Statistics Finland

Compiling Administrative and Survey Data for SBS

BiH, Banja Luka
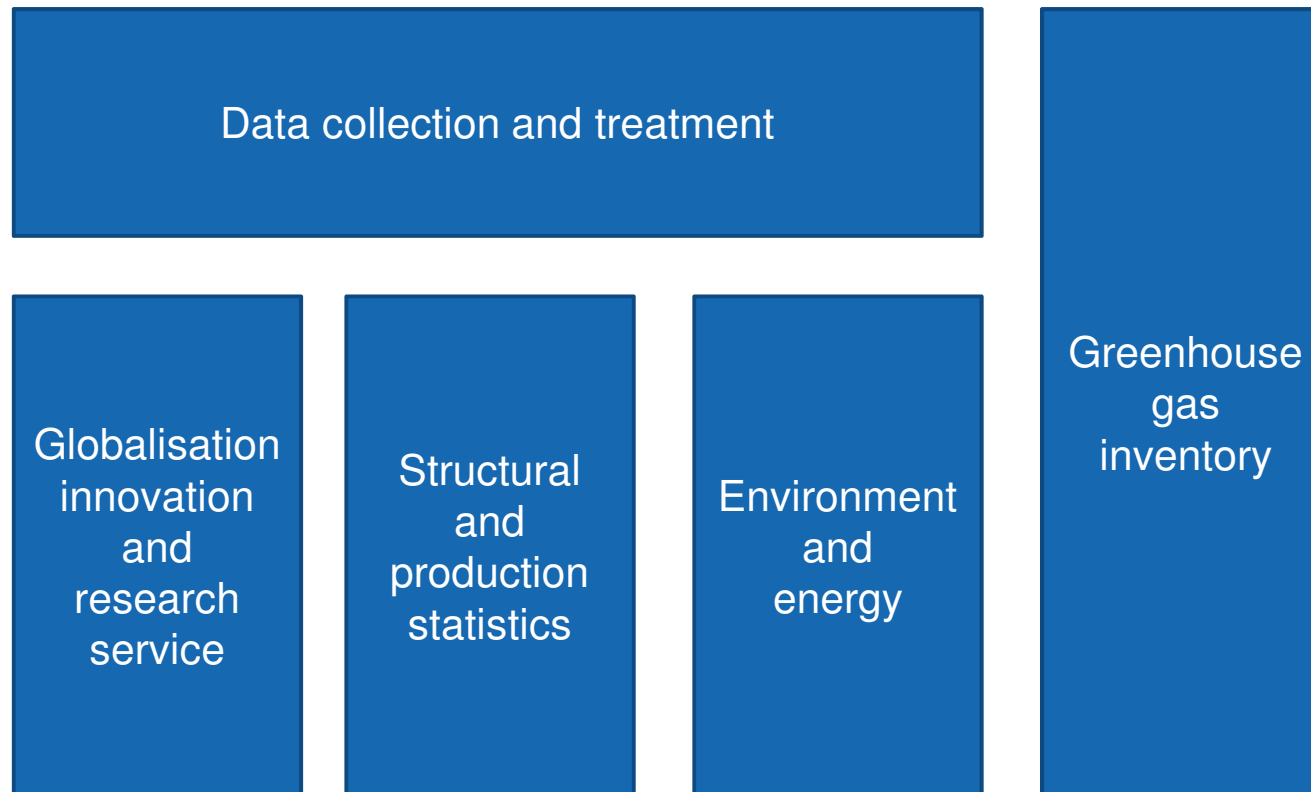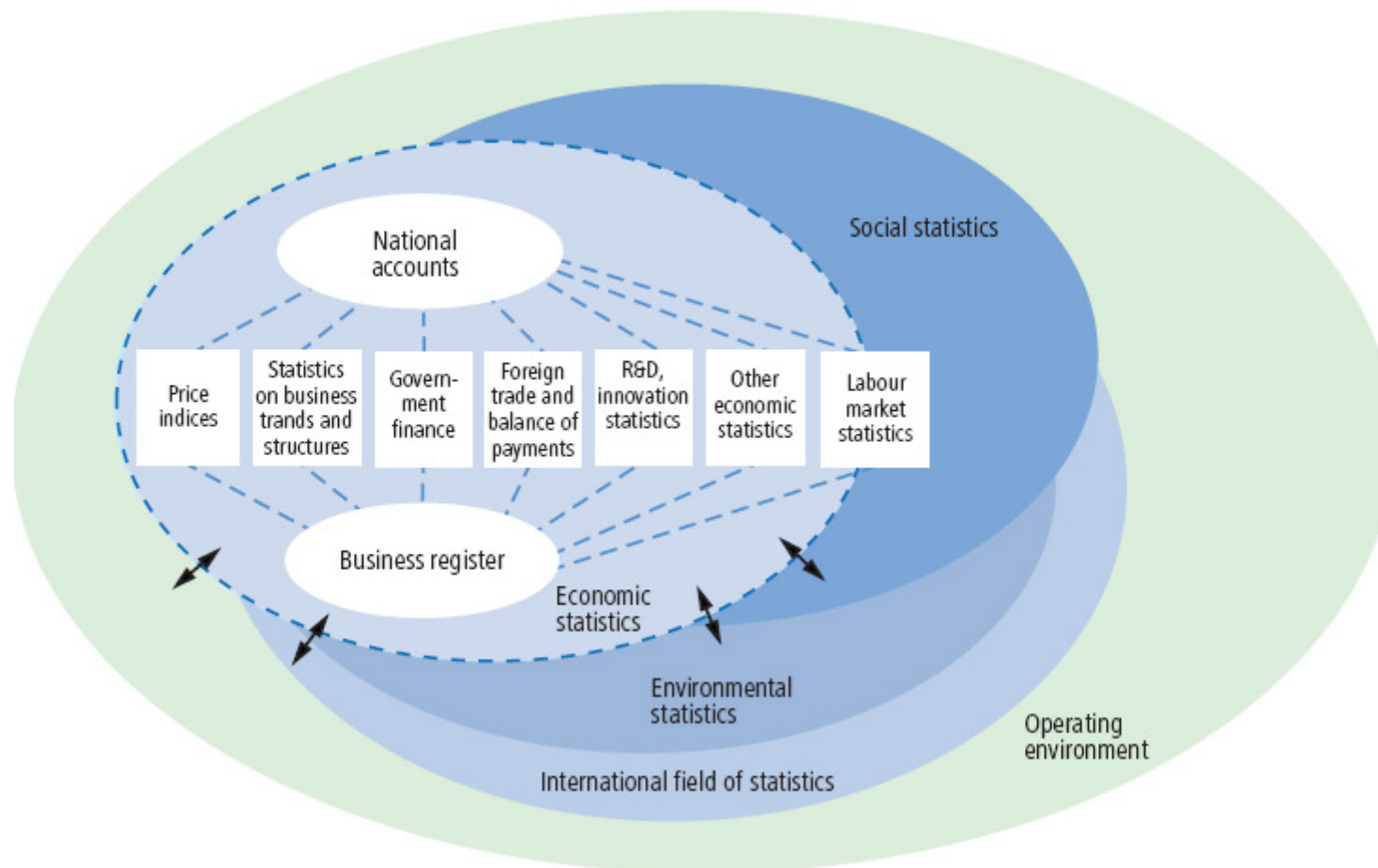
Ville Tolkki

