ENTROPY, MULTIPROPORTIONAL, AND QUADRATIC TECHNIQUES FOR INFERRING PATTERNS OF MIGRATION FROM AGGREGATE DATA

Frans Willekens, András Pór, and Richard Raquillet

1 INTRODUCTION

The lack of adequate regional statistics is frequently given as a reason for not endorsing the development of sophisticated regional or multiregional models. The data problem is particularly severe in the case of studies of interregional migration disaggregated by category of migrant.

Improved modeling of migration and of multiregional demographic phenomena in general can frequently only be carried out if additional data requirements are met. However, in cases where the collection of the necessary data is either impossible or too expensive, estimation procedures may sometimes be used as substitutes.

This paper presents a particular class of estimation methods which have great potential in migration analysis. They may be used whenever detailed migration flows (e.g., flows classified into various migrant categories) are required, but only aggregate data on migration patterns are initially available. For instance, how can we disaggregate a total origin–destination migration-flow matrix into age-specific flows if the only information available is a national estimate of the age composition of migrants? How would this estimate differ from one based on information about the age structure of migrants arriving in and departing from each region, or on an initial crude estimate of the age-specific flow matrices? Does additional information improve our estimates significantly? (This is of particular relevance in deciding on the extent of data collection required for a given migration analysis.) Questions like these may be answered by applying the techniques proposed in this paper. In addition, these methods are useful for updating migration tables, because this updating consists of nothing more than estimating migration flows under new conditions (constraints) when estimates under existing conditions are available.

The applicability of the methodology is not limited to migration studies alone, but may also be extended to input–output analysis, transportation analysis (e.g., trip distribution, freight flows), regional economics (e.g., journey-to-work tables), and other fields: in short, to the analysis of various kinds of interaction tables or contingency tables. In fact these techniques, in their simplified (two-dimensional) forms, have received considerable attention in the above-mentioned fields. Quite recently it was realized that the mathematical and statistical characteristics of the basic problem addressed in this paper have also been investigated in contingency-table analysis (the analysis of qualitative, cross-classified data) (see, for example, Bishop *et al.* 1975). A number of results from this research may be used to study the problem of estimating detailed migration flows from aggregate data and, in particular, to interpret the accuracy of the estimates. The relationship between the estimation methods presented in this paper and the techniques of contingency-table analysis has been discussed by Willekens (1980).

After this introduction, the paper consists of four main sections (Sections 2–5) and two appendixes. In Section 2, the general problem is dealt with in mathematical terms, and some special cases, which are of particular practical interest, are formulated. Section 3 presents the solutions to the basic problem and its variants. In Section 3.1 it is shown that, in terms of mathematical programming, the basic or primal problem is equivalent to another nonlinear problem (dual problem) that may be solved using standard methods. The solution algorithms, which are implemented on the computer, follow in Section 3.2. Section 4 presents some numerical illustrations, namely, the estimated values of interregional migration flows in Austria, disaggregated by age; the estimates are compared with the observed values to investigate the validity of the estimation procedure. A similar validity analysis using Swedish data is also discussed. Conclusions and some suggestions for further research are presented in Section 5. Proofs of five theorems used in the paper, and multiproportional solutions for three special cases of the entropy problem are given in Appendixes A and B, respectively. This paper is a shortened and modified version of an earlier report (Willekens *et al.* 1979), which contains additional applications and a description and listing of the computer program MULTENTROPY that was used to obtain the numerical results presented in this paper.

The entropy method discussed here (the 3F variant) was used in the IIASA Comparative Migration and Settlement Study to obtain estimates of interregional age-specific migration flows for countries with incomplete migration data (Bulgaria and the Netherlands). The applications to Bulgaria and the Netherlands have been described by Philipov (1978) and by Drewe and Willekens (1980), respectively. The method was also used by Tan (1980) to estimate missing migration data for Belgium.

2 PROBLEM FORMULATION

The estimation procedures proposed in this paper for inferring detailed migration patterns from aggregate data have one common feature: the aggregate data appear as sums (hereafter referred to as "marginal sums") of different groups of elements of two- or n-dimensional arrays, the elements of which are unknown and must be estimated. (In the two-dimensional case these marginal sums usually represent row or column sums.) Two main groups of problems may be distinguished: (i) the entropy problem, and (ii) the quadratic adjustment problem.

(i) *The entropy problem.* The problem here is to produce a "maximally unbiased" estimate of the elements of an array under the given marginal conditions. In the application of entropy models, two model types may be distinguished: (a) entropy maximizing models, and (b) information-divergence (I-divergence) minimizing models.

(a) In the entropy maximizing problem one tries to determine the "most probable" elements of an array, under the given marginal conditions. No initial array is known *a priori* and the values of the elements of the array are seen as equally likely, apart from the marginal constraints specified. The entropy maximizing method was introduced in regional science by Wilson (1967, 1970). The two-dimensional case has received considerable attention in the literature and has been elaborated in several ways to recover interregional flow matrices (of people or of commodities) from various forms of aggregate data (Chilton and Poet 1973, Evans and Kirby 1974, Nijkamp 1975, Willekens 1977).

(b) In the I-divergence minimizing problem one tries to estimate an "*a posteriori*" array which is as "close" as possible to an "*a priori*" array, and which satisfies some given constraints (row and column sums).
The distance function used here to measure the "closeness" of the arrays is the I-divergence or Kullback–Leibler information number (Kullback 1959), also called information for discrimination, *information gain*, or entropy of an *a posteriori* distribution relative to an *a priori* distribution (Renyi 1970). If the known *a priori* array is uniformly distributed, i.e., all elements of the array are equal, then the I-divergence measure is equivalent to the negentropy (entropy with a negative sign). I-divergence with a negative sign was also defined as a measure of the average conditional entropy (Nijkamp and Paelinck 1974a, Theil 1967).
A procedure for minimizing I-divergence in spatial interaction models has been developed by Batty and March (1976). In the two-dimensional case,* the biproportional adjustment method, better known as the RAS method, is the technique most extensively used to solve this type of problem. It was developed independently by Leontief and Stone to update input–output tables and has been studied in detail by Bacharach (1970). Hewings (1977), Thumann (1978), and Hewings and Janson (1980) have provided recent evaluations of RAS and related techniques with special emphasis on their capabilities for exchanging, updating, and predicting regional input–output coefficients. The RAS method is equivalent to the Furness method used in traffic models (e.g., Evans 1970, Evans and Kirby 1974). In the analysis of contingency tables the method is known as iterative proportional fitting and is attributed to Bartlett in a paper written as early as 1935.

(ii) *The quadratic adjustment problem.* This problem applies to situations where initial estimates of the entries of an array are available but where the estimates do not conform with predefined (measured) row and column totals. The reason may be that different sources have been used to calculate the estimates and the totals. The sum-constrained adjustment problem is then to find the array which is as close as possible to the initial array and also satisfies the predefined totals.

The distance function used here to measure the "closeness" of the arrays is a quadratic-type function. Examples can easily be found to show that, unlike the I-divergence minimizing problem, the fact that the elements of the initial array are positive does not ensure that the unique solution of the quadratic adjustment problem will also be positive. To overcome this weakness of the quadratic adjustment method, we suggest an adjustment technique which is not quadratic but which is very closely related to the quadratic adjustment method.

---

* A two-dimensional array is a matrix.

2.1 Entropy Maximization

The techniques described in this section address problems in which no input matrix is given *a priori*. Wilson (1967, 1970) used an index to produce the matrix entries which are *most probable*. This index is called the entropy of the matrix.

2.1.1 The Entropy Concept

Suppose that we are given the total number of arrivals I_i and departures O_i disaggregated by region in a two-region system, and that the problem is to estimate the complete origin–destination migration-flow matrix **M** with elements m_{ij}. For example, to estimate

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{21} \\ m_{12} & m_{22} \end{bmatrix}$$

where

$$I_1 = m_{11} + m_{21} = m_{.1} = 3 \qquad O_1 = m_{11} + m_{12} = m_{1.} = 4$$

$$I_2 = m_{12} + m_{22} = m_{.2} = 3 \qquad O_2 = m_{21} + m_{22} = m_{2.} = 2$$

$$m_{11} + m_{12} + m_{21} + m_{22} = m_{..} = 6$$

In contrast to the biproportional adjustment method, no initial estimates of the matrix elements are available. Therefore, our information is limited to the row and column totals of the matrix to be estimated. For given row and column totals of a matrix, there may be a large number of arrangements of entries that satisfy the marginal conditions. For example, if for a 2×2 migration matrix the row sums are three and three, and the column sums are four and two, then there are three possible arrangements of the entries

$$\mathbf{M}_a = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \qquad \mathbf{M}_b = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \qquad \mathbf{M}_c = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$$

Each arrangement of the entries of **M** is called a *macrostate* of the system.

The true migration flow is represented by one of the three macrostates, $\mathbf{M}_a$, $\mathbf{M}_b$, or $\mathbf{M}_c$. Given the limited information we have about the migration behavior, we do not know which macrostate is the true one. Therefore, we must make a guess. It is here that the entropy method is useful. It selects the macrostate which has the highest probability of occurring. A certain macrostate may be generated by various so-called *microstates*. A microstate is an assignment of *individual* migrants to the origin–destination table. In other words, a microstate is a description of the location of every individual in the system, whereas a macrostate gives the number of people in each cell of the table. Consider, for example, the matrix $\mathbf{M}_a$, and denote the individual migrants by $m1$, $m2$, $m3$, $m4$, $m5$, and $m6$. According to $\mathbf{M}_a$, three people migrate from region 1 to region 1, i.e., move within the region. We can select these three from the six migrants in 20 different ways. The number of possible combinations of three people from six can easily be computed using the well-known combinatorial formula

$${}^{3}C_{6} = \frac{6!}{3!(6-3)!}$$

Once we have chosen three people to constitute m_{11}, we must select one person out of the remaining three to constitute m_{12}. There are only three possible ways of doing this. Finally, the two remaining individuals constitute m_{22}, since $m_{21} = 0$. Therefore, the total number of ways of selecting three out of six, one out of the remaining three, and two out of the remaining two, is

$$\frac{6!}{3!(6-3)!}\,\frac{3!}{1!(3-1)!}\,\frac{2!}{2!} = 60$$

Each of the 60 ways constitutes a separate microstate, or assignment of individuals. In general, the number of ways in which we can select a *particular* macrostate from the total number of migrants $m_{..}$ is given by the combinatorial formula

$$W = \frac{m_{..}!}{m_{11}!\,m_{12}!\,m_{21}!\,m_{22}!} \tag{1}$$

$$W = \frac{m_{..}!}{\prod_{i,j} m_{ij}!}$$

Using eqn. (1), we obtain $W = 60$ for $\mathbf{M}_a$, $W = 180$ for $\mathbf{M}_b$, and $W = 60$ for $\mathbf{M}_c$. The value W is the number of microstates which give rise to a particular macrostate, and is called the *entropy* of the macrostate.

The macrostate $\mathbf{M}_b$ with the highest entropy value is then chosen as the best estimate of the true migration flow. The use of this selection criterion relies on two critical assumptions:

1. The probability that a macrostate represents the true migration-flow matrix is proportional to the number of microstates of the system which give rise to this macrostate (entropy) and which satisfy the marginal conditions.
2. Each microstate is equally probable.

The formal solution to the two-dimensional entropy maximization problem is derived in the next section.

2.1.2 Solving the Entropy Problem

The first assumption given above may be stated in a slightly different way: the true arrangement of a system is one which maximizes the entropy, i.e., one in which the elements can be organized in as many ways as possible (maximum "disorder"). This is the second law of thermodynamics. The analogy between the behavior of social and physical systems is not accidental. Several authors (e.g., Isard 1960) have attempted to describe social phenomena by laws borrowed from physics. This approach, known as *social physics*, was developed in the early regional-science literature.

The use of the entropy concept in the social sciences may be illustrated by information theory (Jaynes 1957), by Bayes's theorem for conditional probabilities (Hyman 1969), and by the use of maximum-likelihood estimators (Evans 1971, Batty and MacKie 1972).* In information theory, entropy represents expected

---

* For a comparison, see Wilson (1970, pp. 1–10) and Nijkamp (1977, pp. 18–20).

information. It indicates the degree of uncertainty about the occurrence of events in information systems. Hence, a high entropy value (low uncertainty) is associated with events which are likely to occur. In the maximum-likelihood approach, entropy maximization is equivalent to maximizing the likelihood of a macrostate.

The problem of estimating the most probable migration-flow matrix to satisfy given row and column totals may now be formulated as follows: find the macrostate with maximum entropy W, subject to the marginal conditions. The solution is given by

$$\max W = \frac{m_{..}!}{\prod_{i,j} m_{ij}!} \tag{2}$$

subject to

$$\sum_j m_{ij} = m_{i.} = O_i \qquad \text{for all } i \tag{3}$$

$$\sum_i m_{ij} = m_{.j} = I_j \qquad \text{for all } j \tag{4}$$

Since the maximum of eqn. (2) coincides with the maximum of any monotonic function of W, we may replace W by the Naperian logarithm of W ($\ln W$) in the objective function

$$\begin{aligned} \ln W &= \ln m_{..}! - \ln \prod_{i,j} m_{ij}! \\ &= \ln m_{..}! - \sum_i \sum_j \ln m_{ij}! \end{aligned} \tag{5}$$

To make differentiation of the complex function (5) easier, we replace $\ln m_{ij}!$ by Stirling's approximation

$$\ln m_{ij}! = m_{ij} \ln m_{ij} - m_{ij}$$

Since $\ln m_{..}!$ is a constant, we may write the objective function as

$$\max \ln \hat{W} = -\sum_i \sum_j (m_{ij} \ln m_{ij} - m_{ij}) \tag{6}$$

or equivalently

$$\min \ln \hat{W} = \sum_i \sum_j m_{ij} \ln m_{ij} - m_{ij} \tag{7}$$

2.2 I-Divergence Minimization (Biproportional Adjustment)

The biproportional adjustment method, independently developed by Leontief (1941) and Stone (1963), and equivalent to the iterative proportional fitting procedure of Bartlett (1935), uses the I-divergence measure as a distance function

$$I(\mathbf{M} \| \mathbf{M}^0) = \sum_i \sum_j m_{ij} \ln (m_{ij}/m_{ij}^0) \tag{8}$$

which is defined for $m_{ij}^0 \neq 0$. This technique is better known as the RAS method (Stone 1963).

