TWINNING CONTRACT

Institutional Capacity Building for the Central Agency for Public Mobilisation and Statistics (CAPMAS) and Developing the Legal Framework for Statistics in Egypt

EG/07/AA/F106

MISSION REPORT

on

Activity 5.5.7
Workshop on organising sample frame based on the business register

and

Activity 5.5.8
Business Demography

Mission carried out by
Ms. Tuula Viitaharju, Statistics Finland,
Mr Mark Rantala, Statistics Finland,
and
Mr Søren Netterstrøm, Statistics Denmark

Cairo, 22 – 26 August 2010

Version: Draft
EE 2010:29

		STATISTICS DENMARK
Central Agency for Public Mobilisation and Statistics		Statistics Denmark

Author's name, address, e-mail

Mr. Søren Netterstrøm
Farverstræde 6
DK-4850 Stubbekøbing
Denmark
Tel. +45 5444 0475
sne@viv.dk

Ms. Tuula Viitaharju
Työpajankatu 13
00022 Statistics Finland
Finland
Tel. + 358 9 1734 3378
tuula.viitaharju@stat.fi

Mr. Mark Rantala
Työpajankatu 13
00022 Statistics Finland
Finland
Tel. +358 9 1734 2706
mark.rantala@stat.fi

Table of contents

List of Abbreviations

ABR	Administrative Business Register
BC	Beneficiary country
CAPMAS	Central Agency for Public Mobilisation and Statistics
CoC	Chamber of Commerce
IBR	Integrated Business Register
MoAD	Ministry of Administrative Development
MS	Member States
SBR	Statistical Business Register
SD	Statistics Denmark
SNA	System of National Accounts
ToR	Terms of Reference

Main conclusions and highlights from findings

1. Status

Since the last mission in June there has been a conference organised by Ministry of Administrative Development related to the Unified Register and the Prime Minister Resolution 856 related to sharing of data between administrative units. The intention seems to be to set up a network between the partners allowing this sharing in online mode. IBM is selected for this task. It has also to be clarified what data each provider is willing to share and what data each participants would like to have access to. It was noted, that CAPMAS may not have a high priority in relation to this project. However, for CAPMAS getting online access to data may not be the first priority. What CAPMAS needs are extracts in batch mode, that can be established much simpler, as it was demonstrated when CAPMAS was involved with the encoding of activities in the Commercial register. During the last meeting where MoAD was present it did not seem that they really had understood this major difference but are focused on online access only.

There is no news regarding the accessibility of data regarding turnover and employment, but some investigation is done related to insurance data that may be used for employment figures for large enterprises (?).

The consultants pointed out, that even if the problems of access to turnover and employment data may not be solved quickly, then this should not stop the continuing work of preparing a Statistical Business Register and prepare for its usage, in order to be able to move rapidly forward at the time the data on turnover and employment may become available. The consultants do repeat the recommendation that all business related surveys as soon as possible do include the Tax-ID when collection data in order to utilize these data collections for improvement of the SBR, most important the activity codes.

2. Business demography.

During the workshop the concepts, methods and purpose of Business Demography was presented. See Annex 3 for further details.
The relation between Business Demography and the Statistical Business Register was highlighted. It should be ensured, that the Enterprise Unit in SBR as far as possible is created and maintained with the definitions of birth and death from Business Demography in mind. The ideal situation would be, that SBR contained all the data needed for Business Demography, as SBR will have to determine the kind of demographic event (birth, merge, take over ..). In the SBR administrative starts of the units of the real operational starts should be separated as far as possible. CAPMAS should also find the methods to recognize real births and deaths concerning business demography.

3. Sampling

Sampling has so far not been used in the area of Business Statistics in CAPMAS as a good frame for sampling has not been available. As the business register becomes operational and includes data on employment and/or turnover, such a frame will be available and CAPMAS should take the opportunity to base business surveys on samples.

During the workshop, the basic principles of sampling was introduced and some examples on the actual implementation of sampling were given, see appendix 4.
It was understood that the necessary precondition for sampling is the existence of survey frame (SBR) which covers business population and the most important stratification variables like activity code, size variable and in most cases also some region code. The consultants stressed that to implement sampling in a proper way it is necessary to combine knowledge of the area of statistics to be covered, based on experience and on analysis of historic data, with knowledge of frame (SBR) and knowledge of sampling.
The consultants further stressed the need to know the intended use of the sampled data, what figures and to what level of detail, i.e. geographical distribution or level of ISIC, results will be disseminated in order to determine the stratification and the needed size of the sample.

The consultants pointed out, that sampling opens up possibilities for CAPMAS to widen there coverage of business statistics without the need for further resources. Short Term Statistics (monthly series) could be produced at a low cost and more specialised areas could be covered.

Compared to the present situation, sampling gives the advantages of being able to cover the part of the business community that presently is below the threshold of the surveys carried out.

It is however important to keep in mind, that to produce reliable results from sample surveys, you need a reliable and complete survey frame, i.e. a well established Statistical Business Register.

4. Conclusions and recommendations.

CAPMAS should still pursue the goal to establish a Statistical Business Register based on administrative data and using data collected by CAPMAS surveys to improve the quality of the data for statistical purposes. As there is presently no viable alternative in order increase the quality of the survey frames used by business statistics, CAPMAS should, in spite of the fact that all data is presently not available to CAPMAS, continue the preparations for the system.

CAPMAS should as soon as possible include TAX-ID in all data collections where it is applicable in order to build a base for improved quality of SBR once it can be established.

CAPMAS should continue to examine the options of business demography

CAPMAS should prepare to use samples in the area of business statistics as in order to be able to implement samples as soon as the conditions for doing so, that is a fully functional SBR, is in place. Both sampling specialist and staff responsible for business surveys and SBR may need to have there skills in the area improved.

