Research access to confidential data

David Thorogood
Eurostat

Improving safe research access to data

• Why increase the use of data by researchers?
• What are the barriers?
• Examples of good practice
• Can this be done for EU-level data?
• Current EU legislation
• New draft EU legislation
Why increase researcher use of detailed statistical data?

- To make the best use of the data
  - topic experts
  - better, more sophisticated, research methods
  - better policy decisions
  - better data
- But the priorities:
  - protect respondent confidentiality
  - be seen to have done so

Potential barriers

- Methodological and operational
- Legal
- Respondent perception
- Administrative
  - ‘...not been done before’
  - Loss of control
Methodological and operational

• Tried and tested methods exist:
  – protected microdata for release
  – controlled access to original data
  – submitted analysis commands
• Methods are not perfect but have been successfully used - they are safe

Legal

• Research access may simply not be allowed
• Old legislation: such use of the data not envisaged
• Solutions:
  – Change legislation
  – Re-interpretation of existing laws
Respondent perceptions

• Whatever is done should be acceptable
• Need for further research in this area
  – particularly in Europe

Administrative

• ‘it’s not been done before’
• Loss of control
  – potential personal risk
• Facilitating research may not be a high priority
• Cost
Examples of good practice

- US Census Bureau - Center for Economic Studies
- Statistics Netherlands - Centre for Research of Economic Micro-data
- European Community Household Panel
- UK ONS - Longitudinal Study

Common themes

- Communication: statistical institute, researchers and respondents
- On-site access seen as more appropriate for business data
- Geographical issues relating to access
- Need for on-going review
- It can be done
Role of Eurostat

- Statistical Office of the European Communities
- Does not generally collect data from individual respondents
- Produces statistics based on data supplied by statistical authorities in member states and other countries

Researcher access to EU-level data

- Why?
  - Increasingly EU economy rather than separate countries
    - Single Market
    - Euro zone
  - Companies increasingly operating across internal frontiers
  - EU policy initiatives
Researcher access to EU-level data

• Potential barriers
  – differing national legislation and procedures
  – who gives permission?
  – who knows best if data are sufficiently protected?

Current European legislation

• Regulation 1588/90
  – Eurostat as a trusted partner may receive confidential data from national authorities
  – decision on data release rests with national authority

• Regulation 322/97
  – Access to confidential data permitted for scientific research
  – with permission of national authority
Draft regulation on access to confidential data for scientific purposes

• Aims:
  – Clarify legal position
  – Simplify process of granting permission for access
  – Maintain high level of protection for confidential data
  – Work with guidance and permission from the national authorities

Draft regulation on access to confidential data for scientific purposes (2)

• Envisages both microdata release and on-site access
• For scientific purposes
  – Defines categories of researcher, research institution, and commissioning organisation
• Confidentiality agreements and contracts
• Specific datasets
Draft regulation on access to confidential data for scientific purposes (3)

- Some element of standing permission from national authorities
  - for social data
  - authorised categories of researchers
  - specific datasets
- Dispersed research centres for on-site access to data

Conclusions

- Need for research access at European level
- Build upon best practice
- Research into respondent attitudes
- Researchers must make their needs for data widely understood
- Statistical institutes need to consider innovative ways of fulfilling these needs