

# *Seminar on Registers in Statistics - methodology and quality*

*21 - 23 May, 2007 Helsinki*

## *Possibilities to use the VAT data for production of short-term statistics*

*Epp Karus  
Statistics Estonia  
Email- epp.karus@stat.ee*

### *Structure of presentation*

Results of the Pilot Project on short-term statistics (STS)  
Results of analyses on statistical and VAT turnover  
Plans for production of monthly trade statistics using VAT data

## *Chapter 1 Results of the Pilot Project on short-term statistics (STS)*

### *Chapter 1.1 STS Pilot Project*

Implementation time: May 2004 – May 2005

Data analysed: 2003

Objective: to analyse the usability of VAT data collected by Estonian Tax and Custom Board (ETCB) for the production of short-term statistics according to STS regulation

Analysed activities: construction, trade, services

Sources: SP, surveys of Statistics Estonia (SE), ETCB

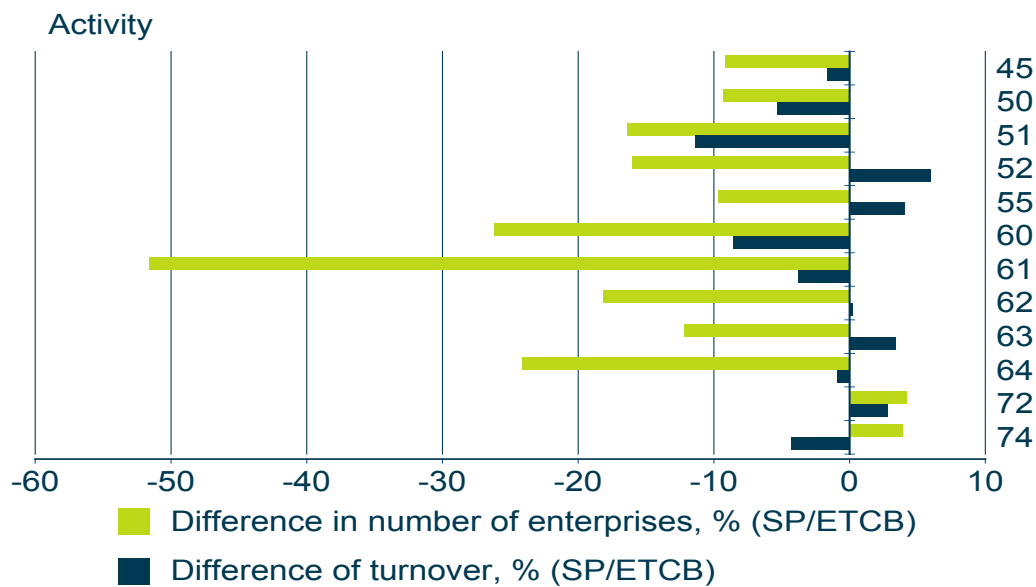
#### Actions executed

Estimation of the quality of VAT data:

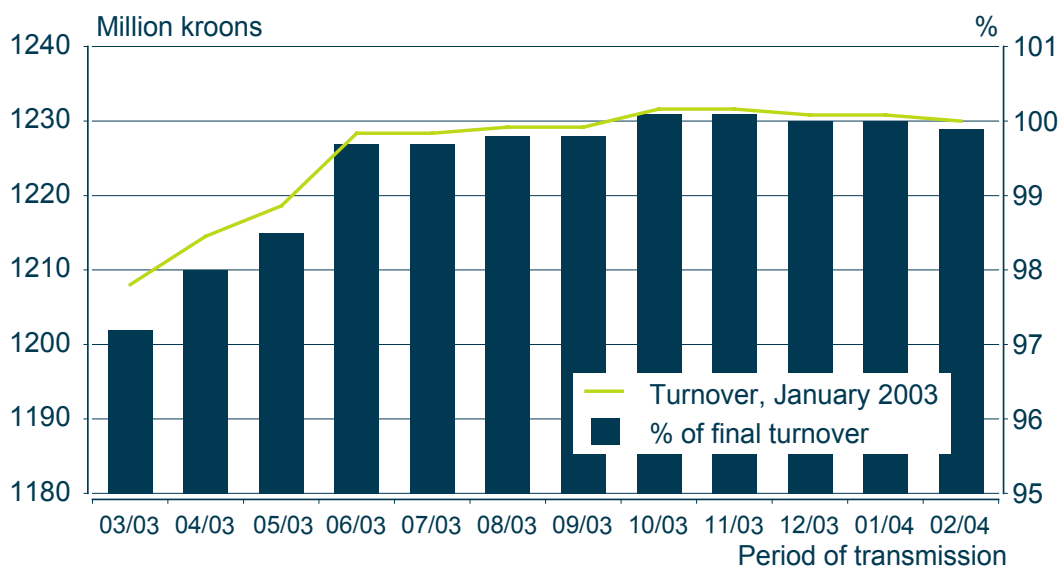
- Main activity of enterprise in SP and ETCB
- Change of VAT data of certain period during one year
- comparison of statistical and VAT data