Statistics Finland

Director General
Heli Jeskanen-Sundström

Advisory Board of Statistics Finland

Scientific Advisory Board

Co-operation Committee

Secretariat of Director General

Management Services

Information Services

Library of Statistics

IT and Statistical Methods

Statistical Methods

Population Statistics

Social Statistics

Interview and Survey Services

Prices and Wages

Economic Statistics

Business Trends

Business Structures

27.2.2012

Statistics Finland

Organisation till 2012

- Sub-Unit: Structural Business Statistics, Profit-Unit: Business Structures
- Two products
  - Financial Statements Statistics for enterprises (FSS)
    - for 7 branches
  - Statistics by Region and Activity for LKAU's (SRA)
    - For 3 branches
- 13 persons
  - There of 7 in FSS and 3 in SRA
  - 2 persons working on maintenance and develompment of database and automated editing and imputation methods
  - 1 person working on IT sector

Organisation Business Structures 2012 onwards

Data collection and treatment			Greenhouse gas inventory
Globalisation innovation and research service	Structural and production statistics	Environment and energy	

Economic statistics at Statistics Finland
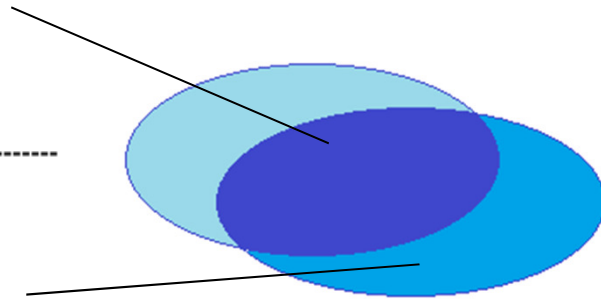
SBS versus Financial Statements Statistics

A common module of annual SS
A detailed module for SS in
- industry
- trade
- construction

- insurance
- credit institutions
- pension funds
- business services
- business demography

- SBS obligated by EC regulation 295/2008 (SBS regulation)

- consists of 9 modules
  - a common modulee for annual SS
  - detailed modules for SS in 8 different branches

- Financial statements statistics is based on domestic demand (National accounts, ministries)

- Produced from 1974

- 4 of SBS modules are compiled from FSS

- The rest modules are produced in other units at Statistics Finland

- Financial statements statistics includes several variables which are excluded from SBS

27.2.2012

Statistics Finland

Data Sources

- Administrative data
  - Business register
  - Income tax files
  - Official financial statements
  - Value-added tax data
- Own direct inquiry
  - Inquiry for enterprises
  - Inquiry for industrial establishments

Statistics Finland

Business Register

- Frame for FSS, SRA and SBS
- BR to SBS: Principal activity, personnel, location, owner type, institutional sector etc
- SBS to BR: Turnover
- Updates constantly, completed in November
- In structural statistics: smallest enterprises are excluded
  - Minimum constraints by: turnover, balance sheet, number of employees
  - The enterprise is excluded if none of the three constraints are full filled

Statistics Finland

Income Tax Files

- The main data source
- A census incuding all the enterprises paying income taxes in Finland
- Profit and loss account and statement of assets (incl items from balance sheet)
- Several data batches in order to improve the timeliness of the data
  - Taxation year is not necessarily the same as calendar year
  - Taxation year = accounting year

Statistics Finland

Direct Inquiry

- 1) Inquiry for enterprises:
  - Subsets of turnover and costs, investments
  - Used for National Accounts and FSS
- 2) Inquiry for industrial establishments
  - Turnover, costs, wages, operating profit and investments
  - Used for National Accounts and SRA
- Both samples drawn from business register
- Auxiliary information: Industrial branch, size class
  - Nace 2
  - Size class: enterprises with 50+ employees are all included
  - A random sample of enterprises with 10-50 employees is drawn
  - Sample for SRA is clustered: An enterprise picked in SRA sample must respond for every industrial LKAU's

Direct Inquiry, Step By Step

- Sampling frame: BR
- Units with 50 employees or more will all be included in the sample
- The rest of the units will be stratified by Nace (rev 2) into 140 stratas
- Relative allocation
- $n \approx 5\ 000$
- Atleast 10 units from each strata
- SRS within the stratas
- Ratio estimator
  - Auxiliary information: turnover, total costs (from tax data)

Direct Inquiry, tools

- Electronic questionnaire since 2006
- Electronic questionnaire is outsourced
- Electronic responses 2006 (50%), 2008 (60%), 2010 (95%)
- Paper questionnaire is not sent since statistical year 2008
- Response rate objective is 80%
- Data response control and treatment with inhouse software (powerBuilder)
- Database SQL server