Annex 1. Terms of Reference

Terms of Reference
for a short-term mission to the Central Agency for Mobilisation and Statistics
On

Activity 5.5.7
Workshop on organising sample frame based on the business register

Joined with

Activity 5.5.8
Business demography

Background

CAPMAS and Statistics Denmark with partners have established a fruitful cooperation in the framework of Twinning. This twinning project is EG/07/AA/F106.

This activity is part of component 5, *Development of certain statistical areas* in the subcomponent dealing with Statistical Business Register. The objective for this component is development of statistical business register.

This activity will contribute to this objective and especially to the benchmark set out in the contract: *By the end of the sixth quarter, MS and BC experts have organised sample frames based on the business register and by the end of the seventh quarter, MS and BC experts have developed the methodology on business demography.*

The first benchmark cannot be achieved since a business register will not be established within the time frame of the project. However some work can be done in establishing the role of the business register in producing sample frames.

Purpose of the mission

This mission is a workshop where the MS and BC experts will work on the method to use the business register in sampling for business surveys and use of the business register before and after the survey.

Also the MS will share experience in using business register data to analyse business demography and work on the content and methods in business demography.

Expected Results

- A methodology on business demography
- Description of the use of a business register when organising sample frames

Activities

A tentative schedule for the mission is:

Sunday 22 August 2010
Discussion on the activity with the RTA
Meeting with component leader and the relevant staff within CAPMAS
Determining the agenda for the mission

Status of the work with the business register, among other things, the internal organisation and establishment of a business register unit in CAPMAS and the cooperation with the unified business register and the customs authorities.

Monday 23 August 2010
Business demography: the experiences in Finland
Possibilities in Egypt with the expected set up of the future statistical business register and existing statistics in Egypt

Tuesday 24 August 2010
The uses of the Danish and Finnish statistical business register in surveys
Using a statistical business register in setting up surveys and to produce the final statistical results adjusting for missing data, non-response and other differences between the sample frame and the actual population in Egypt

Wednesday 25 August 2010
Using a statistical business register in setting up surveys and to produce the final statistical results adjusting for missing data, non-response and other differences between the sample frame and the actual population (cont.)

Thursday 26 August 2010
Final discussions and presentation of the conclusions and result

Tasks to be done by CAPMAS to facilitate the mission

The beneficiary will arrange meetings with the relevant staff in CAPMAS.

Consultant and counterpart

The mission will be carried out jointly by:

Ms. Tuula Viitaharju, Statistics Finland
And
Mr. Mark Rantala, Statistics Finland
And
Mr. Soren Netterstrom, Statistics Denmark

The beneficiary's counterpart will be:
Ms. Awatef Hussein Emam Abou Gendy
And
Ms. Iman El Hitta

Timing

The mission will be carried out during 22- 26 August 2010 in Cairo.

Report

A final report from the mission should be made available not later than two weeks after the termination of the mission.

Annex 2. Participants / persons met

Consultants

Name	Title
1- Mr. Søren Netterstrøm	Consultant
2- Mr. Mark Rantala	Statistics Finland
3- Ms. Tuula Viitaharju	Statistics Finland
4- Ms. Iman El Hitta	CAPMAS

Annex 3. Business Demography

Statistics Finland

Introduction to Business Demography

Mark Rantala
Cairo, Egypt 22.8.2010

Statistics Finland

Contents

- Business Demography in nutshell
  - What is BD
  - Why BD series
  - A brief history of BD
  - Harmonised data collection and employer & high-growth series
  - BD methodology
  - Twinning BD in Egypt

Statistics Finland

What is BD

- Definition in BR Glossary
  "Business demography covers events, like births and other creations of units, deaths and other closings of units, and their ration to the business population. It covers follow-up of units in time dimension, thus gaining information on their survival or discontinuity. It also covers development in time dimension according to certain characteristics like size, thus gaining information on the growth of units, or a cohort of units, by type of activity."
- What does it say?
  - Events ->true births and deaths separated from other (administrative or other) creations
  - Ratio to populations -> relative importance
  - Follow up in time -> survival, success 1
  - Development in time -> success 2

Statistics Finland

What is BD

- Variables
  - Hard figures of the businesses based on (real) births and deaths of enterprises, their survival from one year to another years and growth in different phases of the life span
- Indicators
  - Descriptions of the businesses based on actual and relative number of business start-ups and closures, their impact on employment and turnover in different sectors
- Focusing
  - Enterprise dynamics between sectors
  - Business churn impact on employment, on success
  - Enterprise specific characteristics and their relation to business success/ failure
  - Combining BD attributes (birth, death, survival, growth) of the individual enterprise to other unit level data produce valuable source if information for researchers, etc.

Statistics Finland

Why BD data is needed?

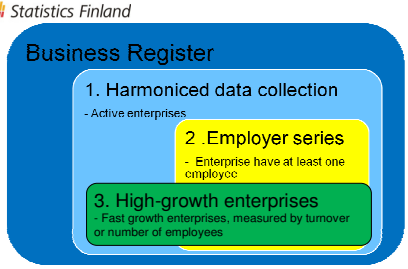
- Coherent and comparable data across the members of the European Statistical System (ESS)
- Policy objects
  - promoting entrepreneurship is instrument for improving competitiveness thus generating economic growth and job opportunities
  - formation of new enterprises and the growth of small and medium-sized enterprises are essential for job creation
  - goal: EU most dynamic knowledge based economy in the world, sustainable growth, more and better jobs in the EU
  - economic growth and productivity, information for tackling social demographic issues
- => Support for entrepreneurship and entrepreneurial dynamism
- => Revealed by BD statistics and studies utilising BD base data

Statistics Finland

A brief history of BD

- Voluntary collection of existing data Autumn 1999
- The BSDG meeting of 27 of October 1999 and the SBS Steering Group of 17-18th of February approved the project plan emphasising the need for methodological work.
- Final proposal for harmonisation January 2001
- Signature of grant contracts for feasibility study April 2001
- ...
- 2008 7th data collection round
- 2007 first data collection on new pilot series, employer and high-growth enterprises
- 2007 publication of the joint manual of Eurostat and OECD about business demography
- Common BD series on SBS regulation from 2008. Effect on data collection from 2009 onwards