Analysis of results

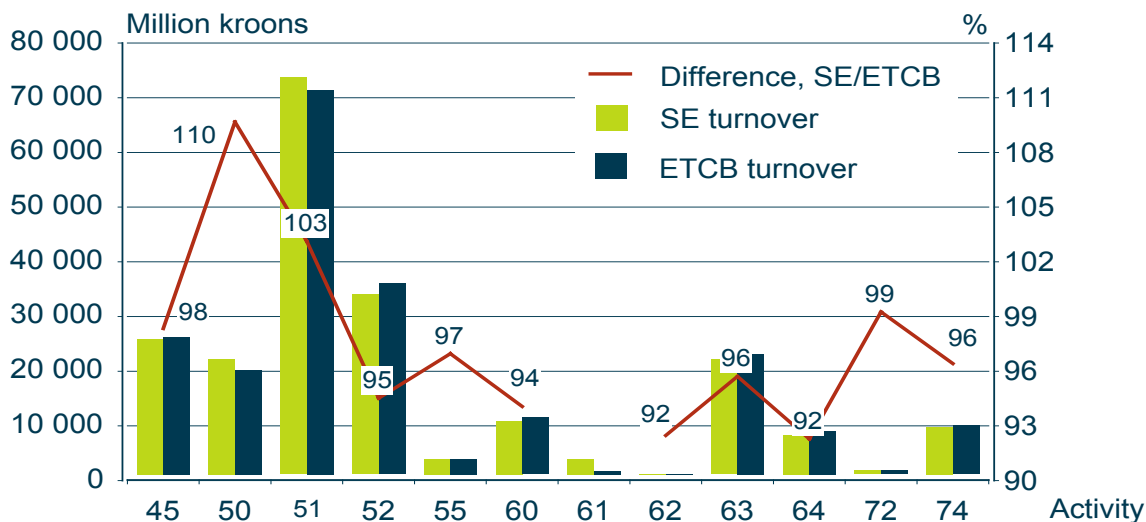
## Chapter 1.2 Difference of data in SP and ETCB, 2003



## Chapter 1.3 Change of ETCB turnover, January 2003



## Chapter 1.4 Turnover by activity, 2003



## Chapter 1.5 Conclusions

- Enterprises' activity in SP and ETCB is encoded differently. In general, different encoding is more frequent between small enterprises (less than 20 persons employed).
- Almost 90% of enterprises report VAT declaration to ETCB during two months after the reference period.
- 98% of final turnover is reported to ETCB during two months after the reference period. In most of activities, the turnover stabilizes 4-6 months after the reference period.
- Production of monthly construction statistics is possible using VAT data.
- Production of monthly retail trade (52) statistics is possible using VAT data, but the transmission of data from ETCB should be shifted to T+25 instead of T+30
- Turnover of transport enterprises (60-63) in SE and ETCB differs quite a lot. Finding out the reasons need further investigation.
- Turnover of service enterprises (55, 72, 74) from ETCB is usable, in some cases further investigation is needed.

## Chapter 2 Results of analyses on statistical and VAT turnover

### Purpose of analysis was to find out:

1. Variance of SE and ETCB turnover by activity and size groups

2. The reasons for differences
3. If and to what extent the VAT data are usable for the production of STS statistics for small enterprises

### ***Chapter 2.1 Availability of VAT data in SE***

Enterprises are obliged to report VAT declaration to ETCB during 20 days after the reference period.

ETCB and SE have data transmission contract. The VAT data are used to update the Business Register for Statistical Purposes (SP).

SE receives from ETCB the individual data of enterprises on the 1st day of the month (T+30).

Approximate statistical turnover is calculated using the components of VAT declaration.