General Overview

- Main data sources:
- Business register, 300 000 units, 40-50 variables
- Income tax files, 270 000 units, about 350 variables
- Direct inquiry, 5000 units (FSS) + 1200 units (SRA), about 100 variables
- Auxiliary/additional information:
- Value-added tax data
- Official financial statements
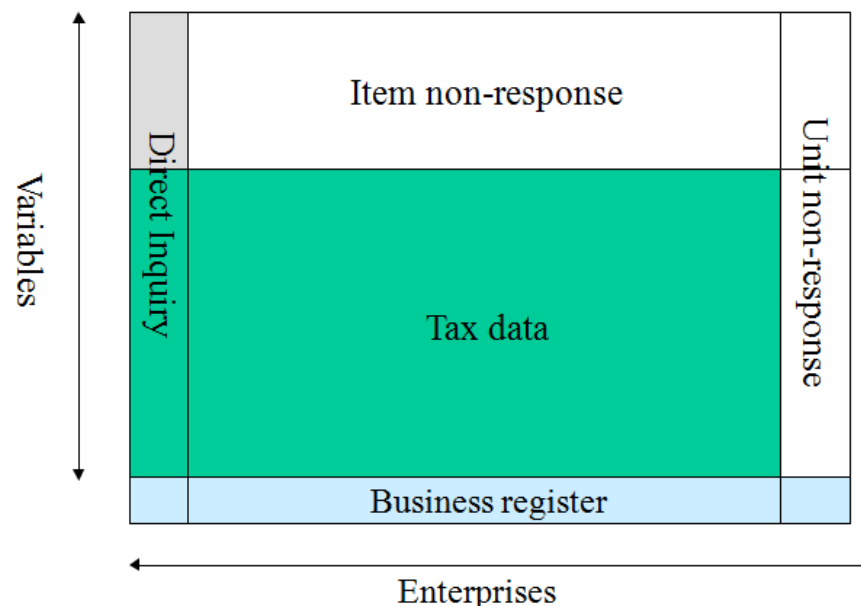
Value-added tax data

- Administrated by business register
- Used if profit and loss account is missing from income tax files
- Used for estimation of turnover
  - Both BR and SBS is using turnover estimated from value-added tax data until the profit and loss account is available
  - Both BR and SBS uses turnover from tax data in final BR and SBS
- Used as auxiliary information in editing and imputation

PDF:s of Official Financial Statements

- Income tax files: not the same as official financial statement
- The official financial statements are collected and stored by National Board of Patents and Registration of Finland
- Transmitted to Statistics Finland in PDF form
- Can not be processed automatically
- PDF:s are used in manual editing for the most influential errors

The Timetable of SBS Process

- Statistical year 2011:
  - Checking the contents of tax data and applying the data (Oct 2011)
  - Defining and updating the frame (Nov 2011)
  - Drawing the samples for own inquiries (Dec 2011)
  - Inquiries sent to reporting units (Feb 2012)
  - First download of tax data (Feb 2012)
  - Tax data batches 2nd to 4th (Apr-May 2012)
  - 5th data batch from tax authority (Aug 2012)
  - Preliminary results for enterprises (Sep 2012)
  - Preliminary results for LKAU's (Sep 2012)
  - Compiling preliminary SBS data and transmission to Eurostat (Oct 2012)
  - Closure of survey data treatment (Nov 2012)
  - Business register is completed. Final frame (Nov 2012)
  - 6th data batch from tax authority (Nov 2012)
  - Survey data, BR and tax data are merged
  - Publishing final results for enterprises (Feb 2013)
  - Publishing final results for establishments/LKAUS's (Feb 2013)
  - Compiling final SBS data and transmission to Eurostat (Jun 2013)
  - Quality report to Eurostat (Apr 2014)

