

Working Paper No. 21  
16 May 2003

ENGLISH ONLY

UN STATISTICAL COMMISSION and  
UN ECONOMIC COMMISSION FOR EUROPE

STATISTICAL OFFICE OF THE  
EUROPEAN COMMUNITIES (EUROSTAT)

CONFERENCE OF EUROPEAN STATISTICIANS

Joint ECE-EUROSTAT Work Session on Population and Housing Censuses  
(Ohrid, The former Yugoslav Republic of Macedonia, 21-23 May 2003)

## **THE INTEGRATED CENSUS MICRODATA INITIATIVE (ICM-EUROPE): AN UPDATE**

Note by the ECE Population Activities Unit

### **ABSTRACT**

This note presents the most recent developments on a project to anonymise, integrate, and make accessible for research use microdata samples based on multiple rounds of population and housing censuses in a number of European countries. It complements the information contained in the paper entitled "Integrating European Census Microdata: a joint project of the ECE Population Activities Unit and the University of Minnesota Population Center, 2004-2008" presented at the 2003 meeting of the EuroStat Working Party on Demographic Statistics and Population and Housing Censuses (DOC. DEM/CEN/E4/3/03-7.5 EN). Readers are invited to consult that paper for more details on the project.

### **I. OBJECTIVES OF THE PROJECT**

1. The overarching aim of the project *Integrated Census Microdata Initiative (ICM-Europe)* is to develop and make accessible for research use anonymised data through an integrated micro- and meta- data infrastructure, based on multiple rounds of population and housing censuses in at least 13 European countries. Thus, this project will address one of the major structural weaknesses of European social science research – the absence of cross-national infrastructure for access to comparable census micro- and meta- data. By virtue of that, it will put European research on par with the American and Canadian social sciences, where census microdata have been available for decades and have resulted in a considerable volume of research.

2. This aim will be attained by working towards the following interrelated objectives:

- i) build the methodological foundations for the integration of the micro- and meta- data, by developing standard concepts, categories, classifications, nomenclatures and coding schemes;
- ii) integrate the micro- and meta- data from the participating countries for all available census rounds;
- iii) provide an access system that will allow researchers to obtain adequately documented custom extracts, tailored to their research needs.

3. The objectives will be achieved through the proposed integrated and internally coherent set of *networking*, *transnational access*, and *joint research* activities. Specifically, the *networking activities* will (a) ensure high quality and coherent management structure for ICM-Europe; (b) facilitate the dissemination of the results of the joint research activities undertaken as part of the project and make a wider audience of users aware of the research potential of the infrastructure access offered by ICM-Europe, thus enlarging the user-base; (c) help to master the expertise of the broadest possible range of researchers and other professionals, so as to ensure that the integration of census micro- and meta- data is performed in accordance with the highest standards; (d) contribute to the creation of an European research area, by establishing regular contacts between researchers coming from different countries and disciplinary backgrounds, who will be using comparable data. Through the *transnational access* activity, four inter-related services based on the integrated census micro- and meta- data infrastructure created as part of the *joint research* activity will be offered in a coherent manner. Finally, the project's joint research activity will develop the methodological foundations and implement a fundamental infrastructure for European social sciences – a transnational database that disseminates fully documented, custom tailored census microdata via the web.

4. The project will build on the experience of the Minnesota Population Center (MPC) of the University of Minnesota and the Population Activities Unit (PAU) of the United Nations Economic Commission for Europe, as well as on the expertise of a consortium of European statistical offices and academic institutions. MPC and PAU are among the pioneers in the creation of cross-nationally harmonized census micro-data samples.

5. Since 1999, MPC in collaboration with national statistical offices and international data archives, has been implementing the IPUMS-International project, which is an effort to inventory, preserve, harmonize, and disseminate census microdata from around the world. This project itself builds on the original Integrated Public Use Microdata Samples (IPUMS) project, which started in the early 1990s and created a high-precision individual-level database describing the characteristics of the US population between 1850 and 2000. It is the world's largest public-access individual-level database on a human population.

