



**Economic and Social
Council**

Distr.
GENERAL

CES/AC.71/2001/26 (SUM)
20 November 2000

Original: ENGLISH

**STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE**

**COMMISSION OF THE EUROPEAN
COMMUNITIES (EUROSTAT)**

CONFERENCE OF EUROPEAN STATISTICIANS

Joint ECE/Eurostat Meeting on the Management of Statistical Information Technology
(Geneva, Switzerland, 14-16 February 2001)

Topic (iv): Integration of statistical (survey) data with registers (administrative) data

**MANAGING STATISTICAL INFORMATION SYSTEMS WITH COMMINGLED SURVEY
AND ADMINISTRATIVE DATA: A U.S. CENSUS BUREAU EXAMPLE**

Submitted by U.S. Bureau of the Census ¹

INVITED PAPER

SUMMARY

1. The maturation of the information age has created new challenges. Governments are faced with rapidly changing economies and societies are forced to make far-reaching economic and social policy decisions. These decisions, however, are based on limited, static, and expensive survey data. At the same time, large administrative and survey data sets are available which are derived from data collected from households, business establishments and governmental entities. These data are not effectively used in the United States, because key issues have not been fully addressed.

2. The integration of administrative records and survey and census data offers a solution for generating timely statistics at much lower costs while reducing respondent burden. Profound advances in computing technology and record linkage accuracy have significantly increased the feasibility of expanded

¹ Prepared by Ron Prevost, Demographic Surveys Division.

uses of administrative records for statistical purposes that avoids repetitive and burdensome inquiries of the public. These advances coupled with spiraling costs of, and public resistance to, traditional data collection have increased the opportunity for significant benefits through using administrative records in data collection, estimation, and evaluation systems.

3. The existence of enormous amounts of personal information in electronic form and the Internet revolution both have added to heightened public awareness of and concern about information privacy and confidentiality. The management of activities that blend respondent information from multiple sources must weigh any, and all, privacy and confidentiality concerns against the potential benefits of such research. Statistical agencies that integrate administrative and survey data must also have a corporate approach to the development and application of policies ensuring that the privacy and confidentiality concerns of the public and their data providers are met.

4. The Census Bureau currently uses administrative records extensively in a wide variety of its economic and demographic programs. The economic census and survey programs rely on administrative records about businesses from the Internal Revenue Service (IRS), and, to a lesser degree, on information from the Social Security Administration (SSA) and the Bureau of Labor Statistics (BLS). The Census Bureau's economic censuses and surveys use information from administrative records to (a) define the universe of business establishments, (b) identify proxy information for both non-responding firms and for those firms not being surveyed in an attempt to manage respondent burden, (c) improve the quality of survey responses through editing and questionnaire design improvements, and (d) reduce the need to conduct an expensive pre-census industry classification operation.

5. While many of the economic census and survey program activities rely extensively upon administrative records relating to businesses, only a few program activities rely upon administrative records for individuals. Perhaps because of the perceived sensitivity and associated higher privacy costs associated with administrative records relating to individuals, their routine use in major program activities has been somewhat restricted. However, the Census Bureau is currently exploring expanded use of administrative records in a number of research activities designed to produce results that, if fruitful, can be used to implement major program activities.

6. In our dynamic economy, the Census Bureau is increasingly unsure about what it should be measuring to capture the underlying activity in the economy. To determine this, and to determine whether existing data collection programs are adequate, this research is critical. The Public Use Micro Sample (PUMS) program on the demographic side has had a tremendous impact on the usefulness of Census Bureau data for internal and external research. The need to safeguard confidentiality with the release of these data into general use has created a core of technical expertise within the Census Bureau.

7. Linked data, if used wisely, can reduce respondent burden, increase data quality, and enhance the information available to the federal, state, and local agencies that rely on Census Bureau data for decision making. These data also provide a valuable tool to the research community. E.g., European restricted-access linked data are inspiring new research programs. These new programs are focused on the interaction of employers and employees over time, the dynamics of social program participation, social organization, population demographics, and health care delivery, to name but a few. The Census Bureau needs the partnership of researchers in many disciplines from the scientific research community to improve linkage technology, to advance the understanding of the statistical properties of linked data, and to fully exploit the many different research opportunities in the social and behavioral sciences that are uncovered by access to such data.