Statistics Finland



Statistics Finland

Harmonised data collection and employer and high-growth series

- 1, Harmonised data collection =HDC = common BD series
  - Figures about births, death, survival. By activity, by legal form, by employee size class
  - Active enterprises defined by having turnover or persons employed at any time during the reference year
  - No thresholds
  - Removal of the non-real events due to administrative registrations, practices in national business registers, etc. Real birth = from scratch, no predecessors. Real death = dissolution of production factors, no successors
- 2. Employer series
  - Only enterprises having paid employees
  - International comparability, removal of extremely small enterprises (~noise)
- 3. High-growth series
  - Fast growing enterprises by turnover, by number of persons employed (3y)
  - Gazelles: high growth enterprise less than 5 years old

Statistics Finland

BD methodology in brief

- Active unit
  - any turnover or any employment =>active
- Birth year: First active year or first active year after two years of inactivity
- Death year: last active year followed by two years of inactivity
- Survival: continuous activity, one inactive year remove from pool of survived enterprise (but does not produce birth nor death)
- Removal of "unreal" births and deaths due to restructuring of existing enterprises, transfer of production factors, discontinuity of the business ID but continuity of the business unit, (mergers, split-offs, break-ups...)
- Removal tool: matching by name, activity, location. Two out of three produce hit for unreal event. Manual verification of the large units.
- National methods accepted

Statistics Finland

Eurostat documentation

- Eurostat - OECD joint Manual on Business Demography Statistics
  - Cook book with basic recipes
- CIRCA web-site: http://circa.europa.eu/Members/irc/dsis/structbus/library
  - Business demography meetings
  - Business demography documents
- Annual delivery of the instructive dokumentation
  - Summary of deliverables, information about the requested data. Delivery to Eurostat with data
  - Transmission format, record and dataset structure, field descriptions, codes, activity and size classes, variables
  - Template for methodological report, delivery to Eurostat with data.

Statistics Finland

Twinning BD in Egypt

- Discussion about the Egyptian BD statistics
  - Goals
  - Data sources
  - Schedule

Statistics Finland

Methodology, defining enterprise births and deaths in BD statistics

Mark Rantala
Cairo, Egypt 26.4.2010

Statistics Finland

Contents

- Births
  - Review of the manual
  - Distinction between entries and births
  - Harmonised BD series
  - Employer enterprise births
  - Identification of the enterprise births
  - Variables in births
- Deaths
  - Review of the manual
  - Distinction between closures and deaths
  - Harmonised and employer deaths
  - Timeliness of the data
  - Variables

Statistics Finland

Review of the manual/ Births

- Chapter 3, defining population of active units
- Chapter 4, defining different events in real world and register
- Chapter 5, defining enterprise births

- Chapter 5 content
  - Introduction to births
    - Timing
    - Employee thresholds in different 'birth' definitions
  - Form enterprise creations to enterprise births
    - Exclusions of register entries
    - Ancillary units
    - Identification of the births, process
  - Employer enterprise births, economic enterprise births
  - Employment in newly born enterprises
  - Indicators

Statistics Finland

Distinction between entries and births

- Statistical world vs. administrative world
  - Unique ID-code? Double-triple… entries
  - Using statistical enterprise or using legal units
  - Convention and legislation in occurrence of changes in businesses and their structures
  - Level of accuracy in BR demography of statistical units
- Typology of events described in detail on manual chapter 4
  - in theory simple formula
  - in practice extremely difficult to carry out in detailed level
  - diversity of the 'events' in real world, non-pure solutions
- Matching used for finding links between predecessors and successors ...
- …completed with national data sources and methodologies
- Timestamp of entry, timestamp for birth
  - Exact date would be nice, level of accuracy one year needed

Statistics Finland

Harmonised BD series, births

- In HDE series birth occurs when
  - 1st year of activity or unit inactive 2 years before reactivation
  - Activity defined by existence of turnover or employment
  - HDE series any employment (paid employees, entrepreneur, entrepreneurs wife brother children etc.) is sufficient
- Birth definition
  - "A count of the number of births of enterprises registered to the population concerned in the business register corrected for errors. A birth amounts to the **creation of a combination of production factors** with the restriction that **no other enterprises are involved** in the event. Births do **not include** entries into the population due to: **mergers, break-ups, split-off or restructuring of a set of enterprises**. It does **not include entries into a sub-population** resulting only from a change of activity."
- What does it say
  - new combination production factors = no transfer of production factors
  - from the scratch = without any proceeding businesses
  - no restructuring of businesses = business reorganisations excluded
  - no entries to total sub-populations = same, same but different is not enough

Statistics Finland

Employer enterprise births

- As before but employment threshold 1 paid employee
- Focus on employment effects
- International comparisons with non-EU OECD countries
- In EMP series birth occurs when
  - 1st year of activity with paid employees or units which existed before but without any paid employees and which hired the first employee
  - Move from the population of non-employer to employer is not due to growth by take-over
- Still, the labour input of entrepreneur is reflected in 'paid employees' if the compensation of the self employed is taken as common salary instead of entrepreneur's profit
- => Economic births with threshold of 2 paid employees (non-implemented series included in BD manual)

Statistics Finland

Identification of the enterprise births, process

- Populations
  - Defining population of active enterprises in year xx (population Nxx)
  - Defining population of active enterprises in xx-1, xx-2 (Nxx-1, Nxx-2)
  - Suggested birth if active only in Nxx and not in Nxx-1, Nxx-2
- Elimination of reactivations by matching of units
  - Name-location, name-activity, activity-location
  - Most detailed level of location and activity
  - Activity-location produce lot of false matches
  - Auxiliary information from national sources and methodology
- Correction of errors
  - Manual verifications, national methods
  - Large enterprises all, samples for smaller enterprises