The basic features of the biproportional adjustment problem are described in this section. Quite often our information is not limited to the row and column sums of a two-dimensional migration matrix; we may also know the migration patterns of specific categories of the population (e.g., sexes, age groups). This information may then be used to adjust the original estimates. A more general version of the biproportional adjustment process, the multiproportional adjustment process, allows more *a priori* information to be taken into consideration; this is treated in Section 2.5.

Suppose we are given the total number of migrants arriving in each region, and departing from each region (I_j and O_i, respectively) in a two-region system, and that, as before, the problem is to estimate the complete origin–destination migration-flow matrix **M** with elements m_{ij}. For example, to estimate

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{21} \\ m_{12} & m_{22} \end{bmatrix}$$

where

$$I_1 = m_{.1} \qquad O_1 = m_{1.}$$

$$I_2 = m_{.2} \qquad O_2 = m_{2.}$$

Suppose we are also given initial estimates of the elements m_{ij}; these estimates are the elements m^0_{ij} of the matrix $\mathbf{M}^0$

$$\mathbf{M}^0 = \begin{bmatrix} m^0_{11} & m^0_{21} \\ m^0_{12} & m^0_{22} \end{bmatrix}$$

The initial estimates may be derived from migration tables of previous years, from experts' opinions, or from other sources.* However, the column sums $m^0_{i.}$ and the row sums $m^0_{.j}$ of $\mathbf{M}^0$ are not necessarily equal to the predefined number of departures $O_i = m_{i.}$, and arrivals $I_j = m_{.j}$. We therefore have to adjust the elements of $\mathbf{M}^0$ so that they add up to the required totals. The sum-constrained biproportional adjustment problem may be formulated as follows:

Find the $p \times q$ matrix **M** such that the I-divergence measure is minimized (in migration tables, $p = q$):

$$\min I(\mathbf{M} \| \mathbf{M}^0) \tag{9}$$

subject to

$$m_{i.} = \sum_j m_{ij} = O_i \qquad (i = 1, 2, \ldots, q) \tag{10}$$

$$m_{.j} = \sum_i m_{ij} = I_j \qquad (j = 1, 2, \ldots, p) \tag{11}$$

Note that since $\sum_i O_i = \sum_j I_j$, the $p + q$ row and column totals are not independent. The number of independent constraints is $p + q - 1$. Constraints (10) and (11) are found in all existing methods for constrained adjustment of a two-dimensional array, i.e., of a matrix.

* This methodology has been particularly successful for updating input–output tables.

2.3 A Comparison of Entropy Maximization and I-Divergence Minimization

This section compares the entropy maximizing and I-divergence minimizing techniques. Both techniques have the same set of constraints,* and any differences may therefore be explained by differences in the objective functions.

(i) Note that the entropy objective function to be maximized

$$\hat{W} = -\sum_i \sum_j (m_{ij} \ln m_{ij} - m_{ij})$$

is also given by

$$\hat{W} = -\sum_i \sum_j (m_{ij} \ln m_{ij}) + m_{..}$$

Since $m_{..}$ is a constant, the maximization of $\hat{W}$ gives the same result as the maximization of the simpler function $\hat{\hat{W}} = -\sum_i \sum_j m_{ij} \ln m_{ij}$. The objective function $\hat{\hat{W}} = -\sum_i \sum_j m_{ij} \ln m_{ij}$ is in fact more widely used. If m_{ij} is interpreted as a probability by scaling $(\sum_i \sum_j m_{ij} = 1)$, then the objective function defines a quantity called *statistical entropy* or the entropy of the probability distribution. This quantity (multiplied by a constant) has been defined by Shannon as a measure of the uncertainty contained in a probability distribution (Shannon and Weaver 1949).

(ii) The biproportional process differs from the entropy maximizing process in that it contains a term m_{ij}^0 representing an initial estimate of the elements. Again, if m_{ij} and m_{ij}^0 can be interpreted as probabilities, the quantity**

$$W = \sum_i \sum_j m_{ij} \ln (m_{ij}/m_{ij}^0)$$

represents the *information divergence*.

(iii) It appears that all these objective functions belong to the same family and merely express the concept of entropy in different ways; the same basic idea is common to biproportional "minimum deviation" and to statistical entropy. In information theory, the biproportional objective function is interpreted as a measure of the "surprise" produced by new values when compared with old values. Therefore, minimizing the "surprise" or the deviation from our initial information and finding the most probable flows are the basic aims of both the entropy and the biproportional approaches. Both of these techniques measure the deviation with the same function

$$\sum_i \sum_j m_{ij} \ln (m_{ij}/m_{ij}^0) \tag{12}$$

when *a priori* information is available, and use the function

$$\sum_i \sum_j m_{ij} \ln m_{ij} \tag{13}$$

---

* If a cost constraint is considered in an entropy problem, then the reciprocal of the cost function plays the role of an *a priori* array $\mathbf{M}^0$ (Willekens 1980).

** When some *a priori* information exists and is expressed in terms of *a priori* probabilities (p_i^0, $i = 1, 2, \ldots, n$), the expected value or information content of a message changing *a priori* information into *a posteriori* information is measured by $-\sum_i p_i \ln (p_i/p_i^0)$ (Jaynes 1957, Theil 1967).

when no such information exists. Therefore, entropy maximizing problems can generally be formulated as I-divergence minimizing problems

$$\min d[\mathbf{M}, \mathbf{M}^0] = \sum_i \sum_j m_{ij} \ln (m_{ij}/m_{ij}^0) \tag{14}$$

where m_{ij}^0 can be uniformly set equal to 1.*

2.4 Quadratic Adjustment

We may formulate the sum-constrained quadratic adjustment problem in the following way.

Find a matrix **M** such that a quadratic distance measure $d[\mathbf{M}, \mathbf{M}^0]$ is minimized, subject to

$$\sum_j m_{ij} = O_i \qquad \text{for all } i$$

$$\sum_i m_{ij} = I_j \qquad \text{for all } j$$

The techniques may be distinguished by the distance measures $d[\mathbf{M}, \mathbf{M}^0]$ used.

(i) *Least-squares adjustment.* The most obvious distance measure is the euclidean norm, defined as

$$d[\mathbf{M}, \mathbf{M}^0] = \tfrac{1}{2} \sum_i \sum_j (m_{ij} - m_{ij}^0)^2 \tag{15}$$

(ii) *Friedlander adjustment.* Friedlander (1961) used a distance norm of the χ^2 type

$$d[\mathbf{M}, \mathbf{M}^0] = \tfrac{1}{2} \sum_i \sum_j (m_{ij} - m_{ij}^0)^2 / m_{ij}^0 \tag{16}$$

to adjust contingency tables. A related method has been developed by Hortensius (1970). The distance norm is the weighted deviation

$$d[\mathbf{M}, \mathbf{M}^0] = \sum_i \sum_j (m_{ij} - m_{ij}^0)^2 / s_{ij} \tag{17}$$

where $1/s_{ij}$ is the weight corresponding to the difference $(m_{ij} - m_{ij}^0)$. The weight used is a measure of the uncertainty. The underlying idea is that the more accurate m_{ij}^0 is, the less adjustment is needed. Hence, for accurate estimates of m_{ij}^0, the weight $1/s_{ij}$ should be small. If the elements of m_{ij}^0 are estimated from a sample, then both the estimate or mean m_{ij}^0 and the whole frequency distribution are known. An appropriate value for s_{ij} is therefore the variance of the distribution (assuming normal distribution). This measure of uncertainty is used by Hortensius.

* In general, the I-divergence or multiproportional problem reduces to an entropy problem whenever the elements of the initial array are multiproportional (with the uniform distribution as a special case). The results are therefore independent of the initial array. Friedmann (1978) recently showed that the RAS solution is the same for different initial biproportional matrices, although he did not connect this with the entropy problem. MacGill (1977) provides a further discussion on the formal equivalence between entropy and biproportional adjustment methods.

(iii) *Modified Friedlander adjustment.* An important weakness of the least-squares and Friedlander approaches is that a strictly positive matrix $\mathbf{M}^0$ does not necessarily yield a strictly positive matrix $\mathbf{M}$ of estimates. Some elements of $\mathbf{M}$ may be negative. To overcome this failure, we suggest the following measure, which is no longer quadratic:

$$d[\mathbf{M}, \mathbf{M}^0] = \tfrac{1}{2} \sum_i \sum_j (m_{ij} - m_{ij}^0)^2 / m_{ij}$$

2.5 Generalization of Flow-Estimation Problems to *N*-Dimensional Arrays

The formulation of the problem can easily be extended to more than two dimensions. Suppose that we are given the migration-flow matrix of the total population, and that we are interested in the migration patterns of subsets of the population, e.g., sexes, age groups, nationalities, professional categories. Suppose that we also know the number of arrivals and departures made by each subset in each region. The migration from i to j by category k is denoted by m_{ijk}. The total migration from i to j is $m_{ij.} = c_{ij}$, the number of departures from i by category k is $m_{i.k} = b_{ik}$, and the number of arrivals in j by category k is $m_{.jk} = a_{jk}$. Therefore, the following marginal constraints must be met:

$$\sum_{i=1}^{n} m_{ijk} = a_{jk} \qquad (j = 1, 2, \ldots, m;\ k = 1, 2, \ldots, l)$$

$$\sum_{j=1}^{m} m_{ijk} = b_{ik} \qquad (i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, l)$$

$$\sum_{k=1}^{l} m_{ijk} = c_{ij} \qquad (i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, m)$$

The information available is, however, not always presented in this way. For example, in an extreme case, we may not know the flow matrix of the total population but only the total number of arrivals and departures disaggregated by region; we may know the composition of the migrant categories only at the national level. In this case the constraints would take on a different form:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} m_{ijk} = u_k \qquad (k = 1, 2, \ldots, l)$$

$$\sum_{i=1}^{n} \sum_{k=1}^{l} m_{ijk} = v_j \qquad (j = 1, 2, \ldots, m)$$

$$\sum_{j=1}^{m} \sum_{k=1}^{l} m_{ijk} = w_i \qquad (i = 1, 2, \ldots, n)$$

where $u_k = m_{..k}$ is the total number of migrants in category k, $v_j = m_{.j.}$ is the total number of arrivals in region j, and $w_i = m_{i..}$ is the total number of departures from region i.

Various combinations of bivariate and univariate marginal sums are possible: a known total-flow matrix with the composition of migrant categories at the national level, numbers of arrivals and departures disaggregated by type of migrant only, etc.

Before proceeding to specific problems, we may formulate our problem in general terms. The mathematical formulation of the *basic adjustment problem* is as follows:

$$\min \sum_i \sum_j \sum_k d[m_{ijk}, m^0_{ijk}] \tag{18}$$

subject to

$$\sum_i m_{ijk} = a_{jk} \qquad (\forall j \in J, k \in K) \tag{19}$$

$$\sum_j m_{ijk} = b_{ik} \qquad (\forall i \in I, k \in K) \tag{20}$$

$$\sum_k m_{ijk} = c_{ij} \qquad (\forall j \in J, i \in I) \tag{21}$$

$$\sum_i \sum_j m_{ijk} = u_k \qquad (\forall k \in K) \tag{22}$$

$$\sum_i \sum_k m_{ijk} = v_j \qquad (\forall j \in J) \tag{23}$$

$$\sum_j \sum_k m_{ijk} = w_i \qquad (\forall i \in I) \tag{24}$$

$$\sum_k u_k = \sum_j v_j = \sum_i w_i = ST \tag{25}$$

$$m_{ijk} \geq 0 \qquad (\forall i \in I, j \in J, k \in K) \tag{26}$$

where $I = \{1, 2, \ldots, n\}$, $J = \{1, 2, \ldots, m\}$, and $K = \{1, 2, \ldots, l\}$ are the index sets; a_{jk}, b_{ik}, and c_{ij} are bivariate marginal sums; u_k, v_j, and w_i are univariate marginal sums; and ST is the total sum.

The elements to be estimated, m_{ijk}, may be arranged in a three-dimensional array, $\mathbf{M} = [m_{ijk}]$; the initial estimates constitute the array $\mathbf{M}^0 = [m^0_{ijk}]$. Both arrays contain only non-negative elements. Some of the elements m_{ijk} may be known exactly. For instance, if intraregional migration is not considered, then the diagonal elements $m_{iik} = m^0_{iik} = 0$. If other migration flows are known *a priori* (i.e., are fixed to the initial estimate m^0_{ijk}), we have to consider the cell constraints

$$m_{ijk} = m^0_{ijk} \qquad (i, j, k) \notin \Gamma \tag{27}$$

where the set Γ is defined by setting $\Gamma = \{(i, j, k) | \text{migration from } i \text{ to } j \text{ by category } k \text{ is possible and not fixed}\}$.

The right-hand sides of the constraints are matrices $\mathbf{A} = [a_{jk}]$, $\mathbf{B} = [b_{ik}]$, and $\mathbf{C} = [c_{ij}]$, and vectors $\boldsymbol{U} = [u_k]$, $\boldsymbol{V} = [v_j]$, and $\boldsymbol{W} = [w_i]$. We shall call the constraints (19)–(21) *face constraints*, because they present given values for the three faces of the "cube" or three-dimensional array $\mathbf{M}$. Analogously, the univariate constraints (22)–(24) will be labeled *edge constraints* because they allocate values to the edges of the "cube" or three-dimensional array $\mathbf{M}$. If the face constraints are given, the edge constraints are redundant since the elements on the edges are the sums of the elements on the faces. We shall therefore refer to the basic problem as the "three prescribed faces" problem, or, more concisely, the "three-face (3F)" problem.

Three special cases are of particular interest:

1. The two-face (2F) problem—minimize (18) subject to (19), (21), (26), and (27);
2. The one-face, one-edge (1FE) problem—minimize (18) subject to (21), (22), (26), and (27);
3. The three-edge (3E) problem—minimize (18) subject to (22)–(26), and (27).

