

Editorial

Official statistics and microdata – access and confidentiality

National Statistical Offices (NSOs) around the world are recognising the importance for researchers to access unit record files, or microdata, for statistical research and analyses and for the development and monitoring of policies.

The challenge NSOs are faced with is to provide access to rich microdata without breaching or compromising the confidentiality of the information provided by individual respondents. With the proliferation of databases and advances in information technology to facilitate efficient and large scale data matching, the technique used in the past to remove names and addresses from microdata is proving inadequate to provide the needed protection of confidentiality of individual's information.

How do NSOs address this challenge?

The paper by Ahmad, Yoo and De Back provides an overview of microdata access in OECD countries using a 2006 OECD survey, having regard to the countries' differences in legislation, technology and attitudes towards confidentiality. The paper also notes that, until recently, there was widespread acceptance amongst statistical agencies that the risks posed by deliberate or accidental disclosure outweighed the potential benefits that could be gained from providing access to microdata. However, there has been a recent welcome change in many NSOs to take a less conservative stance towards providing access to microdata for statistical research and analyses. Welcoming this development, the authors caution that there is a real risk that NSOs may become marginalised if they are unable to meet the requirements of researchers in this increasingly important area of statistical data provision.

The paper by Tam, Farley-Larmour and Gare describes the more commonly used approach to microdata access in Australia as one that requires the microdata to be highly confidentialised before they are

released to users for statistical analyses. Whilst less highly confidentialised microdata are available for access, the output from these analyses are subject to automatic confidentiality protection and sample manual checking before they are released to users. Access to less highly confidentialised microdata is only possible under an on-site data laboratory arrangement where the outputs are subject to detailed manual checking.

The Goldmann paper argues that informed decision making requires both access to relevant data and a community of qualified researchers to conduct the analyses. Statistics Canada has therefore embraced a microdata access model that encompasses not only provision of microdata but active engagement with the Canadian social science research community. This engagement includes promotion of interdisciplinary research and facilitation of collaboration among researchers as well as a role in helping train the next generation of researchers in quantitative analysis of complex social-science data. For Statistics Canada the key microdata access mode is its network of Research Data Centres.

The paper by Bujnowska and Museux provides an overview of the release of European Union (EU) microdata sets for research purposes and also describes several important European Statistical System projects currently underway that are related to remote microdata access. Microdata access in the EU was only enabled legally in 2002 and continues to evolve with changes in the governance structure for statistical confidentiality in the European Statistical System. The authors note that remote access to data stored at a secure server elsewhere is favoured by researchers but that this is not yet available at the EU level.

Noting the large-scale shift in Statistics Netherlands from survey-based statistics to register-based statistics, the Hoeve paper highlights the opportunities to perform advanced research on microdata, including da-

ta linking. A secure microdata access environment is achieved in the Netherlands by ensuring controls of all inputs to and output from the microdata access system are detailed and information. Access is provided on-site in Statistics Netherlands premises and also remotely from special dedicated terminals set up in designated research institutes.

Statistics New Zealand, like the ABS, places strong emphasis on protecting confidentiality through confidentialisation of the microdata before its release to approved users. The paper by Ng notes the success to provide standard microdata files rather than to provide the time-consuming files specific to the needs of users, and aspires to provide a data laboratory system aiming to provide microdata sets with minimal or no modification to the values, with confidentiality protected by output checking and verification.

The paper by Richie provides interesting outlines of the microdata data access and security models developed by the UK Office of National Statistics (ONS).

These models form an overarching framework used to underpin the development and provision of a range of different microdata access channels in the ONS.

The Jarmin paper notes a benefit to NSOs by granting access to microdata is in the improvements in the quality and utility of data products through feedback from users. Exposing the conceptual and processing assumptions that are embedded in the Census Bureau's microdata constitutes a core element in the Census Bureau's commitment to quality and improves the publicly-available microdata and tabulations that are available to the broader community of users. Importantly, the paper notes that microdata access also potentially enables data linking not possible with aggregates – both cross-survey linkages and longitudinal linkages – thus improving the richness of the datasets and power of statistical analyses.

Siu-Ming Tam
Editor-in-Chief