More microdata by Eurostat for scientific purposes

Frank Espelage – Eurostat
frank.espelage@ec.europa.eu
Contents

1. Current release situation
2. AES, SES, CIS
3. New regulation on access to microdata
4. Further surveys to be released soon
Current release situation

- 12 surveys covered by relevant regulation 831/2002 but only six currently released:
  - European Community Household Panel (ECHP)
  - European Union Statistics on Income and Living Conditions (EU-SILC)
  - Labour Force Survey (LFS)
  - Adult Education Survey (AES)
  - Structure of Earnings Survey (SES) – enterprise survey
  - Community Innovation Statistics (CIS) – enterprise survey

- More than 200 contracts per year, permanently increasing
- Mostly DVDs (anonymised data), some Eurostat Safe Centre (non-anonymised data, mainly enterprise surveys)
Adult education survey (AES)

- EU statistics on lifelong learning (together with LFS - educational variables - and CVTS - enterprise survey)
- 2006 AES: voluntary pilot between 2005-2008
- 2011 AES: compulsory survey under regulation
- AES now planned every 5 years
- Target group: people aged 25-64 living in private households
- Both rounds (29 countries): ~ 220 000 respondents
Adult education survey (AES)

Detailed information on participation in education and lifelong learning activities (formal, non-formal and informal learning), incl.

- Job-related activities
- Characteristics of learning activities
- Self-reported skills
- Modules on social and cultural participation
- Foreign language skills
- IT skills and
- Core variables related to main socio-demographic characteristics of the respondents (gender, age, education, occupation, economic activity etc.)
Adult education survey (AES)

Microdata access:

- Anonymised microdata on DVD
- Anonymisation criteria: technical variables dropped (sampling design), modalities grouped (nationality, country of birth, two highest level of education...)
- 2006 AES data for 28 countries
- 2011 AES data to be made available in 2013 (#countries ?)
Structure of earnings survey (SES)

- Main data source for information on the level of remuneration, individual characteristics of employees and their employers in Europe
- Since 2002 every four years
- Target group: enterprises with at least 10 employees in the areas of economic activity defined by
  - sections C to O (excluding L) of NACE Rev. 1.1 (2002, 2006)
  - sections B to S (excluding O) of NACE Rev. 2 (2010)
- First round 2002 (28 countries): ~ 7.9 mio records
- 2006 (30 countries)/2010 (32 countries): > 13 mio records
Structure of earnings survey (SES)

Information on
- Economic activity
- Enterprise size
- Location
- Number of employees of the local unit,
- Age, sex, occupation, education, length of service of the employee,
- Hourly, monthly, annual earnings etc.
Structure of earnings survey (SES)

Microdata access:

- Both DVD and Safe Centre access
- Anonymisation criteria on DVD: sensitive NACE-NUTS-SIZE combinations (less than 3 enterprises) recoded by collapsing the level of detail for NACE/NUTS/SIZE, citizenship removed, age in six age groups, etc.
- SES 2002 and 2006: 23 (DVD) and 24 countries (Safe Centre) respectively
- SES 2010: available via Safe Centre (24 countries), DVD being prepared (13 countries already confirmed) – June 2013
Community Innovation Statistics (CIS)

- Main data source for direct measurement of innovation activities in Europe
- Since 2004 full survey every 4 years, reduced survey every 2 years in between. From 2012 biennial
- Microdata transmission optional
- Target group: Enterprises with at least 10 employees in selected high innovative NACE activities
- CIS 2008 (30 countries): ~ 200 000 records transmitted
Community Innovation Statistics (CIS)

Information on

- Enterprises' product and process innovations
- From 2008 also organizational and marketing innovations
- Core statistics: innovation activities, cooperation, development, expenditures and turnover
- Additional other aspects (eco-innovation in 2008, skills and creativity in 2010)
- Classification of enterprises by economic activity, size classes, type of innovation activity
Community Innovation Statistics (CIS)

Microdata access:

- Both CD and Safe Centre access
- Anonymisation criteria on CD: perturbation of numerical variables, collapsing the level of detail for some variables
- 2000 to 2008 data (four waves): 14 to 16 countries (CD) and 20 to 23 countries (Safe Centre) respectively
- CIS 2010 data: Safe Centre (~ 20 countries): spring 2013, CD later after anonymisation (# countries ?)
New regulation on access to microdata

Aims:

- Better adaptation to the revised Statistical Law (Regulation on European statistics) from 2009
- Allow for new modes of access
- Offer more datasets
- Improve access procedures
- Avoid burdensome admissibility procedure (for entities not defined as research entities by old regulation)
- Increase flexibility and sustainability of the microdata access system
New regulation: process

1. Recognition of research entity (user)

- Application form, confidentiality undertaking and terms of use to be signed by the head of research entity
- Assessment by Eurostat

-> Publication of the name of the research entity on Eurostat website
New regulation: process

2. Research proposal

- Application form and confidentiality declaration to be signed by researchers having access to confidential data
- Consultation of the national authorities in Member States

-> Access
New regulation: impact

- Currently directly admissible bodies need to be recognised as research entity
- Higher responsibility of research entity hosting confidential data
- Shorter delays and smaller administrative burden once the confidentiality undertaking is signed – no contract amendments any more
- Entry into force expected July, old procedure to be stopped 1 April (cp. Eurostat Website)
Further surveys to be released soon

- New, faster and easier procedures will allow release of further surveys
- 6 pending surveys:
  - European Health Interview Survey (EHIS)
  - Community Statistics on Information Society (CSIS) – Module 2 Individuals / Households
  - Continuing Vocational Training Survey (CVTS)
  - Household Budget Survey (HBS)
  - Statistical returns in respect of the Carriage of Goods by Roads
  - Farm Structure Survey (FSS) - postponed
- EHIS, CSIS, CVTS probably from July onwards
European Health Interview Survey (EHIS)

- Data on the health status, life style (health determinants) and health care services use of EU citizens, plus their demographic and socio-economic status (incl. education)

- EHIS 2008 - voluntary survey. Mandatory with planned 5-year periodicity in the future – EHIS 2014?

- Target group: individuals aged 15+ living in private households

- EHIS 2008 (19 countries): ~ 100 000 individuals
European Health Interview Survey (EHIS)

Microdata access:

- Both DVD and Safe Centre access planned
- Anonymisation criteria on DVD: aggregation of AGE (mainly 5-year age groups), region, occupation, economic activity, big households etc.
- EHIS 2008 (17 countries)
Community Statistics on Information Society (CSIS) – Module 2 Individuals/Households

- Statistics on the use of Information and Communication Technologies at European level
- Target group: individuals aged 16-74; households with at least one person aged 16-74
- CSIS 2011: ~ 215,000 individuals, ~ 150,000 households
Community Statistics on Information Society (CSIS) – Module 2 Individuals/Households

Microdata access:

- Anonymised microdata on DVD
- Anonymisation criteria: mainly 10-year age groups
- Microdata for all available countries and survey years 2009 to 2012
Continuing vocational training survey (CVTS)

- Data on continuing vocational training in enterprises
- Target group: enterprises with 10 or more persons employed belonging to certain NACE categories
- CVTS 2005: 28 countries, 2010: 29 countries
Continuing vocational training survey (CVTS)

Microdata access:

- Anonymised microdata on DVD
- Final decision on anonymisation criteria: spring 2013
- At least CVTS 2005 and 2010 planned to be released
Final remark

- Ever more requests for use of more than one microdata collection for a given project
- All Eurostat units responsible for a survey are asked for assessment

-> Please explain for what purpose the individual data collections should be used and how!

Easier for Eurostat -> faster for all...
Further information

General information on microdata access:


Additional information for the individual surveys can also be found in the metadata files linked to their database tables on the Eurostat website.
Database links for AES, SES, CIS

AES:

SES:

CIS:
http://epp.eurostat.ec.europa.eu/portal/page/portal/science_technology_innovation/data/main_tables