To find the best estimates of the elements m_{ijk}, given the constraints and given an array of initial estimates m^0_{ijk}, we generalize the different distance measures as follows:

(i) *Three-dimensional I-divergence measure*

$$d[\mathbf{M}, \mathbf{M}^0] = \sum_{(i,j,k)\in\Gamma} m_{ijk} \ln (m_{ijk}/m^0_{ijk}) \tag{28}$$

Note that the entropy function is equivalent to the I-divergence measure when the original array m^0_{ijk} over the index set Γ consists of ones. The problem of minimizing the I-divergence measure subject to various marginal-sum constraints is the *multi-proportional adjustment problem.*

(ii) *Three-dimensional* χ^2 *measure*

$$d[\mathbf{M}, \mathbf{M}^0] = \tfrac{1}{2} \sum_{(i,j,k)\in\Gamma} (m_{ijk} - m^0_{ijk})^2/m^0_{ijk} \tag{29}$$

The problem of minimizing the χ^2 measure subject to various marginal-sum constraints is the *multidimensional Friedlander adjustment problem.*

(iii) *Modified three-dimensional* χ^2 *measure*

$$d[\mathbf{M}, \mathbf{M}^0] = \tfrac{1}{2} \sum_{(i,j,k)\in\Gamma} (m_{ijk} - m^0_{ijk})^2/m_{ijk} \tag{30}$$

It is understood that $(0 - m^0_{ijk})^2/0 = 0$ if $m^0_{ijk} = 0$, and $(0 - m^0_{ijk})^2/0 = +\infty$ if $m^0_{ijk} > 0$. The problem of minimizing the modified χ^2 measure subject to various marginal-sum constraints is the *modified multidimensional Freidlander adjustment problem.*

In the next section, the basic 3F problem and its variants, considered as nonlinear mathematical programming problems (primal problems), are converted into equivalent dual problems by the duality correspondence, in order to derive the solution algorithms more easily. In establishing the duality results and the solution algorithms for our basic adjustment problem, we rely on results obtained by Rockafellar (1970) in the field of "perturbation" functions and separable programming. We assume that there exists a strictly positive feasible solution [$m_{ijk} > 0$ for all $(i, j, k) \in \Gamma$] which satisfies the constraints (19)–(26).

We prove the existence of a strictly positive feasible solution for the case in which all migration flows are possible in Appendix A. In this case we also estimate the number of people remaining in the same region; the same proof applies however when migration flows from one region to the same region are impossible. For the case in which no strictly positive feasible solution exists, we can establish an asymptotic duality result and prove the convergence of an iterative solution procedure for the I-divergence minimizing problem.

3 SOLUTION OF THE MULTIPROPORTIONAL AND THE MODIFIED FRIEDLANDER (QUADRATIC) ADJUSTMENT PROBLEMS

In order to solve the multiproportional and the modified multidimensional Friedlander adjustment problems, we first derive the corresponding dual problems, for whose solution simple algorithms have been developed. A solution algorithm for the primal multiproportional adjustment problem is also given.

3.1 Duality Results

Before deriving the dual problems, we formulate a number of theorems that demonstrate the existence and uniqueness of the optimal solutions. Proofs of these theorems are given in Appendix A.

Theorem 1

If the solution set defined by constraints (19)–(27) contains a feasible solution $\mathbf{M} = [m_{ijk}]$, such that

$$m_{ijk} > 0 \qquad (i, j, k) \in \Gamma$$

where Γ is the index set for which m_{ijk} is not fixed, then both the multiproportional and the modified multidimensional Friedlander adjustment problems have unique optimal solutions. Further, the optimal solutions are strictly positive for all indices $(i, j, k) \in \Gamma$.

In order to establish the duality results, we introduce some new notation:

$$I(j, k) = \{i \in I \mid (i, j, k) \in \Gamma\} \qquad (\forall j \in J, k \in K)$$

$$J(i, k) = \{j \in J \mid (i, j, k) \in \Gamma\} \qquad (\forall i \in I, k \in K)$$

$$K(i, j) = \{k \in K \mid (i, j, k) \in \Gamma\} \qquad (\forall i \in I, j \in J)$$

Further,

$$\mathbf{\Lambda} = (\lambda_{ij})_{i,j=1}^{n,m}$$

$$\mathbf{N} = (\nu_{ik})_{i,k=1}^{n,l}$$

$$\mathbf{H} = (\zeta_{jk})_{j,k=1}^{m,l}$$

are real-valued matrices denoting the Lagrangian multipliers for the face constraints.

(i) *Unconstrained dual multiproportional adjustment problem*

Minimize

$$L_1(\mathbf{\Lambda}, \mathbf{N}, \mathbf{H}) = \sum_{(i,j,k)\in\Gamma} m_{ijk}^0 \exp[-(1 + \lambda_{ij} + \nu_{ik} + \zeta_{jk})] + \sum_{i=1}^{n}\sum_{j=1}^{m} \bar{c}_{ij}\lambda_{ij} + \sum_{i=1}^{n}\sum_{k=1}^{l} \bar{b}_{ik}\nu_{ik} + \sum_{j=1}^{m}\sum_{k=1}^{l} \bar{a}_{jk}\zeta_{jk} \tag{31}$$

subject to

$$\lambda_{11} = 1$$

$$\lambda_{ij}, \nu_{ik}, \zeta_{jk} \in R \qquad (\forall i \in I, j \in J, k \in K)$$

where

$$\bar{c}_{ij} = c_{ij} - \sum_{k \notin K(i,j)} m^{0}_{ijk}$$

$$\bar{b}_{ik} = b_{ik} - \sum_{j \notin J(i,k)} m^{0}_{ijk}$$

$$\bar{a}_{jk} = a_{jk} - \sum_{i \notin I(j,k)} m^{0}_{ijk}$$

Theorem 2

Let the multiproportional adjustment problem have an optimal solution $\hat{\mathbf{M}} = (\hat{m}_{ijk})$ such that

$$\hat{m}_{ijk} > 0 \qquad (i, j, k) \in \Gamma$$

If $(\bar{\Lambda}, \bar{\mathbf{N}}, \bar{\mathbf{H}})$ is an optimal solution of the dual problem (31), then

$$\hat{m}_{ijk} = m^{0}_{ijk} \exp[-(1 + \bar{\lambda}_{ij} + \bar{\nu}_{ik} + \bar{\zeta}_{jk})]$$

(ii) *Dual modified multidimensional Friedlander adjustment problem*
Minimize

$$L_2(\Lambda, \mathbf{N}, \mathbf{H}) = -2 \sum_{(i,j,k) \in \Gamma} m^{0}_{ijk}(1 + \lambda_{ij} + \nu_{ik} + \zeta_{jk})^{1/2} + \sum_{i=1}^{n} \sum_{j=1}^{m} \bar{c}_{ij}\lambda_{ij} + \sum_{i=1}^{n} \sum_{k=1}^{l} \bar{b}_{ik}\nu_{ik} + \sum_{j=1}^{m} \sum_{k=1}^{l} \bar{a}_{jk}\zeta_{jk} + 2|\Gamma| \tag{32}$$

subject to

$$\lambda_{11} = 1$$

$$\lambda_{ij}, \nu_{ik}, \zeta_{jk} \in R \qquad (\forall i \in I, j \in J, k \in K)$$

$$\lambda_{ij} + \nu_{ik} + \zeta_{jk} > -1$$

Theorem 3

Let the modified multidimensional Friedlander adjustment problem have an optimal solution $\hat{\mathbf{M}} = (\hat{m}_{ijk})$ such that

$$\hat{m}_{ijk} > 0 \qquad (i, j, k) \in \Gamma$$

If $(\bar{\Lambda}, \bar{\mathbf{N}}, \bar{\mathbf{H}})$ is an optimal solution of the dual problem (32) then

$$\hat{m}_{ijk} = \frac{m^{0}_{ijk}}{(1 + \bar{\lambda}_{ij} + \bar{\nu}_{ik} + \bar{\zeta}_{jk})^{1/2}}$$

3.2 Solution Algorithms

The algorithms presented here solve the multiproportional and the modified multidimensional Friedlander adjustment problems for the basic three-face (3F) case. The solutions for the special cases (3E), (1FE), and (2F) may be obtained in a similar way.*

3.2.1 Multiproportional Adjustment Problem

Algorithm for solving the unconstrained dual problem (31)

Step 0

Set

$S(\text{step}) = 0$

$\lambda_{ij}^{(0)} = 1$

$\nu_{ik}^{(0)} = 1 \qquad (\forall i \in I, j \in J, k \in K)$

$\zeta_{jk}^{(0)} = 1$

Step 1

Set

$\lambda_{11}^{(S+1)} = \lambda_{11}^{(S)} = 1$

$$\lambda_{ij}^{(S+1)} = \ln \sum_{k \in K(i,j)} m_{ijk}^{0} \exp[-(1 + \nu_{ik}^{(S)} + \zeta_{jk}^{(S)})] - \ln c_{ij}$$

for all $i \in I$, $j \in J$, except for $i = j = 1$.

Step 2

Set

$$\nu_{ik}^{(S+1)} = \ln \sum_{j \in J(i,k)} m_{ijk}^{0} \exp[-(1 + \lambda_{ij}^{(S+1)} + \zeta_{jk}^{(S)})] - \ln b_{ik}$$

Step 3

Set

$$\zeta_{jk}^{(S+1)} = \ln \sum_{i \in I(j,k)} m_{ijk}^{0} \exp[-(1 + \lambda_{ij}^{(S+1)} + \nu_{ik}^{(S+1)})] - \ln a_{jk}$$

If

$|\lambda_{ij}^{(S+1)} - \lambda_{ij}^{(S)}| \le \varepsilon$

$|\nu_{ik}^{(S+1)} - \nu_{ik}^{(S)}| \le \varepsilon \qquad (\forall i \in I, j \in J, k \in K)$

$|\zeta_{jk}^{(S+1)} - \zeta_{jk}^{(S)}| \le \varepsilon$

is fulfilled then STOP, otherwise $S \leftarrow S + 1$, and go to Step 1.

---

* If the elements m_{ijk}^{0} are uniformly distributed, then the special cases of the multiproportional adjustment problem become entropy problems with closed-form solutions (Appendix B).

Theorem 4

Let $\mathbf{M} = (m_{ijk})$ be a feasible solution of the basic problem such that

$$m_{ijk} > 0 \qquad (i, j, k) \in \Gamma$$

Then the above algorithm converges to an optimal solution of the unconstrained dual problem (31).

Using the result of Theorem 2, we can easily derive the direct primal algorithm.

Algorithm for solving the primal multiproportional adjustment problem

Step 0

Set

$$m_{ijk}^{(0)} = m_{ijk}^{0} \qquad (\forall i \in I, j \in J, k \in K)$$
$$S = 0$$

In the entropy problem, the *a priori* distribution m_{ijk}^{0} is not known and may therefore be set uniformly equal to unity, i.e., $m_{ijk}^{0} = 1$, $\forall i, j, k$.

Step 1

Set

$$m_{ijk}^{(3S+1)} = m_{ijk}^{(3S)} c_{ij} \Big/ \sum_{k=1}^{l} m_{ijk}^{(3S)}$$

for all $i \in I$, $j \in J$

Step 2

Set

$$m_{ijk}^{(3S+2)} = m_{ijk}^{(3S+1)} a_{jk} \Big/ \sum_{i=1}^{n} m_{ijk}^{(3S+1)}$$

Step 3

Set

$$m_{ijk}^{(3S+3)} = m_{ijk}^{(3S+2)} b_{ik} \Big/ \sum_{j=1}^{m} m_{ijk}^{(3S+2)}$$

If

$$\left| \frac{m_{ijk}^{(3S+3)}}{m_{ijk}^{(3S+2)}} - 1 \right| \leq \varepsilon \qquad (i, j, k) \in \Gamma$$

is fulfilled then STOP, otherwise $S \leftarrow S + 1$, and go to Step 1.

3.2.2 *Modified Multidimensional Friedlander Adjustment Problem*

Algorithm for solving the dual problem (32)

Step 0

Set

$S = 0$

$\lambda_{ij}^{(0)} = 0$

$\nu_{ik}^{(0)} = 0 \qquad (\forall i \in I, j \in J, k \in K)$

$\zeta_{jk}^{(0)} = 0$

Step 1

Let $\lambda_{ij}^{(S+1)}$ be the solution of

$$\sum_{k \in K(i,j)} \frac{m_{ijk}^{0}}{[1 + \lambda_{ij}^{(S+1)} + \nu_{ik}^{(S)} + \zeta_{jk}^{(S)}]^{1/2}} = c_{ij}$$

for all $i \in I$, $j \in J$, except for $i = j = 1$.
(The left-hand side of this equation is a strictly monotonically decreasing function of the variable $\lambda_{ij}^{(S+1)}$ in the range $(0, +\infty)$. Therefore, there is a unique solution $\lambda_{ij}^{(S+1)}$. A possible method of solution is the Newton method.)

Step 2

Let $\nu_{ik}^{(S+1)}$ be the solution of

$$\sum_{j \in J(i,k)} \frac{m_{ijk}^{0}}{[1 + \lambda_{ij}^{(S+1)} + \nu_{ik}^{(S+1)} + \zeta_{jk}^{(S)}]^{1/2}} = b_{ik}$$

for $i \in I$, $k \in K$, and where $\lambda_{ij}^{(S+1)}$, $\zeta_{jk}^{(S)}$ are considered to be given.

Step 3

Let $\zeta_{jk}^{(S+1)}$ be the solution of

$$\sum_{i \in I(j,k)} \frac{m_{ijk}^{0}}{[1 + \lambda_{ij}^{(S+1)} + \nu_{ik}^{(S+1)} + \zeta_{jk}^{(S+1)}]^{1/2}} = a_{jk}$$

If

$|\lambda_{ij}^{(S+1)} - \lambda_{ij}^{(S)}| \leq \varepsilon$

$|\nu_{ik}^{(S+1)} - \nu_{ik}^{(S)}| \leq \varepsilon \qquad (\forall i \in I, j \in J, k \in K)$

$|\zeta_{jk}^{(S+1)} - \zeta_{jk}^{(S)}| \leq \varepsilon$

is fulfilled then STOP, otherwise $S \leftarrow S + 1$, and go to Step 1.

