International Collaboration for Microdata Access: Lessons from OECD work

Mariarosa Lunati
Presented by Jennifer Ribarsky
OECD Statistics Directorate

Group of Experts on National Accounts:
Measuring Global Production
Geneva, May 31 - June 2, 2017
Microdata for research and policy at the OECD - the users’ needs

• Strong demand for access to microdata
  Microdata analysis as a key analytical tool

• Broad range of OECD studies based on microdata
  e. g. Income inequality; Employment; Business dynamics; Trade; Innovation; Education and skills; Consumption taxes

• Several approaches to exploit microdata
  – Networks
  – Safe Centres
  – Public Use Files (PUFs)
  – Confidential files
  – Commercial microdata
OECD research and transnational access to official microdata

• **Access in multiple countries** needed for cross-country analysis

• **Challenges**
  - **Locating information** on microdata (coverage, variables, ..)
  - **Language barriers**
  - **Different accreditation procedures** (terms and length)
  - **Different access systems**
  - **Differences in content and structure of microdata**
How are microdata accessed by OECD analysts

- Choice based on type, source and planned use of the microdata

- Most frequent approaches
  - Networks
    - National researchers
    - Statisticians in NSOs
  - Public Use Files (PUFs)

- Other approaches
  - Confidential files (anonymised individual data)
  - Safe Centres

- Commercial microdata

- Special surveys (ex. BEEPS, PISA)
## Main challenges of different approaches to microdata access

<table>
<thead>
<tr>
<th></th>
<th>Advantages</th>
<th>Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>National teams of</strong></td>
<td>Access and analysis done by national experts</td>
<td>Coordination costs</td>
</tr>
<tr>
<td>researchers</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NSO statisticians</strong></td>
<td>Mainstreaming of new microdata indicators</td>
<td>Resource commitment by NSOs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PUFs</strong></td>
<td>Direct access</td>
<td>Lack of data harmonisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Confidential files</strong></td>
<td>Direct access</td>
<td>Lack of data harmonisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Safe centres</strong></td>
<td>Direct access</td>
<td>Travel costs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Commercial</strong></td>
<td>Direct access</td>
<td>Quality and costs of microdata</td>
</tr>
<tr>
<td><strong>microdata</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Special surveys</strong></td>
<td>Direct access</td>
<td>Coverage, costs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Microdata for research and policy at the OECD – addressing the needs

• In 2011, the CSSP created the OECD Expert Group for International Collaboration on Microdata Access

• 26 members: Australia, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, France, Greece, Germany, Hungary, Israel, Italy, Japan, Korea, Mexico, Netherlands, New Zealand, Norway, Slovenia, Sweden, Switzerland, Turkey, United Kingdom, United States, Eurostat
Mandate of the Expert Group

• Increasing coordination and communication between NSOs and other expert groups to adopt best practices, promote a common understanding and minimise duplication of work in the area of cross-border access to microdata.

• Establishing procedures for efficient cross-border access to microdata held by statistical institutes while respecting confidentiality constraints.

• Advising and making recommendations to Chief Statisticians on technical and non-technical issues.
Scope of the work

• Collaboration with other international groups and initiatives

• Areas of investigation
  – Common language
  – Legal and policy framework
    • Sanctions for breach of confidentiality of statistical data
  – Technical issues

Executive Summary and Final Report available on:
http://www.oecd.org/fr/std/microdata.htm
Main outputs of the Expert Group -1

• Recommended consolidated terminology in microdata access, and metadata standards

• Proposed standardised application process for microdata access

• Description of the process flow for microdata access and analysis of cost recovery models of microdata services
Main outputs of the Expert Group -2

• Analysis of applicability of sanctions in international microdata exchange
• Analysis of issues relating to exchange of administrative microdata
• Review of portfolio of access solutions and dissemination strategies
• Two helpful notions
  - Application of maturity models to guide changes
  - Development of the “circle of trust” concept
## Maturity modeling to guide change - Example: Metadata on microdata

<table>
<thead>
<tr>
<th>Maturity in (eg. metadata)</th>
<th>Naïve</th>
<th>Heroic</th>
<th>Embedded</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Indicator</strong></td>
<td>Some metadata is available in Word format on request</td>
<td>Some datasets have machine-readable metadata available on the NSI website</td>
<td>All datasets have SDMX/DDI metadata, published in a resource discovery service in an Open format</td>
</tr>
</tbody>
</table>
The “circle of trust”

- Country 1
  - Zone 1
  - NSI / NSA / data producer original microdata
  - Access on site → secure use files at safe centre /remote access
  - Access off site → scientific use files
  - Access off site → public use files / remote execution

- Country 2
  - Zone 2
  - Level 1: Accreditation to use official microdata for scientific

- Country 3
  - Zone 3
  - Level 2: Accreditation to use highly confidential microdata for scientific
  - Level 3: Only data producers
Thank you!

Report available at
http://www.oecd.org/std/microdata.htm