### ***Chapter 2.2 Compared data***

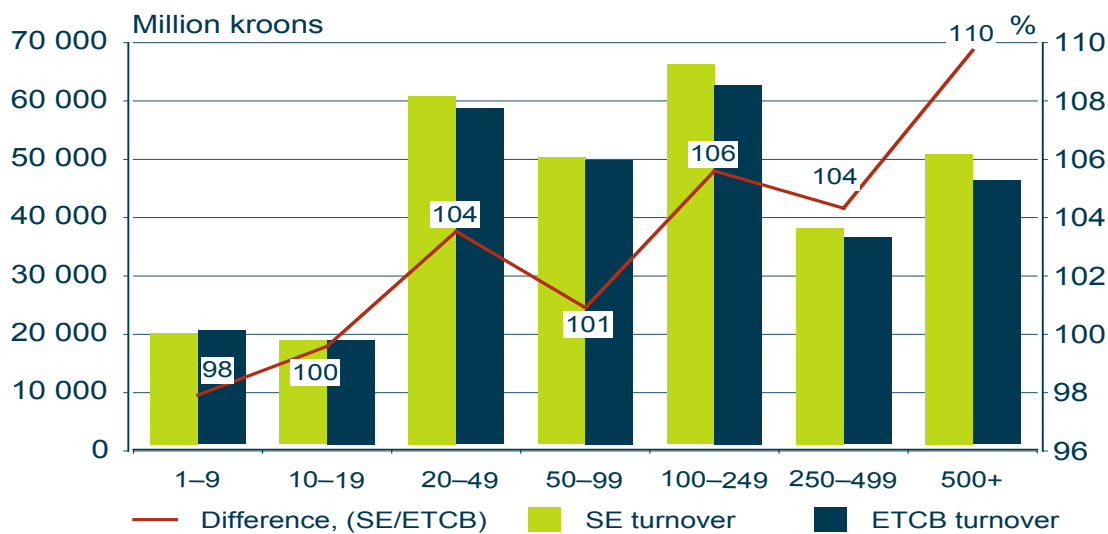
2005 data of enterprises, which reported to:

- SE the data of survey “Economic activity” for 4 quarters
- ETCB the VAT declarations for 12 months

Number of common enterprises was 4716.

Activity	Number of enterprises	% of enterprises
Manufacturing	1373	29
Construction	512	11
Trade	1221	26
Transport	322	7
Services	1288	27

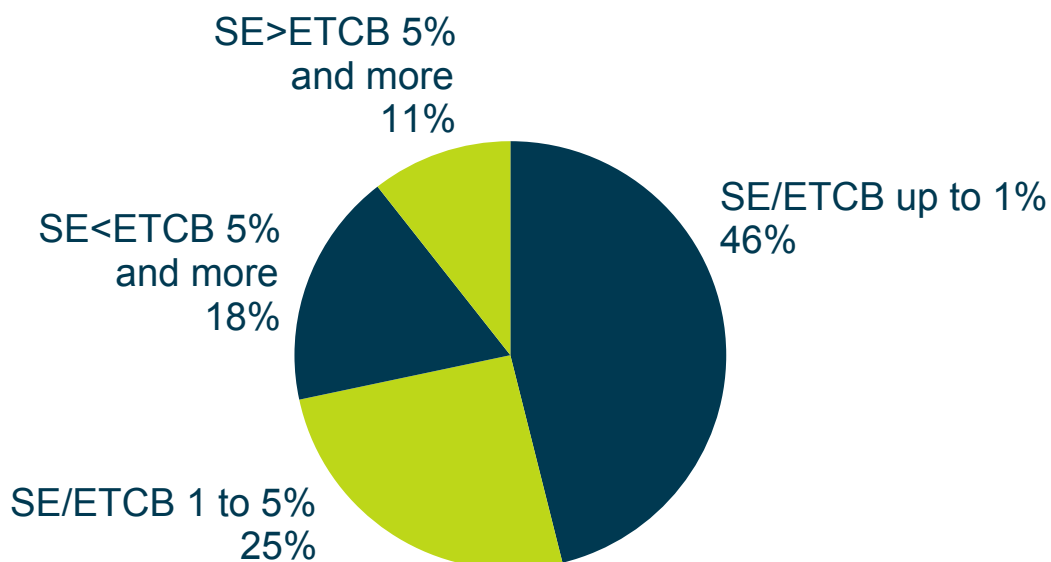
### ***Chapter 2.3 Comparison of SE and ETCB turnover by size of enterprise, 2005***



### Results:

- Total turnover of SE and ETCB differed 4%.
- The turnover in size groups 1-9 persons employed differed only by 2%.
- The turnover in size groups 10-19 persons employed coincide 100%.
- By 1% differed the turnover in size groups 50-99 persons employed. In remaining size groups the difference was 4-10%.