Statistics Finland

Variables in births

List E – Variables

Label	Code	Notes
Population of active enterprises in t	11910	Population Nxx
Number of births of enterprises in t	11920	Population Rxx
Number of deaths of enterprises in t	11930	Population Dxx
Number of enterprises newly born in t-1 having survived to t	11941	Subset of population R(xx-1)
Number of enterprises newly born in t-2 having survived to t	11942	Subset of population R(xx-2)
Number of enterprises newly born in t-3 having survived to t	11943	Subset of population R(xx-3)
Number of enterprises newly born in t-4 having survived to t	11944	Subset of population R(xx-4)
Number of enterprises newly born in t-5 having survived to t	11945	Subset of population R(xx-5)
Number of persons employed in the population of active enterprises in t	16910	Head count – annual average Population Nxx
Number of persons employed in the population of births in t	16920	Head count – annual average Population Rxx
Number of employees in the population of births in t	16921	Head count – annual average Population Rxx
Number of persons employed in the population of deaths in t	16930	Head count – annual average Population Dxx
Number of persons employed in the population of enterprises newly born in t-1 having survived to t	16941	Head count – annual average Subset of population R(xx-1)
Number of persons employed in the population of enterprises newly born in t-2 having survived to t	16942	Head count – annual average Subset of population R(xx-2)
Number of persons employed in the population of enterprises newly born in t-3 having survived to t	16943	Head count – annual average Subset of population R(xx-3)
Number of persons employed in the population of enterprises newly born in t-4 having survived to t	16944	Head count – annual average Subset of population R(xx-4)
Number of persons employed in the population of enterprises newly born in t-5 having survived to t	16945	Head count – annual average Subset of population R(xx-5)
Number of persons employed in the year of birth in the population of enterprises newly born in t-1 having survived to t	16951	Head count – annual average Subset of population R(xx-1)
Number of persons employed in the year of birth in the population of enterprises newly born in t-2 having survived to t	16952	Head count – annual average Subset of population R(xx-2)
Number of persons employed in the year of birth in the population of enterprises newly born in t-3 having survived to t	16953	Head count – annual average Subset of population R(xx-3)
Number of persons employed in the year of birth in the population of enterprises newly born in t-4 having survived to t	16954	Head count – annual average Subset of population R(xx-4)
Number of persons employed in the year of birth in the population of enterprises newly born in t-5 having survived to t	16955	Head count – annual average Subset of population R(xx-5)

Statistics Finland

Distinction between exits and deaths

- Administrative de-registration vs. cessation of the activities
  - Administrative recording too slow or not present at all
  - Flagging inactive units is problem everywhere.
  - No information about deterministic cessation because in many cases such does not exist.
  - Problems especially with minor enterprises, sole proprietors and others
  - 1 year of inactivity= dormant, 2 year of inactivity=cessation
- Typology of events described in detail on manual chapter 4
- Re-activation slow accumulation of final data on business deaths
- Matching used for finding links between predecessors and successors ...
- …completed with national data sources and methodologies
- Timestamp for death, timestamp for administrative de-registration, timestamps for periods of inactivity
  - Exact date would be nice, level of accuracy one year needed

Statistics Finland

Harmonised BD series, deaths

- In HDE series death occurs when
  - Last year of activity followed by 2 years of inactivity
  - Preliminary data based on 1 year of inactivity
  - Estimation methods, ratio of preliminary deaths and 'final' deaths, (etc) for the sake of timeliness
  - If reactivation within 2 years, preliminary death will be cancelled
- Death definition
  - "A count of the number of deaths of enterprises registered to the population concerned in the business register corrected for errors. A death amounts to the ***dissolution of a combination of production factors*** *with the* ***restriction*** *that* ***no other enterprises*** *are involved in the event. Deaths do* ***not include*** *exits from the population due to* ***mergers, take-overs, break-ups and restructuring of a set of enterprises****. It does* ***not include exits from a sub-population*** *resulting only from a change of activity.*"
- What does it say
  - dissolution of production factors = not complete transfers of production factors
  - no other enterprises involved = activities will be closed down for good
  - no restructuring of businesses = no business reorganisation
  - no exits from sub-populations = activity changes even to agriculture is not an excuse

Statistics Finland

Employer enterprise deaths

- As before but employment threshold 1 paid employee
- In EMP series death occurs when
  - Last year of activity with paid employees or units which had employees in previous year but moved in population of non-employers for two consecutive years
  - Move to the population of non-employer is not due to split-off
- Labour input of the entrepreneur still hindering the 'true' employer death data
- => Economic deaths with threshold of 2 paid employees (non-implemented series included in BD manual)

Statistics Finland

Identification of the enterprise deaths, process

- Populations
  - Defining population of active enterprises in year xx (population Nxx)
  - Defining population of active enterprises in xx+1, xx+2 (Nxx+1, Nxx+2)
  - Suggested death if active only in Nxx and not in Nxx+1, Nxx+2
- Elimination of reactivations by matching of units
  - Name-location, name-activity, activity-location
  - Most detailed level of location and activity
  - Activity-location produce lot of false matches
  - Auxiliary information from national sources and methodology
- Correction of errors
  - Manual verifications, national methods
  - Large enterprises all, samples for smaller enterprises