Theorem 5

Let $\mathbf{M} = (m_{ijk})$ be a feasible solution of the basic problem (3F) such that

$m_{ijk} > 0 \qquad (i, j, k) \in \Gamma$

Then the above algorithm for solving the dual program of the modified Friedlander adjustment problem converges to an optimal solution.

Because of the nature of the modified Friedlander adjustment problem, we could not find a direct primal algorithm. However, using the results of Theorem 3 we can compute the solution of the primal problem $\hat{m}_{ijk}$ for $(i, j, k) \in \Gamma$ from the dual solution using the formula

$$\hat{m}_{ijk} = \frac{m^{0}_{ijk}}{(1 + \bar{\lambda}_{ij} + \bar{\nu}_{ik} + \bar{\zeta}_{jk})^{1/2}}$$

4 VALIDITY ANALYSIS AND NUMERICAL ILLUSTRATIONS (USING AUSTRIAN AND SWEDISH DATA)

The techniques developed in the previous sections may be used to deduce detailed migration patterns from aggregate data. But how accurate are the estimates, and which method gives the best results under a given set of conditions? Although research has not yet proceeded far enough to give definite answers to these questions, this section attempts to answer them numerically.

The accuracy of the estimation procedures is examined by applying them to multiregional systems for which complete data sets exist. Using only the marginal conditions, the detailed migration flows are estimated and the estimates are then compared with the observed data. Austria and Sweden are among the few countries that make detailed migration data available. These data may be aggregated in various ways to yield different sets of marginal totals, which may in turn be used to simulate different levels of data availability. Estimates of the detailed flows are calculated using the entropy maximizing method and the quadratic adjustment method.

The "validity" test of the multiproportional and modified Friedlander adjustment methods is basically a comparison between the estimated and observed migration flows and a judgment on the "closeness" of both sets of values. The quality of an estimation procedure is determined by the accuracy with which it can replicate observed data. A key problem in validity analysis is the definition of a composite index that measures the "closeness" of two arrays. In this paper, two such indices are used: the χ^2, and the absolute percentage error. Note that the χ^2 is the distance measure used in the modified Friedlander method.

The χ^2 statistic measures the relative squared deviation between the estimates and the observed values

$$\chi^2 = \sum_{i,j,k} (m_{ijk} - m^{0}_{ijk})^2 / m_{ijk} \qquad m_{ijk} \neq 0 \tag{33}$$

The absolute percentage error (APE) measures the deviation in absolute terms

$$APE = \sum_{i,j,k} d_{ijk}$$

where

$$d_{ijk} = |m_{ijk} - m^{0}_{ijk}| / m^{0}_{ijk} \qquad m^{0}_{ijk} \neq 0 \tag{34}$$

The average absolute percentage error ($\overline{APE}$), also known as the relative mean deviation or the mean prediction error, is given by

$$\overline{APE} = \sum_{i,j,k} |m_{ijk} - m^{0}_{ijk}| \Big/ \sum_{i,j,k} m^{0}_{ijk} \qquad m^{0}_{ijk} \neq 0$$

Both measures give similar results, but the χ^2 statistic attaches a greater penalty to large deviations.

An important observation of the validity study was that the error is not uniformly distributed among the elements of the array. The relative error, expressed by both the χ^2 and the *APE*, is greatest among the small elements (i.e., minor flows). The reason is the small denominator in (33) and (34). This observation is consistent with results obtained in input–output analysis (e.g., Hinojosa 1978). In addition, we found that a small number of elements of the array are responsible for most of the error; most elements are in low-error categories.

Because of these observations, and to get a better idea about the composition of the overall squared or absolute percentage deviation, the error analysis is carried out for subgroups of migration flows. The subgroups are formed on the basis of age (three broad age categories are considered: 0–14, 15–64, 65+), and volume of migration flow (eleven size classes are distinguished: 0–199, 200–399, 400–599, . . . , 1800–1999, 2000+). Note that the size classes are in terms of *observed* flows and not in terms of the estimates. The error analysis is also carried out for twelve error categories ranging from less than 2% to greater than 100%.

4.1 Entropy Maximization

The Austrian migration data are taken from the 1971 census and compare the place of residence in 1966 with that in 1971. The data were kindly provided by Dr. M. Sauberer, of the Austrian Institute for Regional Planning. Various aggregations of the migration data were made, and the techniques presented in the previous sections were then used to reproduce the original flow matrix.

The multidimensional entropy maximization method is a special variant of the multiproportional adjustment method: the initial estimates of the elements of the array are set equal to any scalar value, for example, unity: $m_{ijk}^{0} = 1$, for all i, j, k, except if $i = j$.* In other words, it is assumed that initial estimates are only available for the diagonal elements.

4.1.1 The Three-face (3F) Problem

Suppose the problem is to deduce origin–destination migration flows disaggregated by age for Austria (four regions) from the available information on the flow matrix of the total population and on the age composition of the arriving and departing migrants for each region. The data are given in Table 1, and present the three faces of a cube (array), the content (elements) of which must be estimated. The estimates can only be obtained by iteration. The computer program solves the 3F problem using the direct primal algorithm (see the second algorithm in Section 3.2.1).

The results of the (3F) estimation procedure are shown in Table 2, together with the observed migration-flow data. The algorithm took only five iterations to converge (tolerance level 10^{-4}).

---

* Intraregional migration is assumed to be nonexistent (structural zero). The contingency tables obtained therefore belong to the category of the so-called "incomplete contingency tables" and the analysis is one of "quasi-independence" (Bishop *et al.* 1975, p. 177).

TABLE 1 Internal migration in Austria,[a] 1966–1971.

(a) Migration flow matrix of total population

To	From: East	South	North	West	Total
East	0	12564	10587	3091	26242
South	7460	0	4532	3543	15535
North	11471	7715	0	3629	22815
West	3272	7494	4158	0	14924
Total	22203	27773	19277	10263	79516

(b) Departures ("from") and arrivals ("to") by region and age

	Region: East		South		North		West		Total	
Age	To	From	To	From	To	From	To	From	To	From
0	1896	1783	1428	1909	1733	1445	985	905	6042	6042
5	1086	930	734	1115	981	917	577	416	3378	3378
10	2264	1597	1052	3662	2118	1568	1961	568	7395	7395
15	7424	4172	3212	8323	4801	5297	4595	2240	20032	20032
20	4582	4227	2806	4625	4391	3250	2558	2235	14337	14337
25	2612	2807	1883	2625	2757	1885	1438	1373	8690	8690
30	1122	1123	835	1062	1155	924	603	606	3715	3715
35	906	915	622	903	929	750	519	408	2976	2976
40	950	871	540	807	824	688	413	361	2727	2727
45	625	579	374	517	515	447	237	208	1751	1751
50	601	618	441	514	541	478	268	241	1851	1851
55	675	689	458	521	552	489	251	237	1936	1936
60	576	700	434	455	568	439	201	185	1779	1779
65	431	543	331	340	433	328	147	131	1342	1342
70	274	353	213	217	280	209	94	82	861	861
75	145	194	114	118	156	109	52	46	467	467
80	49	68	37	39	53	35	17	14	156	156
85	24	34	21	21	28	19	8	7	81	81
Total	26242	22203	15535	27773	22815	19277	14924	10263	79516	79516

[a] The nine Austrian federal states (Bundesländer) have been aggregated here into four regions, as follows: East = Wien, Niederösterreich, Burgenland; South = Steiermark, Kärnten; North = Salzburg, Oberösterreich; West = Tirol, Vorarlberg.

The estimates are generally very close to the observed values. The average absolute percentage error is 4.27%, a very low figure compared with that obtained using existing biproportional or two-dimensional entropy methods (Nijkamp and Paelinck 1974b, Hinojosa 1978). About half of the number of migration flows (number of cells in the migration table) and almost two-thirds of the migration volume are estimated with less than 4% error (Table 3b). About 69% of the total absolute percentage error is due to minor migration flows (less than 200 migrants) representing only 11% of the flow volume (Table 3a). A similar pattern is obtained if the χ^2 statistic is used. The error distribution is, however, more explicit. The minor flows account for 34% of the total χ^2 value.

Similar observations were obtained in applying the estimation procedure to Swedish data. For a system composed of eight regions and eighteen age groups, the average absolute percentage error is 6.32%. About 28% of the cells in the migration table and about half of the migration volume are estimated with an error of less than 4%. The lower level of accuracy as compared with the Austrian estimates is a consequence of the larger share of small flows in the Swedish table. Whereas 52% of the migration flows for Austria contain less than 200 migrants, this figure is 79% for Sweden. About 94% of the total absolute percentage error is due to these minor flows, which account for 72% of the total χ^2 value.

The contribution of minor flows to the overall error is further illustrated by the cross-classification of error categories and flow-size classes (Table 3c). Minor flows are concentrated in the larger error categories.

The importance of small flow values in the overall error raises an additional problem, namely, rounding. Since the most probable estimates are rounded to the nearest integer to represent the number of migrants, error due to rounding may be substantial in the case of minor flows. The error measures in this study do not take into account the effects of rounding. The value of m_{ijk} used in the calculation of the error statistics is the original estimate before rounding.

The effect of age on the error distribution was also investigated. The results are not shown, since no significant difference between the contribution of each age category to the overall error could be observed. This may be a consequence of the uniform age pattern of the migrants.

4.1.2 The Three-edge (3E) Problem

In the (3E) problem, it is assumed that the only known information is for the edges of the box, i.e., the total number of arrivals and departures disaggregated by region, and the migrant age structure at the national level. The data are shown in the row and column sums of Table 1(a) and in the last column of Table 1(b).

The entropy or most probable estimates of the migration flows by age are calculated directly using eqn. (B1) from Appendix B. The (3E) entropy method yields estimates with an average absolute percentage error of 31%. The error is not concentrated in the minor flows but is evenly distributed among the classes of flow size. The high inaccuracy of the estimates may be traced back to the lack of data. Estimating migration flows from a uniform distribution as an initial guess ($m^0_{ijk} = 1$), and with only three edges given as new marginals, yields a migration pattern in which the three variables, region of origin, region of destination, and age, are independent when taken all together or in pairs.

4.1.3 The One-face, One-edge (1FE) Problem

In the (1FE) problem we estimate the values of m_{ijk} if the total flow matrix c_{ij} and the age structure of the migrants at the national level u_k are given; the estimates are obtained from eqn. (B2) and are shown in Table 4. This may occur when the available *a priori* information consists of the total population and a model migration schedule (see the paper by Rogers and Castro in this volume). By introducing information on the total flow matrix, the average absolute percentage error drops by half, from 31 to 16%. The contribution of minor flows to the overall error increases

TABLE 2 Estimated (3F) and observed migration flows by age for four regions of Austria, 1966–1971.

Age	Total		Migration from east to East	South	North	West
0	1783	Est.	0	674	874	234
		Obs.	0	670	877	236
5	930	Est.	0	328	482	120
		Obs.	0	328	468	134
10	1597	Est.	0	483	821	293
		Obs.	0	537	828	232
15	4172	Est.	0	1351	2029	793
		Obs.	0	1280	2192	700
20	4227	Est.	0	1336	2246	645
		Obs.	0	1289	2231	707
25	2807	Est.	0	939	1477	392
		Obs.	0	910	1464	433
30	1123	Est.	0	378	596	149
		Obs.	0	368	588	167
35	915	Est.	0	295	493	127
		Obs.	0	293	480	142
40	871	Est.	0	280	474	117
		Obs.	0	312	438	121
45	579	Est.	0	204	306	69
		Obs.	0	222	289	68
50	618	Est.	0	229	314	75
		Obs.	0	241	312	65
55	689	Est.	0	261	346	81
		Obs.	0	280	331	78
60	700	Est.	0	259	373	67
		Obs.	0	269	357	74
65	543	Est.	0	203	290	51
		Obs.	0	210	280	53
70	353	Est.	0	131	189	33
		Obs.	0	138	182	33
75	194	Est.	0	71	105	18
		Obs.	0	74	101	19
80	68	Est.	0	24	37	6
		Obs.	0	26	35	7
85	34	Est.	0	13	18	3
		Obs.	0	13	18	3
Total	22203		0	7460	11471	3272

Age	Total		Migration from south to East	South	North	West
0	1909	Est.	882	0	556	471
		Obs.	853	0	575	481
5	1115	Est.	493	0	348	274
		Obs.	530	0	329	256
10	3662	Est.	1371	0	1075	1216
		Obs.	1342	0	1044	1276
15	8323	Est.	3800	0	2022	2501
		Obs.	3760	0	1950	2613
20	4625	Est.	2097	0	1324	1203
		Obs.	2081	0	1381	1163
25	2625	Est.	1187	0	781	656
		Obs.	1225	0	800	600
30	1062	Est.	465	0	333	263
		Obs.	464	0	346	252
35	903	Est.	392	0	281	230
		Obs.	408	0	276	219
40	807	Est.	409	0	223	174
		Obs.	411	0	240	156
45	517	Est.	277	0	140	100
		Obs.	269	0	149	99
50	514	Est.	256	0	147	111
		Obs.	263	0	134	117
55	521	Est.	289	0	133	99
		Obs.	296	0	132	93
60	455	Est.	246	0	133	76
		Obs.	253	0	137	65
65	340	Est.	185	0	100	55
		Obs.	190	0	102	48
70	217	Est.	118	0	64	35
		Obs.	121	0	65	31
75	118	Est.	63	0	35	19
		Obs.	65	0	36	17
80	39	Est.	22	0	11	6
		Obs.	22	0	12	5
85	21	Est.	11	0	7	3
		Obs.	11	0	7	3
Total	27773		12564	0	7715	7494