Treatment on False and Missing Data

Statistics Finland
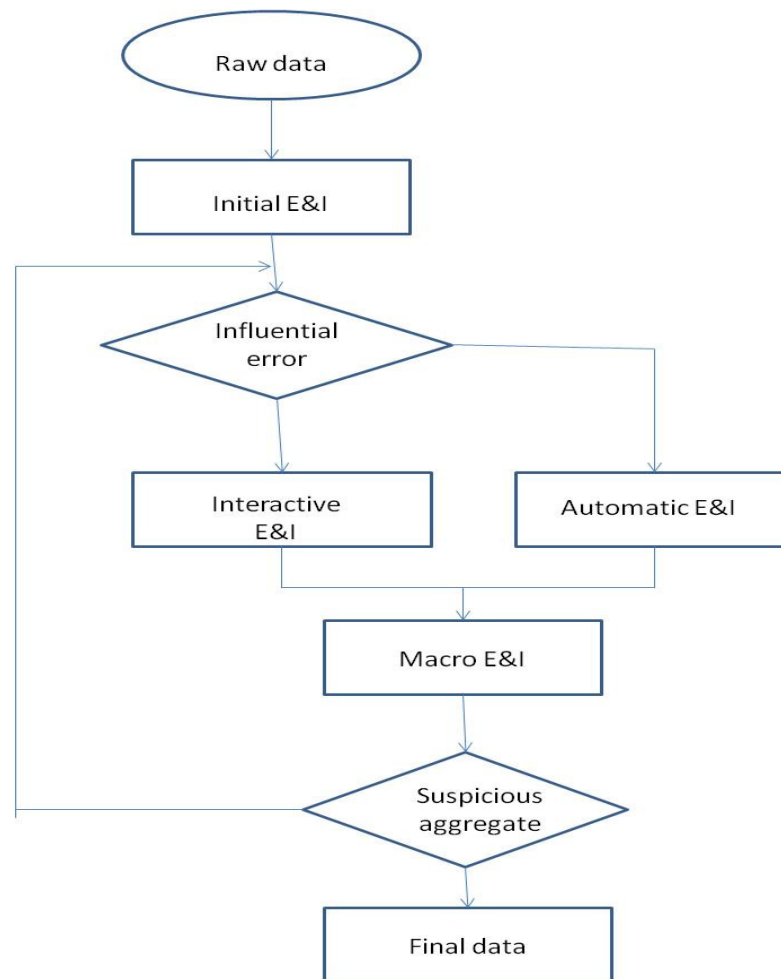
Detecting Influential Errors

- Selective editing
- Each unit will be scored according to following criterias
  - Fixed (local) scores:
    - Number of employees
    - Turnover
    - Balance
  - Relative (local) scores:
    - If the contribution of turnover, balance or number of employees within the branch is high -> high scores
    - The change of turnover, balance or number of employees from the previous year is high -> high scores
- Global score is sum of local scores
- Units with 100 points of more will be checked manually
  - Sorted by number of points -> priority list
- The rest will be E&I automatically

27.2.2012

Statistics Finland

## Treatment of non-influential errors

- Logical edits: Does the subsets of the tax statement sum to the subtotals and to the total
    - If yes, then the unit will be flagged as "valid"
    - If no, and the unit gets less than 100 points, the unit will be automatically fixed by
        - Adjusting outliers, if any
        - Scaling
        - Donor imputation (nearest neighbour)
- The fixed units are flagged

Statistics Finland

Imputing Missing Units

- The preliminary data is compiled when approximately 65-75 percent of the units and over 90 percent of the total turnover is received. This implies that unit non-response imputation has to be applied to overcome the problems of the bias in the data

- Tax data to business register: The non-response imputation process is based on the use of the previous year data for the particular units. Also the value-added tax data files are used to predict the turnover

- Variables from own direct inquiry to BR

  - Subsets of turnover and costs

- Ratio estimates are counted for each subset

  - Auxiliary information: turnover and total costs

- Auxiliary variable is splitted down by multiplying the value with each ratio estimate