Statistics Finland

Variables in deaths

List E – Variables

Label	Code	Notes
Population of active enterprises in t	11910	Population Nxx
Number of births of enterprises in t	11920	Population Rxx
Number of deaths of enterprises in t	11930	Population Dxx
Number of enterprises newly born in t-1 having survived to t	11941	Subset of population R(xx-1)
Number of enterprises newly born in t-2 having survived to t	11942	Subset of population R(xx-2)
Number of enterprises newly born in t-3 having survived to t	11943	Subset of population R(xx-3)
Number of enterprises newly born in t-4 having survived to t	11944	Subset of population R(xx-4)
Number of enterprises newly born in t-5 having survived to t	11945	Subset of population R(xx-5)
Number of persons employed in the population of active enterprises in t	16910	Head count – annual average Population Nxx
Number of persons employed in the population of births in t	16920	Head count – annual average Population Rxx
Number of employees in the population of births in t	16921	Head count – annual average Population Rxx
Number of persons employed in the population of deaths in t	16930	Head count – annual average Population Dxx
Number of persons employed in the population of enterprises newly born in t-1 having survived to t	16941	Head count – annual average Subset of population R(xx-1)
Number of persons employed in the population of enterprises newly born in t-2 having survived to t	16942	Head count – annual average Subset of population R(xx-2)
Number of persons employed in the population of enterprises newly born in t-3 having survived to t	16943	Head count – annual average Subset of population R(xx-3)
Number of persons employed in the population of enterprises newly born in t-4 having survived to t	16944	Head count – annual average Subset of population R(xx-4)
Number of persons employed in the population of enterprises newly born in t-5 having survived to t	16945	Head count – annual average Subset of population R(xx-5)
Number of persons employed in the year of birth in the population of enterprises newly born in t-1 having survived to t	16951	Head count – annual average Subset of population R(xx-1)
Number of persons employed in the year of birth in the population of enterprises newly born in t-2 having survived to t	16952	Head count – annual average Subset of population R(xx-2)
Number of persons employed in the year of birth in the population of enterprises newly born in t-3 having survived to t	16953	Head count – annual average Subset of population R(xx-3)
Number of persons employed in the year of birth in the population of enterprises newly born in t-4 having survived to t	16954	Head count – annual average Subset of population R(xx-4)
Number of persons employed in the year of birth in the population of enterprises newly born in t-5 having survived to t	16955	Head count – annual average Subset of population R(xx-5)

Statistics Finland

Survival

- Some remarks
  - Survival only for consecutive years ->reactivated units are not counted as survivals
  - Continuation of the production factors is the main criteria (once again)
  - Survival by take-over only if continuing enterprise is new: Exclusion of events where existing enterprise take over the activities of the closed unit.
  - Survival by take-over does not require the cessation of the original unit.
- About variables
  - 3 sets of variable, 5 years follow-up
  - 1194X: Number of enterprises newly born in t-X (X=1..5) having survived to t, ref.year R(xx-X)
    1694X: Number of persons employed in the population of enterprises newly born in t-X having survived to t, ref.year R(xx-X)
    1695X: Number of persons employed in the year of birth in the population of enterprises newly born in t-X having survived to t, ref.year t-X
- Reference year
  - Stratification variables (activity, legal form, employee size class) taken from the reference year
  - If applied properly, different variables should be in line with each other

Statistics Finland

3. High-Growth Enterprises

Mark Rantala
Cairo, Egypt 23.8.2010

Statistics Finland

Variables: survival, active units

LEE – Variables

Label	Code	Notes
Population of active enterprises in t	11910	Population Nxx
Number of births of enterprises in t	11920	Population Rxx
Number of deaths of enterprises in t	11930	Population Dxx
Number of enterprises newly born in t-1 having survived to t	11941	Subset of population R(xx-1)
Number of enterprises newly born in t-2 having survived to t	11942	Subset of population R(xx-2)
Number of enterprises newly born in t-3 having survived to t	11943	Subset of population R(xx-3)
Number of enterprises newly born in t-4 having survived to t	11944	Subset of population R(xx-4)
Number of enterprises newly born in t-5 having survived to t	11945	Subset of population R(xx-5)
Number of persons employed in the population of active enterprises in t	16910	Head count – annual average Population Nxx
Number of persons employed in the population of births in t	16920	Head count – annual average Population Rxx
Number of employees in the population of births in t	16927	Head count – annual average Population Rxx
Number of persons employed in the population of deaths in t	16930	Head count – annual average Population Dxx
Number of persons employed in the population of enterprises newly born in t-1 having survived to t	16941	Head count – annual average Subset of population R(xx-1)
Number of persons employed in the population of enterprises newly born in t-2 having survived to t	16942	Head count – annual average Subset of population R(xx-2)
Number of persons employed in the population of enterprises newly born in t-3 having survived to t	16943	Head count – annual average Subset of population R(xx-3)
Number of persons employed in the population of enterprises newly born in t-4 having survived to t	16944	Head count – annual average Subset of population R(xx-4)
Number of persons employed in the population of enterprises newly born in t-5 having survived to t	16945	Head count – annual average Subset of population R(xx-5)
Number of persons employed in the year of birth in the population of enterprises newly born in t-1 having survived to t	16951	Head count – annual average Subset of population R(xx-1)
Number of persons employed in the year of birth in the population of enterprises newly born in t-2 having survived to t	16952	Head count – annual average Subset of population R(xx-2)
Number of persons employed in the year of birth in the population of enterprises newly born in t-3 having survived to t	16953	Head count – annual average Subset of population R(xx-3)
Number of persons employed in the year of birth in the population of enterprises newly born in t-4 having survived to t	16954	Head count – annual average Subset of population R(xx-4)
Number of persons employed in the year of birth in the population of enterprises newly born in t-5 having survived to t	16955	Head count – annual average Subset of population R(xx-5)

Statistics Finland

Contents

- Need of high-growth enterprises data
- Definitions of the high-growth enterprises and gazelles
- Methodology
- Fine details and pitfalls
- Variables
- Twinning BD in Egypt

Statistics Finland

Matching

- Basic approach, matching by
  - Name and location SAFE
  - Name and activity SAFE
  - Location and activity DUBIOUS
- Practices and rules
  - Location and activity in most detailed level (4 digit nace/ national versions, street address)
  - Location, same street addresses in different cities? Street address + postal code/ municipal code/ etc.
  - Multiple matches? Shopping centres, malls, "liberate professions", health centres,...?
  - How to define 'close matches' in automatic procedures. Spelling errors, abbreviations, XXXX Nova...
- Time dimension
  - Matching must be carried against the pool of active enterprises (with some exceptions) in 3 years
  - Birth, years t, t-1, t-2
  - Death years t, t+1, t+2
  - Can be done within the same data steps