TABLE 2 *Continued.*

			Migration from north to							Migration from west to			
Age	Total		East	South	North	West	Age	Total		East	South	North	West
0	1445	Est.	764	402	0	280	0	905	Est.	250	352	303	0
		Obs.	814	363	0	268			Obs.	229	395	281	0
5	917	Est.	482	251	0	184	5	416	Est.	111	155	150	0
		Obs.	448	282	0	187			Obs.	108	124	184	0
10	1568	Est.	745	371	0	452	10	568	Est.	148	197	222	0
		Obs.	771	344	0	453			Obs.	151	171	246	0
15	5297	Est.	2888	1107	0	1302	15	2240	Est.	736	754	750	0
		Obs.	2892	1123	0	1282			Obs.	772	809	659	0
20	3250	Est.	1806	734	0	710	20	2235	Est.	678	736	821	0
		Obs.	1861	701	0	688			Obs.	640	816	779	0
25	1885	Est.	1029	466	0	390	25	1373	Est.	395	479	499	0
		Obs.	998	482	0	405			Obs.	389	491	493	0
30	924	Est.	492	241	0	191	30	606	Est.	165	216	226	0
		Obs.	485	255	0	184			Obs.	173	212	221	0
35	750	Est.	402	187	0	162	35	408	Est.	113	140	155	0
		Obs.	379	213	0	158			Obs	119	116	173	0
40	688	Est.	419	146	0	122	40	361	Est.	121	113	127	0
		Obs.	411	141	0	136			Obs.	128	87	146	0
45	447	Est.	277	101	0	68	45	208	Est.	71	69	69	0
		Obs.	275	102	0	70			Obs.	81	50	77	0
50	478	Est.	273	124	0	81	50	241	Est.	73	88	80	0
		Obs.	273	119	0	86			Obs.	65	81	95	0
55	489	Est.	304	114	0	71	55	237	Est.	82	82	72	0
		Obs.	295	114	0	80			Obs.	84	64	89	0
60	439	Est.	271	110	0	58	60	185	Est.	59	64	61	0
		Obs.	263	114	0	62			Obs.	60	51	74	0
65	328	Est.	204	83	0	41	65	131	Est.	42	46	43	0
		Obs.	198	84	0	46			Obs.	43	37	51	0
70	209	Est.	130	53	0	26	70	82	Est.	26	29	27	0
		Obs.	125	54	0	30			Obs.	28	21	33	0
75	109	Est.	67	27	0	14	75	46	Est.	15	16	16	0
		Obs.	66	27	0	16			Obs.	14	13	19	0
80	35	Est.	22	8	0	4	80	14	Est.	5	5	5	0
		Obs.	22	8	0	5			Obs.	5	3	6	0
85	19	Est.	11	6	0	2	85	7	Est.	2	3	2	0
		Obs.	11	6	0	2			Obs.	2	2	3	0
Total	19277		10587	4532	0	4158	Total	10263		3091	3543	3629	0

TABLE 3 Error analysis of (3F) migration estimates for Austria.

(a) Analysis by size of flow

Size of flow	Number of flows		Number of migrants		Cumulative absolute % error		χ^2	
	Total	%	Total	%	Value	%	Value	%
0-200	112	51.85	8452	10.63	1043	68.56	0.9121e 02	33.70
200-400	45	20.83	12742	16.02	241	15.81	0.5711e 02	21.10
400-600	20	9.26	9481	11.92	74	4.87	0.2255e 02	8.33
600-800	11	5.09	7687	9.67	73	4.81	0.4191e 02	15.49
800-1000	9	4.17	7705	9.69	36	2.36	0.1924e 02	7.11
1000-1200	3	1.39	3330	4.19	8	0.52	0.2466e 01	0.91
1200-1400	7	3.24	9075	11.41	25	1.63	0.1295e 02	4.78
1400-1600	1	0.46	1464	1.84	1	0.06	0.1074e 00	0.04
1600-1800	0	0.00	0	0.00	0	0.00	0.0000e 00	0.00
1800-2000	2	0.93	3811	4.79	7	0.43	0.4201e 01	1.55
2000+	6	2.78	15769	19.83	14	0.95	0.1887e 02	6.97
Total	216	100.00	79516	100.00	1522	100.00	0.2706e 03	100.00

(b) Analysis by error category

Error category	Percentage error[a]	Number of flows		Number of migrants		Average flow
		Total	%	Total	%	
1	0-2	46	21.30	24037	30.23	522.543
2	2-4	57	26.39	24756	31.13	434.316
3	4-6	31	14.35	13604	17.11	438.839
4	6-8	18	8.33	6463	8.13	359.056
5	8-10	12	5.56	4026	5.06	335.500
6	10-15	28	12.96	5021	6.31	179.321
7	15-20	10	4.63	798	1.00	79.800
8	20-30	10	4.63	650	0.82	65.000
9	30-40	3	1.39	158	0.20	52.667
10	40-60	1	0.46	3	0.00	3.000
11	60-100	0	0.00	0	0.00	0.000
12	100+	0	0.00	0	0.00	0.000
Total		216	100.00	79516	100.00	368.130

(c) Analysis by size of flow and error category

Size of flow	Error category												
	0-2	2-4	4-6	6-8	8-10	10-15	15-20	20-30	30-40	40-60	60-100	100+	Total
0-200	21	23	13	8	4	20	10	9	3	1	0	0	112
200-400	9	12	10	5	3	5	0	1	0	0	0	0	45
400-600	6	9	0	2	2	1	0	0	0	0	0	0	20
600-800	1	2	4	0	2	2	0	0	0	0	0	0	11
800-1000	2	4	0	2	1	0	0	0	0	0	0	0	9
1000-1200	1	2	0	0	0	0	0	0	0	0	0	0	3
1200-1400	1	3	3	0	0	0	0	0	0	0	0	0	7
1400-1600	1	0	0	0	0	0	0	0	0	0	0	0	1
1600-1800	0	0	0	0	0	0	0	0	0	0	0	0	0
1800-2000	0	2	0	0	0	0	0	0	0	0	0	0	2
2000+	4	0	1	1	0	0	0	0	0	0	0	0	6
Total	46	57	31	18	12	28	10	10	3	1	0	0	216

[a] Average absolute percentage error (relative mean deviation) = 4.27.

in importance, in particular if the squared deviation is used as a measurement of error (Table 5).

4.1.4 The Two-face (2F) Problem

Assume that the following two faces are known: the total flow matrix (c_{ij}) and the age structure of the arriving migrants disaggregated by region (a_{jk}). The estimates may be obtained directly by applying eqn. (B3). The additional information on the age composition of arrivals reduces the average absolute percentage deviation only slightly, from 16 to 12% (χ^2 drops from 3662 to 2006). However, the error distribution changes. Minor flows get a greater share of the total absolute percentage error (72% of total absolute percentage deviation and 28% of total χ^2 value).

Results similar to those reported for Austria were obtained for Sweden, where the number of regions is twice as large. The average absolute percentage errors in the (3F), (3E), (1FE), and (2F) problems are, respectively, 6.32, 34.58, 15.26, and 11.90%. The share of the minor flows in the total error was always much higher: in the Austrian case the minor flows accounted for 51–72% of the total percentage absolute deviation, while in the Swedish case they amounted to 90–94%. This may be explained in part by the proportion of the total number and volume of flows constituted by minor flows.

The cases of data availability considered here lead to a firm conclusion: expanding the data set not only reduces the estimation error, but also increases the implicit weight attached to minor flows. This is implicit in the error statistics used: as the deviations between estimates and observations decline, the effects of small denominators become more apparent. The error analysis carried out here may also be used to investigate the marginal value of information on migration. Not every subset of the migration data has the same impact on the quality of the estimates. Hence, in further research we should consider the question: what kind of information on the migration pattern is required in order to obtain estimates of the detailed flow with an acceptable minimum level of accuracy?

4.2 Quadratic (Modified Friedlander) Method

As before, we assume that no *a priori* information on the elements m^0_{ijk} is available. The elements are set equal to unity except for the diagonal; hence the objective function of the modified Friedlander problem becomes

$$\min d[m_{ijk}, m^0_{ijk}] = \tfrac{1}{2} \sum_{i,j,k} (m_{ijk} - 1)^2 / m_{ijk}$$

The modified Friedlander method is applied to the 3F problem only. The computer program solves the problem using the dual algorithm (see Section 3.2.2). The estimated values of the migration flows m_{ijk} are shown in Table 6. The algorithm required 60 iterations to converge (tolerance level 10^{-4}). The convergence is therefore much slower than in the entropy algorithm. In addition to slow convergence, rounding errors may cause problems (all variables are single precision). The value of the χ^2 statistic is equal to 1614 (Table 7), and lies between the values obtained for the (2F) and (3F) problems. The χ^2 value is given here for illustrative

TABLE 4 Observed and estimated (IFE) migration flows by age for four regions of Austria, 1966–1971.

			Migration from east to							Migration from south to			
Age	Total		East	South	North	West	Age	Total		East	South	North	West
0	1687	Est.	0	567	872	249	0	2110	Est.	955	0	586	569
		Obs.	0	670	877	236			Obs.	853	0	575	481
5	943	Est	0	317	487	139	5	1180	Est.	534	0	328	318
		Obs.	0	328	468	134			Obs.	530	0	329	256
10	2065	Est.	0	694	1067	304	10	2583	Est.	1168	0	717	697
		Obs.	0	537	828	232			Obs.	1342	0	1044	1276
15	5593	Est.	0	1879	2890	824	15	6997	Est.	3165	0	1944	1888
		Obs.	0	1280	2192	700			Obs	3760	0	1950	2613
20	4003	Est.	0	1345	2068	590	20	5008	Est.	2265	0	1391	1351
		Obs.	0	1289	2231	707			Obs.	2081	0	1381	1163
25	2426	Est	0	815	1254	358	25	3035	Est.	1373	0	843	819
		Obs	0	910	1464	433			Obs.	1225	0	800	600
30	1037	Est.	0	349	536	153	30	1298	Est.	587	0	360	350
		Obs	0	368	588	167			Obs.	464	0	346	252
35	831	Est.	0	279	429	122	35	1039	Est.	470	0	289	280
		Obs.	0	293	480	142			Obs.	408	0	276	219
40	761	Est.	0	256	393	112	40	952	Est.	431	0	265	257
		Obs	0	312	438	121			Obs.	411	0	240	156
45	489	Est.	0	164	253	72	45	612	Est.	277	0	170	165
		Obs.	0	222	289	68			Obs.	269	0	149	99
50	517	Est.	0	174	267	76	50	647	Est.	292	0	180	174
		Obs.	0	241	312	65			Obs.	263	0	134	117
55	541	Est.	0	182	279	80	55	676	Est.	306	0	188	182
		Obs	0	280	331	78			Obs.	296	0	132	93
60	497	Est.	0	167	257	73	60	621	Est.	281	0	173	168
		Obs.	0	269	357	74			Obs.	253	0	137	65
65	375	Est.	0	126	194	55	65	469	Est.	212	0	130	126
		Obs.	0	210	280	53			Obs.	190	0	102	48
70	240	Est.	0	81	124	35	70	301	Est.	136	0	84	81
		Obs.	0	138	182	33			Obs.	121	0	65	31
75	130	Est	0	44	67	19	75	163	Est.	74	0	45	44
		Obs.	0	74	101	19			Obs.	65	0	36	17
80	44	Est.	0	15	23	6	80	54	Est.	25	0	15	15
		Obs.	0	26	35	7			Obs.	22	0	12	5
85	23	Est.	0	8	12	3	85	28	Est.	13	0	8	8
		Obs.	0	13	18	3			Obs.	11	0	7	3
Total	22203		0	7460	11471	3272	Total	27773		12564	0	7715	7494

TABLE 4 *Continued*

Age	Total		Migration from north to East	South	North	West	Age	Total		Migration from west to East	South	North	West
0	1465	Est.	804	344	0	316	0	780	Est	235	269	276	0
		Obs.	814	363	0	268			Obs.	229	395	281	0
5	819	Est.	450	193	0	177	5	436	Est.	131	151	154	0
		Obs.	448	282	0	187			Obs.	108	124	184	0
10	1793	Est.	985	421	0	387	10	954	Est.	287	329	337	0
		Obs	771	344	0	453			Obs.	151	171	246	0
15	4856	Est.	2667	1142	0	1048	15	2585	Est.	779	893	914	0
		Obs.	2892	1123	0	1282			Obs.	772	809	659	0
20	3476	Est.	1909	817	0	750	20	1850	Est.	557	639	654	0
		Obs.	1861	701	0	688			Obs.	640	816	779	0
25	2107	Est.	1157	495	0	454	25	1122	Est.	338	387	397	0
		Obs.	998	482	0	405			Obs.	389	491	493	0
30	901	Est.	495	212	0	194	30	479	Est.	144	166	170	0
		Obs.	485	255	0	184			Obs.	173	212	221	0
35	721	Est.	396	170	0	156	35	384	Est.	116	133	136	0
		Obs.	379	213	0	158			Obs.	119	116	173	0
40	661	Est.	363	155	0	143	40	352	Est.	106	122	124	0
		Obs.	411	141	0	136			Obs.	128	87	146	0
45	424	Est.	233	100	0	92	45	226	Est	68	78	80	0
		Obs.	275	102	0	70			Obs	81	50	77	0
50	449	Est.	246	105	0	97	50	239	Est.	72	82	84	0
		Obs.	273	119	0	86			Obs.	65	81	95	0
55	469	Est.	258	110	0	101	55	250	Est.	75	86	88	0
		Obs.	295	114	0	80			Obs.	84	64	89	0
60	431	Est.	237	101	0	93	60	230	Est.	69	79	81	0
		Obs.	263	114	0	62			Obs.	60	51	74	0
65	325	Est.	179	76	0	70	65	173	Est.	52	60	61	0
		Obs.	198	84	0	46			Obs.	43	37	51	0
70	209	Est.	115	49	0	45	70	111	Est.	33	38	39	0
		Obs.	125	54	0	30			Obs.	28	21	33	0
75	113	Est	62	27	0	24	75	60	Est.	18	21	21	0
		Obs.	66	27	0	16			Obs.	14	13	19	0
80	38	Est.	21	9	0	8	80	20	Est.	6	7	7	0
		Obs.	22	8	0	5			Obs.	5	3	6	0
85	20	Est.	11	5	0	4	85	10	Est.	3	4	4	0
		Obs.	11	6	0	2			Obs.	2	2	3	0
Total	19277		10587	4532	0	4158	Total	10263		3091	3543	3629	0

TABLE 5 Error analysis of (IFE) migration estimates for Austria.