6. Starting in 1992, PAU carried out a project, which resulted in the assembly of a collection of anonymized and partially harmonized microdata samples, based on the 1990-round of censuses in 15 European and North American countries. The samples were designed so as to allow research on a wide range of issues related to population ageing (most of them include oversampling with age), as well as on other social phenomena. A common set of nomenclatures and classifications, derived on the basis of a study of census data comparability in Europe and North America based on the 1990-round of censuses, were adopted as standards for recoding.

## II. RECENT INSTITUTIONAL ARRANGEMENTS

7. In order to ensure an adequate funding for the project, a grant proposal was recently submitted to the European Commission within the 6th Framework Programme for Research, Technological Development and Demonstration (Research Infrastructures Action, Call A.1.: Transnational Access and Integrating Activities). The proposal was submitted on the behalf of a consortium, which includes the following partners:

*Centre d'Estudis Demogràfics, Barcelona, Spain (coordinator)*

*The Cathie Marsh Centre for Census and Survey Research, Manchester, United Kingdom*

*Zentrum fuer Umfragen, Methoden und Analysen, Mannheim, Germany*

*Nederlands Instituut voor Wetenschappelijke Informa-tiediensten, Amsterdam, The Netherlands*

*Minnesota Population Center, Minneapolis, Minnesota, United States of America*

*Population Activities Unit, Geneva, Switzerland*

8. In addition to the consortium members, the project will involve as subcontractors national statistical offices and/or research institutions from the following 13 participating countries: Austria, Belarus, Bulgaria, Czech Republic, France, Germany, Hungary, The Netherlands, Portugal, Romania, Slovenia and Spain. All these countries, with the exception of the Netherlands, have already signed the standard data access protocol. Greece has also signed this protocol. More countries are expected to join the project at a later stage. The text of the proposal is available to interested parties upon request.

9. Separately, the PAU was awarded limited funding from the U.S. National Institute on Aging, which will cover the costs of some of PAU's work related to the 2000-round of population and housing censuses, including the promotion of the idea to create in interested countries higher-density samples of older persons.

### III. PLANS FOR THE FUTURE

10. A decision of the European Commission on whether to fund the project is expected in early autumn. Subject to funding availability, project tasks will be apportioned among consortium partners and subcontractors in each participating country. Statisticians and social scientists with expertise in the field will help develop consensus on the integration of concepts and coding schemes to promote better comparability of results and to enhance connectivity and coherence, independently of the research priorities of potential users.

11. A *Consortium Board* will be constituted as the project's main decision-making body. It will comprise one representative of each of the consortium partners (while up to two persons from each consortium partner may attend the board meetings, they will have only one vote). The decisions of the board are to be taken by consensus and on an equal basis among the partners. The Consortium will also establish a *Scientific Council*, which will play a consultative role and help in steering the work of the project. In addition to the members of the *Consortium Board*, the *Scientific Council* will include experts from the countries participating in the project, as well as a limited number of prominent scientists who will offer expertise in their personal capacity.

12. Subject to funding availability, the first meeting of the *Scientific Council* will be convened in Geneva in the middle of 2004. Subsequently, the *Consortium Board* and the *Scientific Council* will meet annually.

13. Ultimately, the uniform data-access protocols negotiated with each national statistical office will allow approved researchers to access the database via the Internet using the MPC's state-of-the-art, fully-automated, web-based microdata extract system with a functioning mirror-site operated by the Centre d'Estudis Demogràfics (Barcelona). Researchers will obtain custom tailored extracts, by country, census, variable, sample density, and sub-population. Thus, exclusively national access systems will be expanded to all of Europe and beyond by means of two integrated infrastructures, which will service social scientists' needs for microdata and metadata access without charge via the Internet.

-----