Statistics Finland

Need of high-growth enterprises data

- Policy drivers
  - High-growth enterprises (HGE) are seen as one of the main drivers for increasing (economic) welfare
  - View that HGE create more benefits to whole society than 'common' enterprises
  - More and better jobs
- Basic information about the number of HGE
- How to promote the creation of the HGE -> tools for 'academic' studies, combining of different data to HGE

Statistics Finland

Definitions of the high-growth enterprises

- All enterprises with average annualised growth greater than 20% (or 10-20%) per annum, over a three year period should be considered as high-growth enterprises. Growth can be measured by the number of employees or by turnover.
- Small (new) enterprises starting to grow may distort the picture -> need of threshold
- Annual growth rate and threshold have impact of the data confidentiality
  - Small number of HGE ->confidentiality problems
  - Too much confidential data -> loss of information for data users -> no value of the data
- Different variables for HGE by employment and by turnover
- EU definition
  - Only 5+ employee enterprises in the beginning of the growth period
  - Annual growth rate 20+% (or 10-20%)
  - No turnover threshold but employment threshold applied also to turnover series

Statistics Finland

Methodology

- Follow up on three year period
- When trying to identify high-growth enterprises, it is not necessary to check the change in employee numbers or turnover from one year to the next over a three-year period….
- but to calculate ratio between years xx-3 and xx
- In practice, average annualised growth of 20% over three years would be equal to 72.8% growth from xx-3 to year xx
  - check by ID number comparison which enterprises in population $N_x$xx-3 are still in population $N_x$xx, and then
  - check whether the number of employees, or turnover respectively, in year xx is at least 1.728 times higher than in year xx-3.
- If the number of employees drop below 5+ threshold, enterprise is still kept in population

Statistics Finland

Fine details and pitfalls

- Beginning of the growth period have to start after the year of birth
  - If starting operations in the middle of the year -> risk of recording turnover only for part time year basis
  - In HGE defined by employment -> same populations in turnover and employment HGE series
- If the number of employees drop below 5+ threshold, enterprise is still kept in population
- Exclusion of growth due to mergers, take-overs, etc.
  - Focusing on true high-growth instead of growth due to combining of the units
  - Information concerning merged of units involved in take-over created already in production of harmonised data series -> matching

Statistics Finland

Gazelles

- Subset of high-growth enterprises
  - Definition 'up to five years old' in the end of the growth period
    - Some request for changing the definition as 5 yrs old in the beginning of the growth period
  - Otherwise same definitions as with HGE
  - Year of birth not taken in consideration (same as with HGE)

Statistics Finland

Variables

- 11950 Number of high growth enterprises measured in employment
- 11951 Number of high growth enterprises measured in turnover
- 11952 Number of gazelles measured in employment
- 11953 Number of gazelles measured in turnover

Indicators in HGE and GAZ

- Rate of high-growth enterprises: Number of high-growth enterprises as a percentage of the total population of active enterprises with at least t employees.
- Rate of gazelles among newly born enterprises: Number of gazelles as a percentage of all active enterprises with at least t employees that were born four or five years ago.

Statistics Finland

Summary of Deliverables HDC

- Excel table to be filled and delivered with data to Eurostat
- Spreadsheets with
  - Deliverables from the series 9A - 2010 harmonised data collection with breakdown by NACE rev2. activity and legal form
  - Deliverables from the series 9B - 2010 harmonised data collection with breakdown by NACE rev2. activity and employee size class
  - Deliverables from the series 9C - 2010 harmonised data collection with breakdown by NACE rev2. activity and legal form: Estimated data on enterprise deaths in 2008
  - Deliverables from the series 9D - 2010 harmonised data collection with breakdown by NACE rev2. activity and size class: Estimated data on enterprise deaths in 2008
- Variables, labels, reference populations, reference years, breakdowns
- Also available for HGE and GAZ series

Annex 4. Sampling

Statistics Finland

Business Register in statistics production

Business Register as sampling frame

Mission Egypt, August 2010

Tuula Viitaharju

Statistics Finland

Target of this mission / presentation

- to discuss the benefits and need of unified SBR (=statistical Business Register) in statistics production =>
  - better quality of statistical results
  - rational workflows and co-operation in statistical office
  - data sharing (feedback data)
  - Business Register makes sampling possible

Statistics Finland

Our strong recommendation as a result and first step of this Mission component

- to establish a special unit or team for SBR tasks
- to make a data base and data model for SBR
- to start to update some data in SBR
  - survey by survey
  - administrative data source by source (tax-files, UR-files, Insurance-files etc…)
- this means: to be on a way to permanenty maintained SBR
- ***To complete the system will take time! Best to start as soon as possible!!!***

Statistics Finland

Sampling

- Sampling means that you survey only some part of units in population and generalise the result on the whole population level
- BR helps to apply effective samples
- possible to concentrate on large / important units
- you can also survey the units above some threshold

Statistics Finland

Sampling procedure in general

- you have a target: what kind of statistics you want to produce
  - identification of the statistical units
  - identification of the population
  - => FRAME
- Business register is a frame for Enterprise or Establishment statistics

Statistics Finland

Sample design

- choose of sampling method, exsamples
  - a) simple random sampling (SRS)
  - b) stratified random sampling (SSRS)
  - c) propability proportional to size (PPS)

Statistics Finland

Stratification

- size
- activity
- region

- in order to form homogeneous groups, so the accuracy of results will be good

Statistics Finland

Stratification (2)

- size
  - either take all (you choose all units above certain threshold, example: employment > 250)
  - or take some (you form groups (= stratum, strata), like employment between 0-4, 5-10, 11-50, 50-100, 100-250)
- this means => there should be the size variable in the frame (SBR)
  - it should be as real-time as possible
  - correct and
  - good coverage

Statistics Finland

Stratification (3)

- activity
  - ISIC 4
  - reasonable level of breakdown, like 2-digit level
    - to form homogeneous strata
    - important classes can be chosen (those, which you want to study)
- region
  - when you you need to produce regional statistics
- these means => there should be the activity code and some region code in the frame (SBR)