(a) Analysis by size of flow

Size of flow	Number of flows		Number of migrants		Cumulative absolute % error		χ^2	
	Total	%	Total	%	Value	%	Value	%
0–200	112	51.85	8452	10.63	3809	70.07	0.7952e 03	21.72
200–400	45	20.83	12742	16.02	774	14.24	0.6374e 03	17.41
400–600	20	9.26	9481	11.92	232	4.27	0.1920e 03	5.24
600–800	11	5.09	7687	9.67	208	3.83	0.2945e 03	8.04
800–1000	9	4.17	7705	9.69	106	1.96	0.1565e 03	4.27
1000–1200	3	1.39	3330	4.19	49	0.90	0.1751e 03	4.78
1200–1400	7	3.24	9075	11.41	141	2.59	0.7689e 03	21.00
1400–1600	1	0.46	1464	1.84	14	0.26	0.3530e 02	0.96
1600–1800	0	0.00	0	0.00	0	0.00	0.0000e 00	0.00
1800–2000	2	0.93	3811	4.79	3	0.05	0.1222e 01	0.03
2000 +	6	2.78	15769	19.83	99	1.83	0.6055e 03	16.54
Total	216	100.00	79516	100.00	5437	100.00	0.3662e 04	100.00

(b) Analysis by error category

Error category	Percentage error[a]	Number of flows		Number of migrants		Average flow
		Total	%	Total	%	
1	0–2	19	8.80	10024	12.61	527.579
2	2–4	12	5.56	4089	5.14	340.750
3	4–6	18	8.33	5945	7.48	330.278
4	6–8	4	1.85	5277	6.64	1319.250
5	8–10	12	5.56	4602	5.79	383.500
6	10–15	38	17.59	13163	16.55	346.395
7	15–20	28	12.96	14785	18.59	528.036
8	20–30	31	14.35	9764	12.28	314.968
9	30–40	20	9.26	7422	9.33	371.100
10	40–60	16	7.41	3523	4.43	220.188
11	60–100	10	4.63	748	0.94	74.800
12	100 +	8	3.70	174	0.22	21.750
Total		216	100.00	79516	100.00	368.130

(c) Analysis by size of flow and error category

Size of flow	Error category												
	0–2	2–4	4–6	6–8	8–10	10–15	15–20	20–30	30–40	40–60	60–100	100 +	Total
0–200	7	6	7	2	7	19	10	15	8	13	10	8	112
200–400	2	4	7	0	2	7	5	9	8	1	0	0	45
400–600	4	1	2	0	1	5	4	3	0	0	0	0	20
600–800	1	0	0	0	1	1	5	1	2	0	0	0	11
800–1000	2	0	1	0	0	3	1	2	0	0	0	0	9
1000–1200	1	0	0	0	0	0	1	0	1	0	0	0	3
1200–1400	1	0	1	0	0	2	1	0	0	2	0	0	7
1400–1600	0	0	0	0	0	1	0	0	0	0	0	0	1
1600–1800	0	0	0	0	0	0	0	0	0	0	0	0	0
1800–2000	1	1	0	0	0	0	0	0	0	0	0	0	2
2000 +	0	0	0	2	1	0	1	1	1	0	0	0	6
Total	19	12	18	4	12	38	28	31	20	16	10	8	216

[a] Average absolute percentage error (relative mean deviation) = 16.24.

purposes only. It is not an appropriate goodness-of-fit test for the modified Friedlander method, since it is not independent of the objective function used. As a consequence, an objective comparison of the quality of the entropy and the quadratic methods is not straightforward. Further research on the development of appropriate test statistics is needed.

5 CONCLUSIONS

In migration analysis it frequently occurs that data on migration flows disaggregated by regions of origin and destination do not exist for subgroups of the population (such as age groups, income classes, educational levels, etc.) and that one has to infer these detailed flows from some more-aggregate information. This paper presents and generalizes two classes of estimation methods with great potential for solving this type of problem.

The classes and individual techniques may formally be presented as mathematical optimization problems with nonlinear objective functions and linear constraints. The various techniques differ in the objective function used.

The first class, namely, the bi- and multiproportional adjustment methods, combines the entropy maximization and the I-divergence minimizing problems. The two formulations have much in common, although they were developed independently. This paper presents a generalized formulation integrating both estimation techniques. The entropy method may be considered a special case of the more general I-divergence method. In the entropy problem, no initial values of the elements to be estimated are available and they are therefore uniformly set equal to unity. A single-solution algorithm is developed for both problems. Which of the techniques should be used will depend on the data available. The entropy method is more suited to estimating detailed migration flows on the basis of aggregate information only; the I-divergence method lends itself to updating migration-flow tables.

In the quadratic adjustment problems, weighted squared deviations between estimates and initial guesses are minimized. The latter may be derived from outdated migration tables or may represent *a priori* information on the migration pattern. A disadvantage of the quadratic adjustment methods, among them the Friedlander method, is that the estimates are not always of the appropriate sign. To avoid such anomalies as negative gross out-migration flows, a modified Friedlander method is presented.

The validity of the various estimation procedures has been demonstrated using migration data referring to Austria and Sweden. The numerical illustration shows that the estimates obtained by the multidimensional entropy method are very close to the observed data in the (3F) case. The estimates in other cases were not so good, but it was assumed that fewer data were available. However, one remarkable observation was that the (2F) case yielded estimates that did not deviate much from those obtained in the (3F) case, even though the latter required much more initial data. The reason is the high age specificity of migration. There seems to be a threshold level of information about the age composition of migrants that is sufficient to yield good estimates of detailed flows: after this threshold is reached, having more *a priori* information on the age structure does not add significantly to the quality of the estimates. This may lead to some interesting further research.

TABLE 6 Observed and estimated (modified Friedlander) migration flows by age for four regions of Austria, 1966–1971.

			Migration from east to							Migration from south to			
Age	Total		East	South	North	West	Age	Total		East	South	North	West
0	1783	Est.	0	681	799	304	0	1909	Est.	921	0	628	360
		Obs.	0	670	877	236			Obs.	853	0	575	481
5	930	Est.	0	316	428	186	5	1115	Est.	503	0	413	199
		Obs	0	328	468	134			Obs.	530	0	329	256
10	1597	Est.	0	466	729	402	10	3662	Est.	1426	0	1194	1042
		Obs.	0	537	828	232			Obs.	1342	0	1044	1276
15	4172	Est.	0	1522	2156	494	15	8323	Est.	3388	0	1815	3120
		Obs.	0	1280	2192	700			Obs.	3760	0	1950	2613
20	4227	Est.	0	1321	2411	494	20	4625	Est.	2197	0	1143	1285
		Obs.	0	1289	2231	707			Obs.	2081	0	1381	1163
25	2807	Est.	0	915	1486	407	25	2625	Est.	1268	0	790	567
		Obs.	0	910	1464	433			Obs.	1225	0	800	600
30	1123	Est.	0	362	560	201	30	1062	Est.	470	0	388	204
		Obs.	0	368	588	167			Obs.	464	0	346	252
35	915	Est.	0	273	465	176	35	903	Est.	404	0	325	174
		Obs.	0	293	480	142			Obs.	408	0	276	219
40	871	Est.	0	252	471	148	40	807	Est.	439	0	234	134
		Obs.	0	312	438	121			Obs.	411	0	240	156
45	579	Est.	0	190	304	85	45	517	Est.	298	0	143	76
		Obs.	0	222	289	68			Obs.	269	0	149	99
50	618	Est.	0	223	300	95	50	514	Est.	265	0	162	87
		Obs.	0	241	312	65			Obs.	263	0	134	117
55	689	Est.	0	257	339	92	55	521	Est.	305	0	137	80
		Obs.	0	280	331	78			Obs.	296	0	132	93
60	700	Est.	0	253	375	72	60	455	Est.	258	0	132	65
		Obs.	0	269	357	74			Obs.	253	0	137	65
65	543	Est.	0	198	292	53	65	340	Est.	195	0	98	47
		Obs.	0	210	280	53			Obs.	190	0	102	48
70	353	Est.	0	128	191	34	70	217	Est.	124	0	62	30
		Obs.	0	138	182	33			Obs.	121	0	65	31
75	194	Est.	0	68	107	19	75	118	Est.	67	0	34	17
		Obs.	0	74	101	19			Obs.	65	0	36	17
80	68	Est.	0	23	38	6	80	39	Est.	24	0	10	5
		Obs.	0	26	35	7			Obs.	22	0	12	5
85	34	Est.	0	13	19	3	85	21	Est.	11	0	7	3
		Obs.	0	13	18	3			Obs.	11	0	7	3
Total	22203		0	7460	11471	3272	Total	27773		12564	0	7715	7494

TABLE 6 *Continued*

			Migration from north to							Migration from west to			
Age	Total		East	South	North	West	Age	Total		East	South	North	West
0	1445	Est.	678	445	0	322	0	905	Est.	297	302	306	0
		Obs.	814	363	0	268			Obs.	229	395	281	0
5	917	Est.	443	283	0	192	5	416	Est.	140	135	141	0
		Obs.	448	282	0	187			Obs.	108	124	184	0
10	1568	Est.	650	401	0	517	10	563	Est.	188	186	194	0
		Obs.	771	344	0	453			Obs.	151	171	246	0
15	5297	Est.	3481	835	0	981	15	2240	Est.	555	855	830	0
		Obs.	2892	1123	0	1282			Obs.	772	809	659	0
20	3250	Est.	1796	676	0	779	20	2235	Est.	589	809	837	0
		Obs.	1861	701	0	688			Obs.	640	816	779	0
25	1885	Est.	911	510	0	464	25	1373	Est.	433	459	482	0
		Obs.	998	482	0	405			Obs.	389	491	493	0
30	924	Est.	444	282	0	198	30	606	Est.	208	191	207	0
		Obs.	485	255	0	184			Obs.	173	212	221	0
35	750	Est.	362	220	0	168	35	408	Est.	141	129	139	0
		Obs.	379	213	0	158			Obs.	119	116	173	0
40	688	Est.	381	176	0	131	40	361	Est.	130	111	120	0
		Obs.	411	141	0	136			Obs.	128	87	146	0
45	447	Est.	253	119	0	75	45	208	Est.	74	66	68	0
		Obs.	275	102	0	70			Obs.	81	50	77	0
50	478	Est.	251	142	0	86	50	241	Est.	85	77	79	0
		Obs	273	119	0	86			Obs.	65	81	95	0
55	489	Est.	284	126	0	79	55	237	Est.	87	75	76	0
		Obs.	295	114	0	80			Obs.	84	64	89	0
60	439	Est.	253	122	0	64	60	185	Est.	65	59	60	0
		Obs.	263	114	0	62			Obs.	60	51	74	0
65	328	Est.	190	91	0	47	65	131	Est.	46	42	43	0
		Obs.	198	84	0	46			Obs.	43	37	51	0
70	209	Est.	121	58	0	30	70	82	Est.	29	26	27	0
		Obs.	125	54	0	30			Obs.	28	21	33	0
75	109	Est.	61	31	0	17	75	46	Est.	16	15	15	0
		Obs.	66	27	0	16			Obs.	14	13	19	0
80	35	Est.	20	9	0	5	80	14	Est.	5	4	5	0
		Obs.	22	8	0	5			Obs.	5	3	6	0
85	19	Est.	10	6	0	3	85	7	Est.	2	2	2	0
		Obs.	11	6	0	2			Obs.	2	2	3	0
Total	19277		10587	4532	0	4158	Total	10263		3091	3543	3629	0

TABLE 7 Error analysis of the modified Friedlander migration estimates for Austria.

(a) Analysis by size of flow

Size of flow	Number of flows		Number of migrants		Cumulative absolute % error		χ^2	
	Total	%	Total	%	Value	%	Value	%
0–200	112	51.85	8452	10.63	1355	54.16	0.1903e 03	11.79
200–400	45	20.83	12742	16.02	545	21.79	0.2963e 03	18.35
400–600	20	9.26	9481	11.92	153	6.11	0.9538e 02	5.91
600–800	11	5.09	7687	9.67	168	6.73	0.3418e 03	21.17
800–1000	9	4.17	7705	9.69	63	2.50	0.6438e 02	3.99
1000–1200	3	1.39	3330	4.19	51	2.02	0.1301e 03	8.06
1200–1400	7	3.24	9075	11.41	90	3.61	0.2404e 03	14.89
1400–1600	1	0.46	1464	1.84	1	0.06	0.3149e 00	0.02
1600–1800	0	0.00	0	0.00	0	0.00	0.0000e 00	0.00
1800–2000	2	0.93	3811	4.79	10	0.42	0.1237e 02	0.77
2000+	6	2.78	15769	19.83	65	2.60	0.2431e 03	15.06
Total	216	100.00	79516	100.00	2502	100.00	0.1614e 04	100.00

(b) Analysis by error category

Error category	Percentage error[a]	Number of flows		Number of migrants		Average flow
		Total	%	Total	%	
1	0–2	22	10.19	9606	12.08	436.636
2	2–4	37	17.13	10261	12.90	277.324
3	4–6	18	8.33	7370	9.27	409.444
4	6–8	30	13.89	10016	12.60	333.867
5	8–10	17	7.87	10995	13.83	646.765
6	10–15	25	11.57	7359	9.25	294.360
7	15–20	26	12.04	10460	13.15	402.308
8	20–30	33	15.28	12085	15.20	366.212
9	30–40	5	2.31	1064	1.34	212.800
10	40–60	2	0.93	68	0.09	34.000
11	60–100	1	0.46	232	0.29	232.000
12	100+	0	0.00	0	0.00	0.000
Total		216	100.00	79516	100.00	368.130

(c) Analysis by size of flow and error category

Size of flow	Error category: 0–2	2–4	4–6	6–8	8–10	10–15	15–20	20–30	30–40	40–60	60–100	100+	Total
0–200	10	22	7	15	6	13	16	17	4	2	0	0	112
200–400	3	9	5	5	4	5	4	9	0	0	1	0	45
400–600	3	2	3	5	3	3	0	1	0	0	0	0	20
600–800	1	1	1	2	0	1	1	3	1	0	0	0	11
800–1000	3	0	1	1	2	1	1	0	0	0	0	0	9
1000–1200	0	0	0	0	0	2	0	1	0	0	0	0	3
1200–1400	0	2	0	1	0	0	3	1	0	0	0	0	7
1400–1600	1	0	0	0	0	0	0	0	0	0	0	0	1
1600–1800	0	0	0	0	0	0	0	0	0	0	0	0	0
1800–2000	0	1	0	1	0	0	0	0	0	0	0	0	2
2000+	1	0	1	0	2	0	1	1	0	0	0	0	6
Total	22	37	18	30	17	25	26	33	5	2	1	0	216

[a] Average absolute percentage error (relative mean deviation) = 11.03

A number of topics that need more study have been mentioned in this paper. First, there is the question of how much we need to know in advance to generate acceptable or good estimates of migration flows for subgroups of the population. The answer depends on the homogeneity of the population structure with regard to variables underlying the classification scheme. For instance, the age curve of migrants is very homogeneous across populations. Limited information on the age composition may therefore give good estimates of the flows disaggregated by age.

