Statistics on trade by enterprise characteristics
projects linking business statistics in the European Statistical System

Pekka Alajääskö
Eurostat

SESSION II: Linking of statistical business registers and trade statistics

Introduction

- European business statistics compilers face a dilemma:
  - users and policy makers demand additional information
  - budget constraints and reluctance to increase the burden on survey respondents and national statistical institutes put tight restraints on the extension of data requirements

- Micro data linking (MDL) can provide an opportunity to develop new statistics and indicators both when
  - using existing data sets
  - but also when combined with new data collections

Introduction

- Very visible outcomes of the MDL business statistics projects within the European Statistical System (ESS) are the regular compulsory productions of:
    - international trade in goods by enterprise characteristics statistics (TEC)
    - inward foreign affiliate statistics (IFATS)
- There are many other important MDL projects with interesting results
- They will be briefly presented with the view of improving statistics on entrepreneurship and in particular its international aspects

TEC development and EIP

- TEC is a statistical domain, which unlike traditional trade statistics, aims at describing the structure of trade by characteristics of the trading enterprises, for instance by their economic activities, their size or concentration of trade

- It is based on linking international trade in goods (ITGS) micro data with business register (BR) information

- TEC has been part of the regular data collection since 2009

- From the EIP perspective particularly interesting TEC statistics is the enterprise size class information – performance of small exporting enterprises

TEC development and EIP

- Information on young, small and high growth enterprises is the common and central theme for the EIP

- An even better indicator, originally identified by the EIP would have been the export performance of young enterprises

- The planned MDL project in 2016, extending the MDL of SBS with other business statistics will link ITGS and BD to produce statistics on export performance of young (and/or small) enterprises

- EIP indicator on Export performance:

  - currently one gets information on export performance of **small enterprises** from the TEC

  - In the future ITGS and BD data will be linked to produce data on export performance of **young enterprises**

MDL and EIP

- Current and planned (MDL) projects will take the development work done in the TEC framework forward to cover services enterprises i.e. international trade in services by enterprise characteristics (STEC)

- This will make it possible to produce EIP indicators for enterprises engaged in international trade in services (ITSS)

- Additional important development work carried out by the MDL SBS projects is the consistent linking of the enterprise population of ITGS with the SBS population

  - As a result one could compare and produce indicators on the performance and relative shares of traders and non-traders in the SBS population

  - Some of these indicators hopefully will be produced on a more regular basis (voluntary) in the SBS framework

MDL and EIP

- In 2016 it is also planned to extend the linking of ITGS and SBS enterprise level data with data of BD. This would facilitate the production of following EIP indicators:

- EIP indicator on value-added of young or small enterprises

  - Currently one gets information on value-added of **small enterprises** from the SBS data collection

  - In the 2016 MDL project SBS and BD data will be linked to produce data on value-added of **young enterprises**

  - In the 2016 MDL project SBS and BD data will also be linked with ITGS to produce data on value-added of **young exporting and non- exporting enterprises**

MDL in business statistics

- Apart from the standard dimensions e.g. of location and economic activity, the population of enterprises in business statistics can be grouped along the following dimensions:

  - Group status: to determine which enterprises are dependent (part of an enterprises group) and which enterprises are independent

  - Size class: to identify (subgroups of) SMEs based on persons employed

  - Control: to distinguish foreign and domestic controlled enterprises

  - Age: to determine which are the young enterprises

  - Involvement in international trade: to determine importers, exporters, two-way traders and non-traders

EIP data needs and MDL

- Combining groups status with size class to identify **dependent and independent SMEs**

- Combining "control" with size class to compare **foreign and nationally controlled enterprises** in different size classes

- Combining "control" with age to identify **young national and young foreign controlled enterprises** and combine this information on growth to determine **high growth enterprises (HGEs) and gazelles (young HGEs)** which are young enterprises that have had 20 % employment growth or alternatively 10 % employment growth

EIP data needs and MDL

- Using "control" in TEC data to distinguish **foreign controlled and nationally controlled traders**

- Use involvement in trade to divide the **SBS population into exporters, importers , two-way traders and non-traders** and further divide these by size class

- Identify high **growth enterprises and gazelles in above trader populations** by size class

- Ideally we do not implement changes separately in each of our data collections but form a coherent plan on how to use micro data linking and other tools to create rich datasets that cater not only to the data needs identified and presented above but are capable of satisfying future data needs as well

Independent/dependent SMEs

- Data on independent/dependent SMEs are also one of the issues the on-going MDL project linking SBS and other business statistics is working on. First results on this work will be published in September. These results indicate that:

  - Among SMEs, medium sized enterprises are very often part of an enterprise group, most visibly in manufacturing and to a lesser degree in the knowledge intensive business services

  - Dependent SMEs, belonging to a group, are important in terms of employment and gross value added (GVA), especially in smaller countries such as Denmark, Norway and Finland but also in Germany where the dependent SMEs constitute 43 per cent of total GVA created by SMEs and employ 34 percent of total employment in SMEs

  - Dependent SMEs are much more open to international trade than the independent ones; the dependent enterprises behave in this regard very much like the large enterprises

