



Statistics Sweden

Statistiska centralbyrån

# Treatment and editing of tax data for Swedish Structural Business Statistics

Johan Erikson

Statistics Sweden

Economic statistics, methodological department



# Disposition

---

- Background
- The SBS survey today
- Main points using administrative data
- New design – increased use of tax data
- Further work





# Background

---

- Regulation on Structural Business Statistics (1997)
- Other main user: National accounts
- Annual report data, production and consumption, investment, employment





# The SBS survey today

---

- From 1997: total enumeration
- Large (> 50 employees) enterprises: questionnaire
- Smaller enterprises: admin data only
- Two sets of indicators, one common denominator
- SRU = Standardised accounts
- The income declaration





# Population / coverage

---

- The business register and the coordinated sample process gives the frame
- Mis-matches between tax data and the frame
- Missing enterprises: estimated
- Enterprises not in the frame: disregarded





# Missing indicators

---

- 1. Investment
- 2. Employment
- 3. Changes in equity
- Model calculations (1 and 3)
- Other (commercial) source (2)
- Model calculations are based on admin data for two years





# Calculating investments

---

Value at the beginning of the year

+ Investment

- Sales

+/- profit/loss from sales

- Depreciation

+/- Revaluations

= Value at the end of the year





# Editing procedures

---

- Summary checks
- Large changes – effects on aggregate figures
- Single indicators, first and foremost the calculated ones





## Editing procedures (2)

---

- Macro perspective used on micro level
- Does the change in a value for a single enterprise effect the aggregate as a whole?
- Thresholds: 50 million SEK (6 million EUR) and 2 per cent effect.
- Comparing different aggregates (changes in NACE code etc): the effect on both aggregates is checked





# Imputation

---

- “all or nothing”
- Mean values for the relevant activity code and size class
- Hierarchical model





## New design

---

- Only the largest enterprises surveyed by questionnaire only
- Main values from tax data, additional information from questionnaires
- Data collection in the autumn, preprinted values from tax data, less double collection
- Split samples (production/consumption investment, shares)





## Future work

---

- Automatic corrections
- Better and more effective editing
- Better models for imputation
- Improving the frame
- Use tax data in other surveys