A related research topic is the development of ways to improve the initial guesses m^0_{ijk}. One strategy is to derive the *a priori* distribution m^0_{ijk} from a behavioral migration model in which moves are explained on the basis of push and pull factors, intervening factors, and personal factors or migrant characteristics. This approach was applied by Nijkamp (1976) to estimate commuting flows. Another possible strategy may be to introduce *a priori* expert opinions on detailed migration patterns. Yet another approach is to identify a structure in the interaction flows to be estimated, to represent this structure by a specific mathematical expression, and to introduce it as additional *a priori* information to improve the quality of the estimates. In the entropy methods presented in this paper, it is assumed that the estimates m_{ijk} are independent. However, some kind of interdependence may be introduced without over-complicating the estimation procedure. Current research on model migration schedules is of particular relevance in this regard (Castro and Rogers 1979). The increased attention devoted to the analysis of contingency tables or cross-classified data may also be very useful (Bishop *et al.* 1975, Goodman 1978). In fact, the limited data available require implicit or explicit assumptions to be made about the interdependence of the cross-classified variables. In the 3F problem, pairwise interactions between variables are introduced in the estimation procedure through the marginal constraints. These interaction patterns are absent in the 3E problem, and only partially present in the 1FE problem. The potential of contingency-table analysis for improved estimation of complex migration tables has been explored by Willekens (1980).

A third topic for further research is the extension and application of the techniques presented in this paper for updating migration tables. In countries such as the United States, the United Kingdom, Canada, Japan, and France, censuses are the main sources of migration statistics. Changes in migration patterns during the period between censuses are therefore difficult to assess at a detailed level. The multiproportional adjustment method may be relevant for updating census migration tables using marginal constraints derived from aggregate data for the years between censuses.

Finally, we need to know objectively how well the techniques perform. It is necessary, therefore, to develop one or a set of appropriate statistics to measure in objective terms the "goodness-of-fit" of the methods presented. The χ^2 statistic is used most frequently. Related statistics are also used by Nijkamp and Paelinck (1974a), Hinojosa (1978), and others. However, the use of these test statistics poses some severe problems; for example their values are affected heavily by smaller flows. (See also Forslund and Schoettner 1979.)

The research reported in this paper has led to some potentially very useful estimation methods for migration analysis. It has also led to the formulation of a number of challenging research topics to improve our knowledge of migration patterns. The results presented here and the research priorities identified are not limited, however, to the study of migration. They are applicable to any kind of n-dimensional interaction tables, whether they are spatial (e.g., trip distribution, commuting), sectoral (input–output studies), or of some other type. It is not the nature of the data that is important, but their cross-classification; the methods may be applied to all cross-classified data or n-dimensional contingency tables regardless of their subject.

REFERENCES

Bacharach, M. (1970) Biproportional Matrices and Input–Output Analysis. London: Cambridge University Press.

Bartlett, M. S. (1935) Contingency table interactions. Journal of the Royal Statistical Society Series A(2):248–252.

Batty, M., and S. MacKie (1972) The calibration of gravity, entropy and related models of spatial interaction. Environment and Planning A 4:205–233.

Batty, M., and L. March (1976) The method of residues in urban modelling. Environment and Planning A 8:198–215.

Bishop, Y. M., S. E. Fienberg, and P. W. Holland (1975) Discrete Multivariate Analysis. Cambridge, Mass.: M.I.T. Press.

Castro, L., and A. Rogers (1979) Migration Age Patterns: I. Measurement and Analysis. WP-79-16. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Chilton, R., and R. Poet (1973) An entropy maximizing approach to the recovery of detailed migration patterns from aggregate census data. Environment and Planning A 5:135–146.

Drewe, P., and F. Willekens (1980) Maximum Likelihood Estimates of Age-Specific Migration Flows in the Netherlands. Delft Progress Report 5:92–111.

Evans, A. W. (1970) Some properties of trip distribution methods. Transportation Research 4:19–36.

Evans, A. W. (1971) The calibration of trip distribution models with exponential or similar cost functions. Transportation Research 5:15–38.

Evans, S. P., and H. R. Kirby (1974) A three-dimensional Furness procedure for calibrating gravity models. Transportation Research 8:105–122.

Forslund, P., and J. Schoettner (1979) Estimation of Migration Flows: a Validation of Entropy Solutions. WP-79-59. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Friedlander, D. (1961) A technique for estimating a contingency table, given the marginal totals and some supplementary data. Journal of the Royal Statistical Society, Series A:412–420.

Friedmann, R. (1978) Updating input coefficients. Zeitschrift für die Gesamte Staatswissenschaften 134:144–165 (in German).

Goodman, L. (1978) The Analysis of Categorical Data. Cambridge, Mass.: ABT Books, ABT Associates.

Hewings, G. J. D. (1977) Evaluating the possibilities for exchanging regional input–output coefficients. Environment and Planning A 9:927–944.

Hewings, G. J. D., and B. N. Janson (1980) Exchanging regional input–output coefficients: a reply and further comments. Environment and Planning A 12:843–854.

Hinojosa, R. C. (1978) A performance test of the biproportional adjustment of input–output coefficients. Environment and Planning A 10:1047–1052.

Hortensius (1970) Estimation of the Elements of a Table (unpublished manuscript in Dutch). The Hague: Central Planning Agency.

Hyman, G. M. (1969) The calibration of trip distribution models. Environment and Planning A 1:105–112.

Isard, W. (1960) Methods of Regional Analysis. Cambridge, Mass.: M.I.T. Press.

Jaynes, E. T. (1957) Information theory and statistical mechanics. Physical Review 106:620–630.

Kullback, S. (1959) Information Theory and Statistics. New York: Wiley.

Leontief, W. (1941) The Structure of the American Economy 1919–1939. New York: Oxford University Press.

MacGill, S. M. (1977) Theoretical properties of biproportional matrix adjustments. Environment and Planning A 9:687–701.

Nijkamp, P. (1975) Reflections on gravity and entropy models. Regional Science and Urban Economics 5:203–225.
Nijkamp, P. (1976) Spatial Mobility and Settlement Patterns: An Application of Behavioral Entropy. RM-76-45. Laxenburg, Austria: International Institute for Applied Systems Analysis.
Nijkamp, P. (1977) Gravity and Entropy Models: The State of the Art. Research Memorandum No. 1977-2. Amsterdam: Vrije Universiteit, Economische Fakulteit.
Nijkamp, P., and J. Paelinck (1974a) A dual interpretation and generalization of entropy maximizing models in regional science. Papers of the Regional Science Association 33:13–31.
Nijkamp, P., and J. Paelinck (1974b) Some methods for updating input–output tables, in Foundations of Empirical Economic Research. Rotterdam: Netherlands Economic Institute (mimeographed).
Philipov, D. (1978) Migration and settlement in Bulgaria. Environment and Planning A 10:593–617.
Renyi, A. (1970) Probability Theory. Amsterdam: North-Holland Publishing Company.
Rockafellar, R. T. (1970) Convex Analysis. Princeton, N.J.: Princeton University Press.
Shannon, C. E., and W. Weaver (1949) The Mathematical Theory of Communication. Urbana, Ill.: University of Illinois Press.
Stone, R., Ed. (1963) Input–Output Relations, 1954–1966. London: Chapman and Hall.
Tan, E. (1980) On the Estimation of Migration Flows by Migrant Categories. MA Thesis. Inter-university Program in Demography, Brussels.
Theil, H. (1967) Economics and Information Theory. Amsterdam: North-Holland Publishing Company.
Thumann, R. G. (1978) A comment on "evaluating the possibilities for exchanging input–output coefficients." Environment and Planning A 10:321–325.
Willekens, F. (1977) The Recovery of Detailed Migration Patterns from Aggregate Data: An Entropy Maximizing Approach. RM-77-58. Laxenburg, Austria: International Institute for Applied Systems Analysis.
Willekens, F. (1980) Entropy, Multiproportional Adjustment and Analysis of Contingency Tables. Forthcoming in Sistemi Urbani, Vol. 2, No. 2–3.
Willekens, F., and A. Rogers (1978) Spatial Population Analysis: Methods and Computer Programs. RR-78-18. Laxenburg, Austria: International Institute for Applied Systems Analysis.
Willekens, F., A. Pór, and R. Raquillet (1979) Entropy, Multiproportional, and Quadratic Techniques for Inferring Detailed Migration Patterns from Aggregate Data. WP-79-88. Laxenburg, Austria: International Institute for Applied Systems Analysis.
Wilson, A. (1967) A statistical theory of spatial distribution models. Transportation Research 1:253–269.
Wilson, A. (1970) Entropy in Urban and Regional Modeling. London: Pion Limited.

APPENDIX A: PROOFS OF THEOREMS 1–5

Proof of Theorem 1

Let S denote the set of feasible solutions to constraints (19)–(27). S is a nonempty, bounded, closed convex set in R_+^γ ($\gamma = |\Gamma|$; the cardinality of set Γ).

Since the objective functions of both the multiproportional and the modified multidimensional Friedlander adjustment problems are strictly convex over R_+^γ, we have a unique optimal solution $s^* \in S$ in both cases. s^* is a vector containing the elements of the set $\{m_{ijk} | (i, j, k) \in \Gamma\}$.

We now show that if one element $s_i^* = 0$ for some i, then a contradiction ensues. By assumption, we know that there exists $\hat{s} \in S$ such that $\hat{s}_i > 0$ for $0 \leqslant i \leqslant \gamma$.

By definition, the value of the modified three-dimensional χ^2 measure (xe: 30) is $+\infty$ at point s^* and finite at $\hat{s}$, which obviously contradicts the fact that s^* is an optimal value for the modified multidimensional Friedlander adjustment problem.

For the multiproportional adjustment problem, and since S is a convex set, we have

$$\lambda \hat{s} + (1-\lambda) s^* \in S \qquad (\forall 0 \leq \lambda \leq 1)$$

$$\lambda \hat{s} + (1-\lambda) s^* > 0 \qquad (\forall 0 < \lambda \leq 1)$$

Let

$$f(\lambda) = \sum_{i=1}^{\gamma} (\lambda \hat{s}_i + (1-\lambda) s_i^*) \ln \left[\frac{\lambda \hat{s}_i + (1-\lambda) s_i^*}{s_i^0} \right] \qquad 0 < \lambda < 1$$

where the vector s^0 represents the values of the set $\{m_{ijk}^0 | (i, j, k) \in \Gamma\}$ in the same way as the vector $s \in S$ represents the values of $\{m_{ijk} | (i, j, k) \in \Gamma\}$.

Consider now the derivative of $f(\lambda)$ for $0 < \lambda < 1$. We have

$$\frac{\partial f(\lambda)}{\partial \lambda} = \sum_{s_i^*=0} \hat{s}_i \ln \left(\frac{\lambda \hat{s}_i}{s_i^0} \right) + \sum_{s_i^* \neq 0} (\hat{s}_i - s_i^*) \ln \left[\frac{\lambda \hat{s}_i + (1-\lambda) s_i^*}{s_i^0} \right] + \sum_{i=1}^{\gamma} (\hat{s}_i - s_i^*)$$

Since

$$\sum_{s_i^* \neq 0} (\hat{s}_i - s_i^*) \ln \left[\frac{\lambda \hat{s}_i + (1-\lambda) s_i^*}{s_i^0} \right]$$

is bounded over the interval [0, 1] and

$$\lim_{\lambda \to 0} \sum_{s_i^*=0} \hat{s}_i \ln \left(\frac{\lambda \hat{s}_i}{s_i^0} \right) = -\infty$$

there exist such $0 < \lambda_0 < 1$ that

$$\frac{\partial f(\lambda)}{\partial \lambda} < 0 \qquad (\forall 0 < \lambda < \lambda_0)$$

which obviously contradicts the fact that s_i^* is an optimal solution.

Proof of Theorems 2 and 3

Before starting the proof of Theorems 2 and 3, we first summarize some features of convex analysis, as presented by Rockafellar (1970).

Let f be a closed, proper convex function on R^n (i.e., f is convex and lower semicontinuous and never assumes the value $-\infty$, although $+\infty$ is allowed). The effective domain of f is the convex set dom f = $\{x: f(x) < +\infty\}$. The conjugate function of $f(x)$ is denoted by $f^*(x^*)$ and is defined as

$$f^*(x^*) = \sup \{\langle x, x^* \rangle - f(x) | x \in \text{ri (dom f)}\}$$

where ri denotes the relative interior of the set dom f. The proof of the duality results will rely on the decomposition principle suggested by Rockafellar (1970).

First, we shall consider the problem of minimizing

$$f_1(x_1) + \cdots + f_n(x_n) \qquad x = (x_1, x_2, \ldots, x_n) \in R^n \tag{A1}$$

subject to

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$$

and

$$\boldsymbol{x} \geq 0$$

where $f_i(x_i)$ is a proper convex function and $\boldsymbol{b}$ is a given vector in R^n.