Statistics Finland

Allocation

- take all
  - you take all units of the stratum
- take some
  - equal (same number of units from each stratum)
  - proportional allocation (different number of units from each stratum)
  - optimal allocation (like Bankier, Neuman allocations)
    - idea: some auxiliary variable which is correlated with study variables (the greater the variance of the auxiliary variable in the stratum is, the greater number of sample is needed; it means that the more heterogeneous is the stratum the the greater sample is needed)

Statistics Finland

Data collection

- questionnaires sent to chosen units (enterprises, establishments…)
- After response
  - efforts to minimize non-response
  - chechings, imputations…

Statistics Finland

Weighting
Estimation after data collection

- generally by HT estimation
  - inverse of sampling weight (share of sample in the frame) in each stratum
- in practice
  - inverse of share of responded units to frame in each stratum (takes care of unit non-response)

Statistics Finland

Summary: SBR as sampling frame

- frame
  - all units, incl. active, non-active, removals
  - stratification variables
    - classifications
    - size indicator
- mailing data: fresh addresses
- handling survey data / survey results
  - non-response
  - sample vs total population: estimation

Statistics Finland

Size distribution of enterprises according to employment, in Finland, in 2009

Size class	Number of enterprises	Share %	Employment	share %	Turnover milj. euro	Share %
Small and medium 0-249	320 296	99,8	942 831	62,8	193 148	48,7
- small 0-49	317 855	99,0	696 944	46,4	129 195	32,6
- Medium 50-249	2 441	0,8	245 887	16,4	63 953	16,1
Large enterprises > 250	656	0,2	559 382	37,2	203 444	51,3
Total	320 952	100,0	1 502 213	100,0	396 591	100,0

Statistics Finland

Thank you for your attention!

Tuula Viitaharju
Head, Business Register
Statistics Finland
tuula.viitaharju@stat.fi
tel: +358 9 1734 3378

Statistics Finland

Sharing Business Register data - in-house

- BR database is in wide use inside Statistics Finland
- about 300 users from economic as well as from social statistics
- **on-line access** or reading **by batch** programs
- use is supported by BR metadata kept available in Statistics Finland (quality reports, time table of BR updatings, variable descriptions, process descriptions)

Statistics Finland

Business Register in statistics production

Sampling example in Statistics Finland
Mission Egypt, August 2010
Tuula Viitaharju

Statistics Finland

Business surveys

Altogether about 60 direct surveys

Annual inquiry on foreign trade in services
Authority (Virati) data collections (A and T data collections)
Building cost index
Business Register inquiry for multi-establishment enterprises
Business Register inquiry for new enterprises
Business Register inquiry for single-establishment enterprises
Business Register inquiry for small single-establishment enterprises
Business services
CVTS, Continuing Vocational Training Survey
Credit cards
Data collection on occupations
Data collection on personnel of multi-establishment enterprises for employment statistics
Energy production inquiry
Energy use in manufacturing
Enterprises' research and development
Environmental protection expenditure in industry
Finance of housing corporations
Financial leasing inquiry
Financial statements inquiry for enterprises (TILKES)
Financial statements of mutual funds and investment management companies
Forest workers' wages

More information on surveys:
http://stat.fi/keruu/index_en.html

Income and expenditure of foreign sea transport
Information technology and electronic commerce in enterprises
Innovation survey
Inquiry for industrial establishments (T5)
Inquiry on cost indices
Inquiry on incentive stock options
Inquiry on manufacturing commodities
Inquiry on new orders in manufacturing
Inquiry on postal and small freight deliveries
Inquiry on private sector wages and salaries
Inquiry on volume index of industrial output
Inventory inquiry
Job vacancy survey
Monthly inquiry for accommodation statistics
Monthly inquiry for iron and steel statistics
Mutual fund inquiry
Outstanding credit
Periodic survey of motor vehicle trade
Periodical regional survey of trade
Periodical survey of retail trade
Periodical survey of wholesale trade
Price inquiry on producer price indices for manufactured products
Producer price indices for services
Quarterly inquiry on international trade in services
Quarterly inquiry on labour costs
Regional outstanding credit
Renovation building of construction enterprises
Renovation building of housing corporations
Sales inquiry
Statistics on agriculture
Statistics on goods transport
Telecommunications inquiry
The inquiry on financial accounts

Statistics Finland

Use of Business Register in Sampling

- BR database is updated continually, register of each statistical year is finalised in t+10 months
- branch statistics draw their samples by themselves direct from Business Register (their own programs)
- they have on-line access or reading by batch programs
- use is supported by BR metadata (quality reports, time table of BR updatings, variable descriptions, process descriptions)

Statistics Finland

Sampling, one general example of stratification

- activity: 2 digit
- size:
  - large enterprises, employment > 250
    - take all
  - medium enterprises, 50-249
    - sample by PPS
  - small and micro enterprises,
    - 0-49, sample by PPS or not at all (not studied or only administrative data used)

Statistics Finland

Sampling example in Finland

- *Volume Index of Industrial Output, monthly*
- *size*
  - large establishments, employment > 150
    - take all
  - medium establishments, 50-149
    - sample by PPS
  - small enterprises,
    - 0-49, only administrative data used
- *activity*: NACE 3-digit level groups in Manufacturing Industry

Statistics Finland

Business Register; data content (1)

- Identification
  - ID-codes
  - Names
  - Addresses
  - Telephone, fax, e-mail
- Classification
  - activity (NACE)
  - legal form
  - ownership code
  - classification of sectors 2000
  - regional classifications
  - map co-ordinates

Statistics Finland

Business Register; data content (2)

- Size variables
  - employment
  - turnover
  - sum of wages and salaries
- Status variables
  - starting / cessation dates
  - starting / cessation forms
  - date of take-over, merger
  - links between abandoning and receiving units
- Others, like
  - variables for the management of BR own surveys
  - codes for employer, subject to VAT, subject to tax prepayment register
  - exporter/importer, intrastat

Statistics Finland

Thank you for your attention!

