# Administrative data and survey data on electricity use in Hungary

Akos Gerencser

Hungarian Energy and Public Utility Regulatory Authority

gerencsera@mekh.hu

Madrid, May 31 - June 3

# Institutional framework of energy statistics in Hungary

- **Hungarian Energy and Public Utility Regulatory Authority** (HEA) is the regulatory body of the energy and public utility market

- Part of the European Statistical System as **Other National Authority** (ONA) since 2012

- Sole responsibility for producing official energy statistics

- Legal background: Regulation (EC) No 1099/2008/EC of the European Parliament and Council on energy statistics
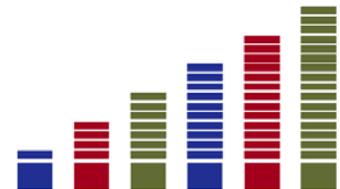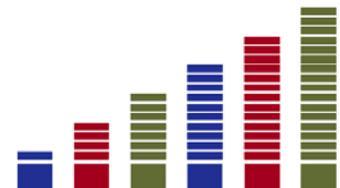
**Final energy consumption** is the total energy consumed by end users, such as households, industry and agriculture. It is the energy which reaches the final consumer's door and excludes that which is used by the energy sector itself. (EUROSTAT)

**Energy end user categories (groups according to NACE):**

- Private households

- Industry

- Transport

- Agriculture

- Commercial and Public Services

# Data sources as defined by UN ESCM

| | Advantages | Disadvantages |
|---|---|---|
| **Statistical Surveys** | • Comprehensive information on all fuels supplied and used<br>• Good data quality<br>• Can be used directly and as an input for model calculations<br>• Good response rates when surveys are covered by legislation | • Resource intensive and expensive<br>• Time consuming<br>• High survey burden<br>• If voluntary, response rates can be low<br>• Data validation required<br>• Reporting of non-metered fuels, often purchases not used |
| **Administrative data** | • Low survey burden<br>• Greater number of records allows more detailed breakdowns<br>• Avoids duplication by making use of existing data<br>• No sample error | • Dependency on third parties<br>• Definitions and information may not match statistical needs<br>• Often requires substantial effort to set up and may be legal barriers to use |
| **Modelling** | • Allows quantification of variables that which cannot be directly measured or observed<br>• Save resources (money and staff)<br>• Low survey burden<br>• Quick results<br>• Can be used to adapt or improve survey results<br>• Can be used to reduce survey frequency | • Worse data quality compared to surveys<br>• No Stand-alone methodology: cannot be calculated without input data<br>• Quality of results depend on accuracy of input data and the design of the model |
| **"In situ" measurements** | • Detailed information on individual appliances, information on patterns of use of the equipment<br>• High quality of the results<br>• Input data for surveys and/or modelling | • Invasive for respondents: difficulties in finding respondents willing to participate<br>• High burden in terms of time and human resources<br>• Expensive, so often small samples, and less representative<br>• Constraints in monitoring equipment: limitation in the number of metering devices and monitoring incidences |

# Statistical surveys on energy end-use

- Population: Energy non-household end-users
- Frame: General Business Register
- Stratified Random Sampling
- Strata according to NACE activity and size
- Cut-off threshold based on number of employees
- Specific questionnaires for sectors (industry, transport, agriculture, commercial and public services)
- Sample size: altogether ca. 7600 entities
- CAWI, significant IT developments, questionnaire re-design
- Improved estimation methodology
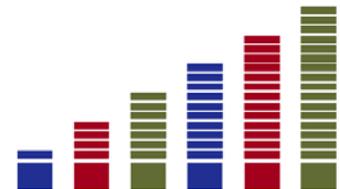
# Administrative data for electricity

- Electricity is generally delivered to end-users via the electricity distribution system

- Administrative records on distributed amounts by Distribution System Operators (DSO-s)

- Wide coverage

- Used traditionally for annual residential electricity consumption
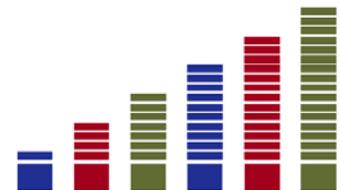
# Data linking exercise

- Idea: Matching DSO billing partner information with General Business Register via VAT-ID

| | No. of reported entities | % of matched entities | Total reported non-household MWh | % of matched MWh |
|---|---|---|---|---|
| DSO 1 | 19 907 | 98,54% | 2 909 620 | 99,94% |
| DSO 2 | 18 163 | 87,57% | 2 746 475 | 96,72% |
| DSO 3 | 29 220 | 98,66% | 5 463 128 | 99,77% |
| DSO 4 | 40 055 | 99,08% | 6 340 123 | 97,12% |
| DSO 5 | 11 274 | 99,00% | 4 181 231 | 97,94% |
| DSO 6 | 18 065 | 98,70% | 2 576 421 | 99,89% |
| TOTAL | 136 684 | 97,32% | 24 216 998 | 98,45% |
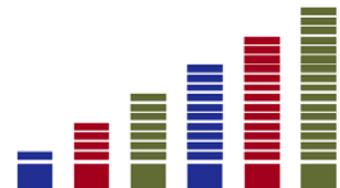
- The linking exercise can be considered as very successful. In the most NACE groups the bias is around or less than 20%

- Consumption is usually higher according to the statistical survey than in the admin data in the Industry and Transport sectors, while vice-versa for Commercial and public services
  - Re-sale of electricity by real estate transactions or in industrial parks distorting admin data
  - Self-generation and subsequent on-site use
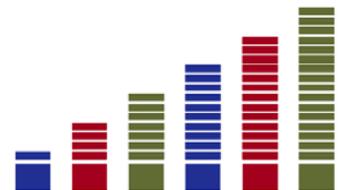  - Supply directly from the transmission system
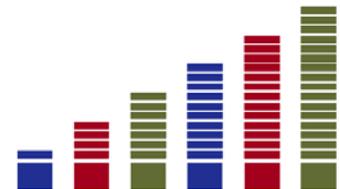
# Conclusion

- Statistical survey data are considered more reliable for producing NACE breakdown of electricity end-use

## HOWEVER

- Admin data are extremely useful:
    - Validity checking of stat survey reported data
    - Producing territorial breakdowns
    - Analysing the consumption of small and medium size enterprises
    - Calculating energy efficiency indicators

# Visualisation of admin data

Thank you for your attention!