Theorem A1

If the infimum in the above problem is finite and there exists $\boldsymbol{x} \in R^n$ such that

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$$

and

$$\boldsymbol{x} > 0$$

(i.e., all the components of $\boldsymbol{x}$ are strictly positive), then by minimizing the convex function

$$w(\boldsymbol{\lambda}) = f_1^*(-\langle \boldsymbol{a}^{(n)}, \boldsymbol{\lambda}\rangle) + \cdots + f_n(-\langle \boldsymbol{a}^{(n)}, \boldsymbol{\lambda}\rangle) + \langle \boldsymbol{b}, \boldsymbol{\lambda}\rangle \qquad \boldsymbol{\lambda} \in R^n$$

(where f_i^* is conjugate to f_i and $\boldsymbol{a}^{(i)}$ represents the *i*th column of matrix $\mathbf{A}$), and using any member of the minimum set $\boldsymbol{\lambda}^* \in R^n$ of it, the following subproblems define an optimal solution to the primal problem

$$\min_{x_i} f_i(x_i) + x_i\langle \boldsymbol{a}^{(i)}, \boldsymbol{\lambda}^*\rangle \qquad (i = 1, 2, \ldots, n)$$

For the proof of Theorem A1, see Section 28 of Rockafellar (1970).

Both the I-divergence measure eqn. (28) and the modified three-dimensional χ^2 measure can be separated into single-valued convex functions of the variables m_{ijk}.

Since the constraints (19)–(27) are linear equations of variables m_{ijk}, we can decompose both the multiproportional and the modified multidimensional Friedlander adjustment problems as discussed above.

The conjugate of the functions

$$f(x) = \begin{cases} x \ln(x/a) & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ +\infty & \text{if } x < 0 \end{cases} \qquad a > 0 \qquad x \in R_+$$

and

$$g(x) = \begin{cases} (x-a)^2/x & \text{if } x > 0 \\ +\infty & \text{if } x \leq 0 \end{cases}$$

can be obtained in the form

$$f^*(x^*) = a\, e^{x^*-1} \qquad \forall x^* \in R$$

$$g^*(x^*) = \begin{cases} -2a\sqrt{1-x^*} + 2 & \text{if } x^* < 1 \\ +\infty & \text{if } x^* \geq 1 \end{cases} \qquad x \in R_+$$

It is now easy to see that the functions W for the multiproportional and for the modified multidimensional Friedlander problems will take the forms $L_1(\mathbf{\Lambda}, \mathbf{N}, \mathbf{H})$ and $L_2(\mathbf{\Lambda}, \mathbf{N}, \mathbf{H})$, respectively.

It is easy to prove that one of the dual variables can be fixed at any constant value in both dual problems without changing the optimal value of the objective function. Theorems 2 and 3 are direct corollaries of Theorem A1.

Proof of Theorems 4 and 5

Before starting the proof of Theorems 4 and 5, we must summarize a number of points from convex analysis.

The recession function of f is denoted by $f0^+$ and defined as

$$(f0^+)(\boldsymbol{x}) = \lim_{\alpha \downarrow 0} \alpha f(\alpha^{-1}\boldsymbol{x}) \quad \boldsymbol{x} \in \text{dom } f$$

If, for a nonzero vector, $\boldsymbol{x} \in \text{dom } f$, $(f0^+)(\boldsymbol{x}) \leq 0$, then vector $\boldsymbol{x}$ is known as the "direction in which f recedes" or the direction of recession of f.

In the following analysis, our attention will be focused on the properties of the parameterized nest of level sets

$$\text{lev}_\alpha f = \{\boldsymbol{x} | f(\boldsymbol{x}) \leq \alpha\} \qquad \alpha \in R$$

belonging to a given proper convex function f.

Let inf f denote the infimum of $f(\boldsymbol{x})$ as $\boldsymbol{x}$ ranges over R^n. For $\alpha = \inf f$, $\text{lev}_\alpha f$ consists of the points $\boldsymbol{x}$ where the infimum of f is attained. We call this level set the minimum set of f.

We shall need the following properties of the level sets:

Lemma 1

Let f be a closed proper convex function which has no direction of recession. The infimum of f is then finite and attained. Moreover, all the level sets $\text{lev}_\alpha f (\alpha \geq \inf f)$ are nonempty, closed, bounded convex sets.

For the proof of Lemma 1, see Theorem 27.1 of Rockafellar (1970).

To prove the special algorithms given in Section 3 it is necessary to formulate a general procedure for a class of differentiable, closed proper convex functions which have no direction of recession and converge to the minimum set.

Let $f(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m)$ be a closed proper convex function on R^n, where $\boldsymbol{x}_i \in R^{n_i}$ for all $1 \leq i \leq m$ and $\sum n_i = n$.

For a given $\boldsymbol{x}^0 \in \text{dom } f$, let us define the sequence of points $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots$ recursively, by letting $\boldsymbol{x}^{(k)}$ be one of the solutions of the minimizing problem

$$\min_{\boldsymbol{x}_l \in R^{n_l}} f(\boldsymbol{x}_1^{(s)}, \ldots, \boldsymbol{x}_{l-1}^{(s)}, \boldsymbol{x}_l, \boldsymbol{x}_{l+1}^{(s-1)}, \ldots, \boldsymbol{x}_m^{(s-1)}) \tag{A2}$$

where

$$k = ms + l \qquad 1 \leq l \leq m$$

Lemma 2

Let f be a closed, proper convex function on R^n as defined in the above algorithm. Further, let $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots,$ be the sequence generated by the algorithm.

If f is continuously differentiable at every $x \in \text{dom f}$ and has no direction of recession, then

$$\lim_{k\to\infty} f(x^{(k)}) = \inf f$$

and $x^{(1)}, x^{(2)}, \ldots$ is a bounded sequence and all its cluster points belong to the minimum set of f.

Proof

By definition of the sequence $x^{(1)}, x^{(2)}, \ldots$ we have

$$+\infty > f(x^{(0)}) \geqslant f(x^{(1)}) \geqslant, \ldots, \geqslant f(x^{(k-1)}) \geqslant f(x^{(k)}) \geqslant \alpha$$

Hence

$$\lim_{k\to\infty} f(x^{(k)}) = \alpha \geqslant \inf f \tag{A3}$$

From Lemma 1, it follows that the sequence $x^{(1)}, x^{(2)}, \ldots$ is bounded and $\inf f < +\infty$.

Let x_c denote a cluster point of the sequence $x^{(1)}, x^{(2)}, \ldots$. From (A3) we obtain

$$x_c \in \{x \mid f(x) = \alpha\} \tag{A4}$$

We shall assume that $\alpha > \inf f$ and demonstrate that this implies a contradiction. Since function f is differentiable over dom f, we have from eqn. (A2)

$$\left.\frac{\partial f(x_1, \ldots, x_l, \ldots, x_m)}{\partial x_{l_i}}\right|_{x=x^{(k)}} = 0 \qquad (\forall 1 \leqslant i \leqslant n_l)$$

for all l, i, and k, where $k = ms + l$ and $x_l = (x_{l1}, x_{l2}, \ldots, x_{ln_l}) \in R^{n_l}$. Hence

$$\left.\frac{\partial f(x)}{\partial x_i}\right|_{x=x_c} = 0 \qquad (\forall 1 \leqslant i \leqslant n) \tag{A5}$$

which contradicts the fact that $\alpha > \inf f$, because a necessary and sufficient condition for a given point x to belong to the minimum set is given by eqn. (A5).

Lemma 3

Let $L_1(\mathbf{\Lambda}, \mathbf{N}, \mathbf{H})$ and $L_2(\mathbf{\Lambda}, \mathbf{N}, \mathbf{H})$ be the functions defined in eqns. (31) and (32).

If there exists a feasible solution to the constraints (19)–(27) such that

$$m_{ijk} > 0 \qquad (i, j, k) \in \Gamma$$

then $L_1(\mathbf{\Lambda}, \mathbf{N}, \mathbf{H})$ and $L_2(\mathbf{\Lambda}, \mathbf{N}, \mathbf{H})$ are closed proper convex functions and have no direction of recession.

Proof

From the simple form of functions L_1 and L_2, it is almost a trivial matter to see that the function is a closed proper function over all values of variables $\mathbf{\Lambda}, \mathbf{N}, \mathbf{H}$.

By definition, the recession function of L_1 is given as

$$(L_1 0^+)(\mathbf{\Lambda}, \mathbf{N}, \mathbf{H}) = \lim_{\alpha \downarrow 0} \alpha L_1(\alpha^{-1}\mathbf{\Lambda}, \alpha^{-1}\mathbf{N}, \alpha^{-1}\mathbf{H})$$

Recall that λ_{11} is not treated as a variable but has a constant value of 1.

If there exist $(i, j, k) \in \Gamma$ such that

$$\lambda_{ij} + \nu_{ik} + \zeta_{jk} < 0 \qquad \text{if } i \neq 1 \text{ or } j \neq 1$$

$$\nu_{1k} + \zeta_{1k} < 0 \qquad \text{if } i = j = 1$$

then

$$(L_1 0^+)(\Lambda, \mathbf{N}, \mathbf{H}) = +\infty$$

In the case when $\nu_{1k} + \zeta_{1k} \geqslant 0$ for all $k \in K$, and $\lambda_{ij} + \nu_{ik} + \zeta_{jk} \geqslant 0$ for all $(i, j, k) \in \Gamma$, we have

$$(L_1 0^+)(\Lambda, \mathbf{N}, \mathbf{H}) = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij}\lambda_{ij} + \sum_{i=1}^{n} \sum_{k=1}^{l} b_{ik}\nu_{ik} + \sum_{j=1}^{m} \sum_{k=1}^{l} a_{jk}\zeta_{jk}$$

Since there exists a feasible solution to the basic problem such that

$$m_{ijk} > 0 \qquad (i, j, k) \in \Gamma$$

$$m_{ijk} = \overset{0}{m}_{ijk} \qquad (i, j, k) \notin \Gamma$$

we have

$$\sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij}\lambda_{ij} + \sum_{i=1}^{n} \sum_{k=1}^{l} b_{ik}\nu_{ik} + \sum_{j=1}^{m} \sum_{k=1}^{l} a_{jk}\zeta_{jk} = \sum_{(i,j,k)\in\Gamma} m_{ijk}(\lambda_{ij} + \nu_{ik} + \zeta_{jk})$$

Since

$$\lambda_{11} + \nu_{1k} + \zeta_{1k} > 0 \qquad k \in K$$

and since there exists $k \in K$ such that

$$(1, 1, k) \in \Gamma$$

we have

$$(L_1 0^+)(\Lambda, \mathbf{N}, \mathbf{H}) > 0$$

The proof for function $L_2(\Lambda, \mathbf{N}, \mathbf{H})$ can be derived in the same way as for function $L_1(\Lambda, \mathbf{N}, \mathbf{H})$. It is easy to see that the solution algorithms for our dual problems (31) and (32) are special cases of our general procedure.

The proof of Theorems 4 and 5 obviously follows from Lemmas 1, 2, and 3.

APPENDIX B: MULTIPROPORTIONAL SOLUTIONS FOR THE SPECIAL CASES (3E), (1FE), AND (2F)

Note that in the two-dimensional entropy problem excluding any cost factor, the most probable values of the elements of interaction tables can be obtained analytically and are given by the simple expression

$$m_{ij} = \frac{O_i I_j}{\sum_i I_i} = \frac{m_{i.} m_{.j}}{m_{..}}$$

The element $m_{i.}$ denotes the total number of departures out of i, and $m_{.j}/m_{..}$ is the proportion of all migrants that go to j. This proportion is independent of the region of origin. The most probable number of migrants from i to j is simply the geometric average of the numbers of departures from i and arrivals at j. In other words, the best estimates are obtained by multiplying the number of out-migrants by a fixed in-migration profile (multiplicative spatial interaction model).

In the multidimensional case, there are also some analytical solutions to the I-divergence (entropy maximization) problem. A necessary condition for these solutions to exist, however, is that the initial distribution is uniform and that the constraints correspond to (3E), (1FE), or (2F).

The Three-edge (3E) Problem

In the (3E) problem it is necessary to minimize (18) subject to (22)–(26). Let i denote the region of origin, j the region of destination, and k the age group. The solution to the (3E) problem is given by

$$m_{ijk} = u_k v_j w_i/(ST)^2$$

or

$$m_{ijk} = m_{.j.} m_{..k} m_{i..}/(m_{...})^2 \tag{B1}$$

where

$$m_{i..} = \sum_{j,k} m_{ijk}$$

$$m_{.j.} = \sum_{i,k} m_{ijk}$$

$$m_{..k} = \sum_{i,j} m_{ijk}$$

Note that the solution assumes that all three classifications are mutually independent (e.g., number of arrivals is independent of number of departures).

The One-face, One-edge (1FE) Problem

In the (1FE) problem it is necessary to minimize (18) subject to (21), (22), and (26). Suppose the given face consists of the flow matrix of the total population c_{ij} and the edge represents the age structure of the migrants at the national level. The best estimates of the flow matrices disaggregated by age are given by

$$m_{ijk} = c_{ij} u_k/ST$$

or

$$m_{ijk} = m_{ij.} m_{..k}/m_{...} \tag{B2}$$

where

$$m_{ij.} = \sum_{k} m_{ijk}$$

The ratio $m_{..k}/m_{...}$ is the national age composition of the migrants. It is applied to all values of the flow matrix $m_{ij.}$ and hence the age structure is uniform for all flows.

The solution (B2) is therefore a model describing the migration flow under the assumption of a uniform age profile for migrants, and may be referred to as the age-profile model. If the given face consists of the age composition of the departing migrants ($m_{i.k}$), and the edge is the number of in-migrants by region ($m_{.j.}$), the model is an in-migration profile model. Conversely, if the age structure of arriving migrants ($m_{.jk}$) is known, together with the total number of out-migrants by region ($m_{i..}$), we have an out-migration profile model. Which of the models yields the best results depends on the homogeneity of the categories considered with respect to the profile adopted. Since the real age profile of migrants is quite uniform, the age-profile model may well yield better estimates.

The Two-face (2F) Problem

In the (2F) problem it is necessary to minimize (18) subject to (19), (21), and (26). The solution is for given values of c_{ij} and a_{jk}.

$$m_{ijk} = c_{ij}a_{jk}/v_j$$

or

$$m_{ijk} = m_{ij.}m_{.jk}/m_{.j.} \qquad \text{(B3)}$$

The ratios $m_{.jk}/m_{.j.}$ are conditional probabilities. The solution therefore implies the assumption of conditional independence between the j (destination) and i (origin) classifications for every k (migrant category, e.g., age group). In eqn. (B3) the age structure of migrants is destination-specific. The ratio $m_{.jk}/m_{.j.}$ represents the age curve of in-migrants, and is independent of the region of origin.