### Chapter 2.4 Comparison of individual data



### Results:

- SE turnover of common enterprises was 4% bigger than ETCB turnover.
- For almost half of enterprises, the turnover differed only by 1%.
- 29% of enterprises had turnover difference by 5%.

- For some enterprises, the turnover differed by times.

## ***Chapter 2.5 Reasons for differences***

VAT accounting differs from bookkeeping accounting

Main reasons of differences:

Excise taxes, sales of tangible assets, prepayments, sales in abroad, a parent undertaking and its subsidiaries are registered as a single taxable person, the moment of creating turnover is different, etc.

Different activities have different reasons

## *Chapter 2.5 Conclusion of analysis*

### Possible to use VAT data

- Compensation of non-response
- Production of statistics in size group 1-9 persons employed manufacturing (15-37), construction (45), trade, hotels and restaurants (50-55), road transport (602), partly services (63, 64, 72, 85, 90, 93)
- Production of statistics in size group 10-19 persons employed manufacturing (15-37), construction (45), trade, hotels and restaurants (50-55), road transport (602), partly services (72, 85, 90, 93)

### Impossible to use VAT data

- Mining and quarrying , electricity, gas and water supply (10, 40, 41)
- Transport via railways, water and air transport (601, 61, 62)
- Real estate, renting and business activities, education, recreational, cultural and sporting activities (70, 71, 73, 74, 80, 92)

## *Chapter 3 Plans for production of monthly trade statistics using VAT data*

### *Chapter 3.1 Reasons for renovation of monthly Retail trade survey*

Official Statistics Act §5. (1)

In conducting official statistical surveys, **data collected in state and other databases**, and data resulting from or collected in the course of the activities of state and local government agencies and other legal persons **shall primarily be used**, provided the composition and quality of collected data, in the opinion of agencies conducting statistical surveys, conform to the methodology of official statistical surveys.

The new data transmission contract will be concluded soon between SE and ETCB

Enterprises' complaints on high statistical burden

### *Chapter 3.2 Methodology development*

Best practices of other countries were followed while developing the methodology.

Population of the survey is based on SP. Motor and retail trade enterprises (50 or 52) are involved.

In 2007, the population is 6100 enterprises and sample is 1100 enterprises.

### *Chapter 3.3 Data sources*

#### Questionnaire of SE

Most important enterprises by size are included in the sample. Covered turnover should be 70-85% for each activity. Sample size is 300-400 enterprises, 700 enterprises will be removed from monthly sample.

#### Data of VAT declarations collected by ETCB

Sample survey or total survey

### *Chapter 3.4 Validation of VAT data*

The quantity of records to be validated increases 6 times - from 1000 enterprises to 6000 enterprises.

Large quantities of data should be validated during 2-3 days.

According to new contrast between SE and ETCB, the data should be transmitted to SE 24 days after the reference period. The data should be disseminated by SE on the 30th day after the reference period.

### *Chapter 3.5 Groups of validations*

Impact of errors to the final data is different. Levels of errors need to be determined and treated differently by focusing on most essential errors.

#### Essential errors

- statistician will decide the amount of corrections by consulting the data provider if necessary

#### Slight errors

- will be corrected automatically

#### Warnings

- will not be corrected

### *Chapter 3.5 Initial variable*

Initial variable could be **turnover or turnover index**. At present the initial variable is turnover, on the basis of which the turnover index is calculated.

The imputation of non-response depends on choice of initial variable.

- In case of SE questionnaire the non-response will be imputed with the data of previous period or with the average of stratum.
- ETCB non-response should be imputed if turnover is the initial variable. In case of turnover index only common enterprises of compared periods will be counted and non-response need not to be imputed.

### *Chapter 3.6 Estimation of Retail sales*

Retail sales is one part of turnover.

With SE questionnaire **turnover and retail sales** are collected.

ETCB data include only turnover. To that part of enterprises retail sales variable should be estimated using following sources:



- monthly retail trade data of important enterprises (SE questionnaire)
- quarterly and annual retail trade data by activity.

### ***Chapter 3.7 Plans for the future***

To receive recommendation for methodology from Methodology Department of SE

Compilation the order for development of data processing software, programming

Testing of methodology and data processing software by conducting two parallel surveys. Adjust necessary modification.

Start to produce monthly retail trade data using VAT data

### ***Further actions***

As a result of analysis it emerged that the VAT data could be used for the production of quarterly STS statistics

- The models and methods for using the ETCB data for the production of quarterly STS statistics should be elaborated
- Consult main users on production of turnover of not correlated variables (investments, inventories, financial leasing) in the future