Tuula Viitaharju
Head, Business Register
Statistics Finland
tuula.viitaharju@stat.fi
tel: +358 9 1734 3378

Accounts statistics DK

- Statistical business register-draw with information of all active enterprises and its local units:
  - identity codes
  - names
  - addresses
  - economic activity codes (NACE)
  - ownership codes
  - no. of employees (in full-time units)
  - VAT turnover

Sample rotation DK

Main principles:

- *5-9 employees. 10% selected*
  Each enterprise is selected for 1 year followed by 9 years of exemption
- *10-19 employees. 20% selected*
  Each enterprise is selected for 2 years followed by 8 years of exemption
- *20-49 employees. 50% selected*
  Each enterprise is selected for 3 years followed by 3 years of exemption
- *>=50 employees. 100% selected*
  All enterprises selected

Sample rotation DK

20-49 employees. 50% selected

Selected (each letter represents 1/6 of total population)

Survey year			
2004	A, B, C		
2005		B, C, D	
2006			C, D, E
2007		D, E, F	
2008	A,		E, F
2009	A, B		F

Stratification DK

The sample population is split up in groups ("strata") using a combination of info from the Business Register:

1.NACE
2.Ownership: company versus
single proprietor/self employed, partnership
3. Size group: number of employees
0-9
10-19
20-49
50+

(main principles: at least 5 enterprises in each stratum)

Grossing-up procedure

1. Final results of questionnaire-based survey
2. Compute sub-strata averages
   (using Business Register's: activity code, size and form of ownership)
3. Interpolate data for units covered only by business tax return summaries
   (using sub-strata averages in missing cells)
4. Compute new sub-strata averages (for 1+3)
5. Estimate results for non-covered units
   (using new-sub-strata averages, account item per employed person or per turnover unit)

1+3+5 = full coverage ☺ (205,000)

Total population in Construction. Number of enterprises by industry, type of ownership and size of employment

		Single proprietor, selfemployed, partnership (P)					Company (S)				
		Number of employees (in full-time units)									
DB07	Stdgrp 127	< 5	'5 - 9	'10 - 19	20 - 49	50 +	< 5	'5 - 9	'10 - 19	20 - 49	50 +
Size code		0	1	2	3	4	0	1	2	3	4
411000	41000	119	0	0	0	0	439	8	6	7	1
412000	41000	3.168	38	19	3	0	3.553	304	224	163	104
421000	42000	162	1	0	1	0	80	22	10	16	9
422000	42000	506	24	11	3	1	254	58	54	51	34
429000	42000	207	2	0	0	0	107	8	7	4	4
431100	43009	93	0	1	0	0	17	0	0	6	1
431200	43009	981	24	4	3	1	327	59	53	29	10
431300	43009	8	0	0	0	0	2	1	2	0	0
432100	43001	2.169	105	30	4	2	914	342	306	187	62
432200	43001	2.125	107	28	6	0	1.137	383	299	138	29
432900	43001	196	5	1	0	0	134	30	23	12	8
433100	43002	16	1	0	0	0	14	0	1	0	0
433200	43002	8.463	259	83	14	2	2.664	612	392	212	58
433300	43002	999	12	1	0	0	337	45	31	17	4
433410	43002	3.381	172	51	13	0	516	185	163	96	12
433420	43002	338	8	0	0	0	121	37	13	11	0
433900	43002	493	4	1	0	0	82	7	9	2	0
439100	43009	565	11	1	0	0	257	41	29	23	5
439910	43009	3.232	121	34	4	1	1.049	294	179	83	14
439990	43009	1.976	43	8	2	1	738	103	80	49	8

Number of enterprises in the sample by groups ("strata")

Industry code	Owner-ship	Size group	No. selected
41000 - 00	P+S	0-10 employees [0+1]	18
		10-19 employees [2]	29
		20-49 employees [3]	82
		50+ employees [4]	98
42000 - 01	P+S	0-19 employees [0+1+2]	5
		20-49 employees [3]	9
		50+ employees [4]	13
42000 - 02	P+S	0- 5 employees [0]	5
		5-10 employees [1]	8
		10-19 employees [2]	16
		20-49 employees [3]	24
		50+ employees [4]	35
43001 - 01	P+S	0- 5 employees [0]	6
		5-10 employees [1]	23
		10-19 employees [2]	38
		20-49 employees [3]	91
		50+ employees [4]	57
43001 - 02	P+S	0-10 employees [0+1]	26
		10-19 employees [2]	36
		20-49 employees [3]	70
		50+ employees [4]	23
43001 - 03	P+S	0-49 employees [0+1+2+3]	14
		50+ employees [4]	8
43002 - 02	P+S	0-19 employees [0+1+2]	90
		20-49 employees [3]	108
		50+ employees [4]	52
43002 - 03	S	0-10 employees [0+1]	10
		10-19 employees [2]	20
		20-49 employees [3]	53
		50+ employees [4]	15
	P	0-10 employees [0+1]	8
		10-49 employees [2+3]	11
43002 - 04	P+S	0-49 employees [0+1+2+3]	10
43002 - 09	P+S	0-49 employees [0+1+2+3]	4
43009 - 00	P+S	0-10 employees [0+1]	8
		10-19 employees [2]	8
		20-49 employees [3]	17
		50+ employees [4]	11
43009 - 01	P+S	0-19 employees [0+1+2]	12
		20-49 employees [3]	10
		50+ employees [4]	5
43009 - 02	S	0-10 employees [0+1]	19
		10-19 employees [2]	18
		20-49 employees [3]	38
		50+ employees [4]	12

	P	0-49 employees [0+1+2+3]	6
		50+ employees [4]	1
43009 - 03	S	0-10 employees [0+1]	9
		10-19 employees [2]	16
		20-49 employees [3]	24
		50+ employees [4]	7
	P	0-49 employees [0+1+2+3]	7
		50+ employees [4]	1