Search Behavior and Expected Utility Theory. *Environment and Planning A*, Vol.14:681-689 and Vol. 14:717-737.

Clark, W.A.V. and J.E. Burt (1980) The impact of Workplace on Residential Relocation, *Annals Association of American Geographers*, Vol.70:57-59.

Curry, L. (1981) Division of Labor from Geographical Competition, *Annals of the Association of American Geographers*, Vol. 71:133-165.

Dendrinos, D.S., and H. Mullally (1981) Evolutionary Patterns of Metropolitan Populations, *Geographical Analysis*, Vol. 13:328-344.

Griffith, D.A. (1982) Dynamic Characteristics of Spatial Economic Systems, *Economic Geography*, Vol. 58:178-196.

Griffith, D.A., and A.C. Lea (eds) (1983) *Evolving Geographical Structures*, Nato ASI Series D, No. 15, Martinus Nijhoff Publishers.

Haag, G., and D.S. Dendrinos (1983) Toward a Stochastic Dynamical Theory of Location: A Non-Linear Migration Process. *Geographical Analysis*, forthcoming.

Leonardi, G. (1983) An Optimal Control Representation of a Stochastic Multistage-Multiactor Choice Process, in *Evolving Geographical Structures*, Nato ASI Series D, No. 15, (eds) D.A. Griffith and A.C. Lea.

MacKinnon, R. (1970) Dynamic Programming and Geographical Systems. *Economic Geography*, Vol. 46:350-366.

Sonis, M. (1981) Flows, Hierarchies, Potentials. *Environment and Planning A*, Vol. 13:413-420.

Weidlich, W., and G. Haag (1980) Migration Behavior of Mixed Population in a Town. *Collective Phenomena*, Vol. 3:89-102.

Weidlich, W., and G. Haag (1983) *Concepts and Models of a Quantitative Sociology, The Dynamics of Interacting Populations*. Springer-Series in Synergetics, Vol. 14. Springer, Heidelberg/New York.