The contribution of valid and imputed units
(Year 2009)
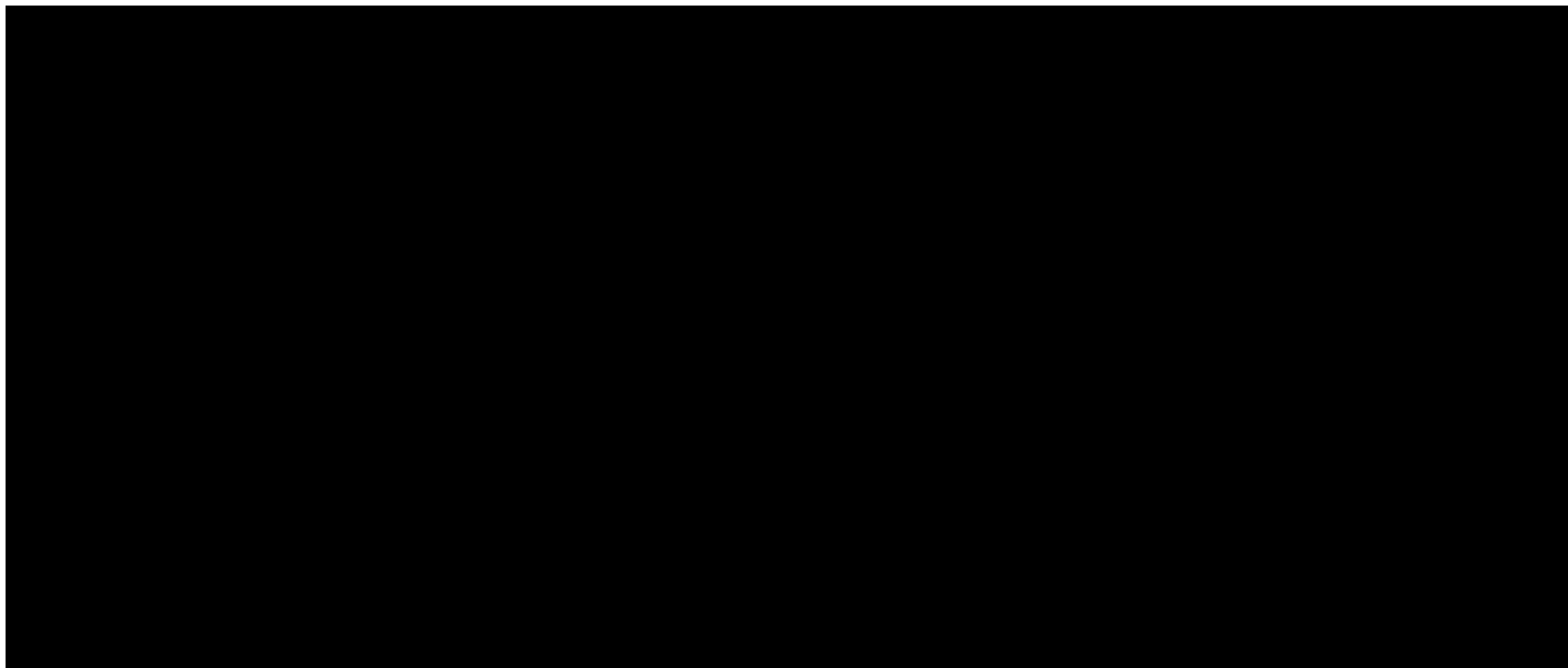
Benefits using tax data

1) Reduced response burden
2) Reduced costs
3) Improved coverage
4) Improved precision

Enterprises has a strong intensive to provide correct data
to the tax authority

Challenges Using Tax Data

- Adnimistrative data may be erroneous
    - Data entry errors
    - Inconsistency of concepts
- Scope of information is limited
- Statistical unit versus legal unit
- Dependency of administrative data

Future Developments

- Statistical unit: From legal unit to operational unit (enterprise, 2014)

- Own inquiry: Automatical E&I for non-initial errors (ongoing)

- Data warehouse project of the business register and business statistics (2009-2014)

- Board of managing taxonomy for mandatory business reporting (ongoing activity)

- Co-operation with tax authority in data collection, that is developing XBRL-tools