Co-ordinated micro data linking

- Two things are vital for a successful linking:

    1. There has to be a large enough intersection of responding units - up to date SBRs and co-ordination of survey samples are important

    2. There has to be a unique identifier or at least a very reliable matching approach - up to date SBRs are vital

- Confidentiality is an important issue thus the linked micro data files are stored locally at participating NSIs and are not shared with third parties

- Therefore it is important to decide on concrete research objectives before each exercise and to decide on the variables to retain and on the aggregates to be produced

Co-ordinated micro data linking

- To achieve a harmonised output, the participants are provided with precise data requirements and standardised guidelines explaining in detail how the linked datasets in each country are to be structured

- The computer code to produce the tables is provided as well, to ensure that identical tables are constructed in all countries

- Micro data linking is also an efficient tool to improve the quality of existing statistics – linking data from different sources reveal inconsistencies and abnormalities

Co-ordinated micro data linking

- Apart from the SBS MDL exercises, TEC, and the FATS linking projects described above, the ESS has engaged in several other business statistics related MDL exercises
- There has been
  - ESSnet on Linking of Micro data on ICT Usage, where business registers were linked with SBS and ICT usage and e-commerce data
  - There has been an ESSnet on data warehousing (DWH) and MDL which touched more on theoretical aspects
  - The following table provides an overview and shows which countries were involved in each round:

	ICT Usage			Business Statistics				Trade						DWH	FATS
	2006-08	2010-12	2013	2010-11	2012-13	2014-15	2015-16	TEC regular	TEC Meets 2009	TEC Meets 2011	TEC Meets 2013	sTEC voluntary	sTEC 2015-16	Data warehous	iFATS vs oFATS
BE								X							
BG							X	X							
CZ	X			X				X				X			
DK	X	X	X	X	X	X	X	X		X	X	X	X		X
DE	X	X	X			X	X	X	X						
EE							X	X				X	X	X	
IE	X	X	X		X			X				X			
GR								X	X						
ES								X							
FR	X	X	X		X			X							
HR															
IT	X	X	X	X			X	X		X				X	
CY								X							
LT					X			X						X	
LV						X	X	X	X						
LU		X	X				X	X				X			
HU								X	X						
MT								X							
NL	X	X	X	X	X	X	X	X			X	X	X	X	
AT	X	X	X			X	X	X			X	X			
PL		X	X					X			X	X			
PT				X		X		X						X	
RO		X		X	X			X							
SI	X	X	X	X				X							
SK								X		X					
FI	X	X	X	X	X	X	X	X							X
SE	X	X	X		X	X	X	X						X	
UK	X	X	X					X				X		X	
NO	X	X	X	X	X	X	X	X				X	X		X
IS							X	X					X		

"Benefits" of MDL

- Micro-data linking has significant potential in gathering new statistics without increasing the burden placed on respondents

- It is an additional way to ensure data quality and consistency between related data sets

- MDL can be a tool to reduce discrepancies that occur in mirroring international statistics e.g. in inward and outward foreign affiliate statistics

- A coordinated approach (most ESS MDL projects in business statistics applied this approach) to micro-data linking is a cost-effective way for national statistical offices to undertake micro-data linking and ensures harmonised, comparable results across countries

- MDL can help reduce the burden on enterprises when conducting new or existing surveys by eliminating all questions that can be answered from existing data sources through MDL. A successful example here is the International Sourcing survey carried out in 2007 and 2012 (to be carried out in 2017)

"Challenges"of MDL

- It is very important to decide on concrete research objectives before the exercise and to design the linked data set accordingly i.e. decide on the variables for the analysis and on the aggregates to be produced

- Sample surveys versus administrative registers. Administrative data is generally found to be superior to information collected by a survey. Micro-data sets based on administrative data are also more straightforward to gross-up to total population; they are often exhaustive census type of data

- Linking survey based data at enterprise level is additionally complicated by the negative coordination of samples

- From the analytical perspective it is desirable not only to link the latest data but also earlier records going back as many years as possible. Business demography events like the creation, termination or acquisition of enterprises affect the BR populations. Over time matching rates are the highest for latest data (in previous exercises more than 95% in all participating countries) and decrease for each year back in time (often just above 50 % a decade earlier)

"Challenges"of MDL

- It is useful to have a common strategy to deal with demographic events
- Statistical units are not always the same; legal units versus enterprise versus enterprise group complicates the micro-data linking. ESS is aiming at one common statistical unit for business statistics: enterprise. It is also important to implement the statistical unit in the same way
- There are different national institutes responsible of the production official business statistics; this often causes unnecessary problems in the micro-data linking
- Important dimension of globalisation: information about belonging to an enterprise group or not, and whether the enterprise group is all national or foreign is already part of the business register regulation in the ESS, however, this regulation is not implemented or it is implemented in different ways in various member states

Thank you for your attention!

EIP site:

http://ec.europa.eu/eurostat/web/structural-business-statistics/entrepreneurship/indicators

EIP indicators:

http://ec.europa.eu/eurostat/data/database

SE-article:

http://ec.europa.eu/eurostat/statistics-explained/index.php/Entrepreneurship_-_statistical_indicators