

CONFERENCE OF EUROPEAN STATISTICIANS

Joint ECE-EUROSTAT work session on Demographic Projections
(Perugia, Italy, 3-7 May 1999)

Working paper No. 16 (Summary)

**MODELING APPROACHES TO THE INDIRECT ESTIMATION OF MIGRATION FLOWS
FROM ENTROPY TO EM**

Submitted by Frans Willekens of the Population Research Centre, University of Groningen,
The Netherlands

ABSTRACT

The paper presents probability models to recover information on migration flows from incomplete data. Models are used to predict migration and to combine data from different sources. The parameters of the model are estimated from the data by the maximum likelihood method. If data are incomplete, an extension of the maximum likelihood method, the EM algorithm, may be applied.

Two models are considered: the binomial (multinomial) model, which underlies the logit model and the logistic regression, and the Poisson model, which underlies the loglinear model, the log-rate model and the Poisson regression. The binomial model is viewed in relation to the Poisson model.

By way of illustration, the probabilistic approach and the EM algorithm are applied to two different missing data problems. The first problem is the prediction of migration flows using spatial interaction models. The probabilistic approach is compared to conventional methods, such as the gravity model and entropy maximization. In fact, spatial interaction models are particular variants of log-linear models. The second problem is one of unobserved heterogeneity. A mixture model is applied to determine the relative sizes of different migrant categories.

Keywords: migration, missing data, probability models, entropy, maximum likelihood, EM