Central Statistical
Bureau of Latvia

Comparison of two different BTS weighting systems in the services sector of Latvia

Ieva Vanaga
Ieva.Vanaga@csb.gov.lv

Central Statistical Bureau of Latvia

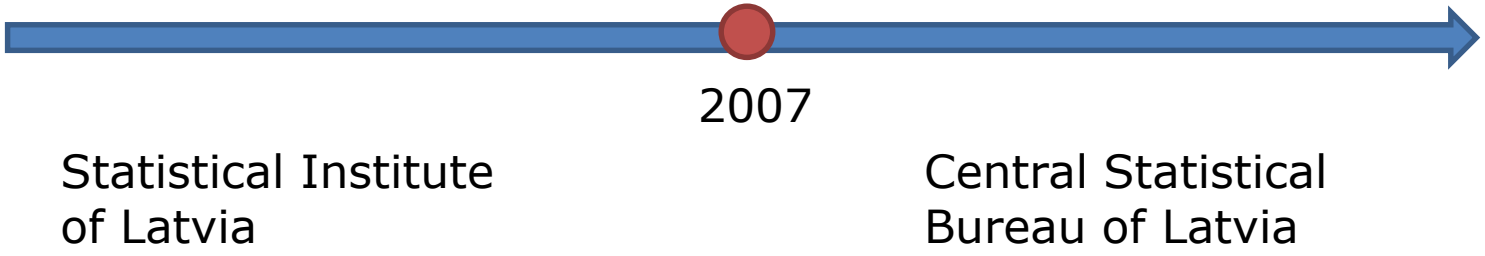
Structure of presentation

- Historical introduction on BTS in Latvia
- Role of CSB in the field of BTS
- Current weighting system of BTS data
- Subject and steps of activity
- First general results
- Results in different NACE groups
- Conclusions and questions to the further activities

Central Statistical Bureau of Latvia

History of BTS in Latvia

- Industry and construction surveys (1993)
- Retail trade survey (1996)
- Investment survey (2001)
- Survey in services sector (2002)



Central Statistical Bureau of Latvia

Role of CSB in the field of BTS

+ Planning and organizing surveys
+ Data collection
+ Calculation and dissemination of results

+ Development in step with the general statistics system of Latvia

- CSB does not perform in-depth economical analysis of BTS data

Central Statistical Bureau of Latvia

Current weighting system of BST data

- Enterprises are stratified by NACE and in 3 size groups by number of employees
- Primary data of each enterprise are weighted with coefficient 1, 2 or 3 depending on the size group. These weights are used in order to calculate the results in the necessary NACE breakdowns
- In order to calculate the overall results of the whole sector the share of every NACE group in the respective sector is used

Central Statistical Bureau of Latvia

Current weighting system of BST data

- Very simplified and unified model of significance of small, medium and large enterprises
- Simplicity as advantage – individual weights of enterprises stay unchanged throughout the whole calendar year
- Correlation with hard data are quite good, slightly lower for Industry survey
- How adequate can the description of real tendencies be if such a simplified method is used?

Central Statistical Bureau of Latvia

Comparison of two different weighting systems

1. Calculation of BTS results by using multiplication of sample weights and the number of employees as individual weights for primary data

1. Comparing results of both weighting systems

2. Comparing BTS data (both systems) with hard data

- At the moment first two steps of this list are completed.
- For the realizing of the last step we must accumulate data calculated by using sample/employees weights for seasonal adjustment.

Central Statistical Bureau of Latvia

Subject of activity – Survey of services sector

1. Services sector includes many subsectors (30)

2. Subsectors are very different:
   - by nature of economic activity
   - by number of enterprises in sample frame
   - by distribution of enterprises in size classes
   - by share of subsector in total

Central Statistical Bureau of Latvia

Sample frame of Services sector

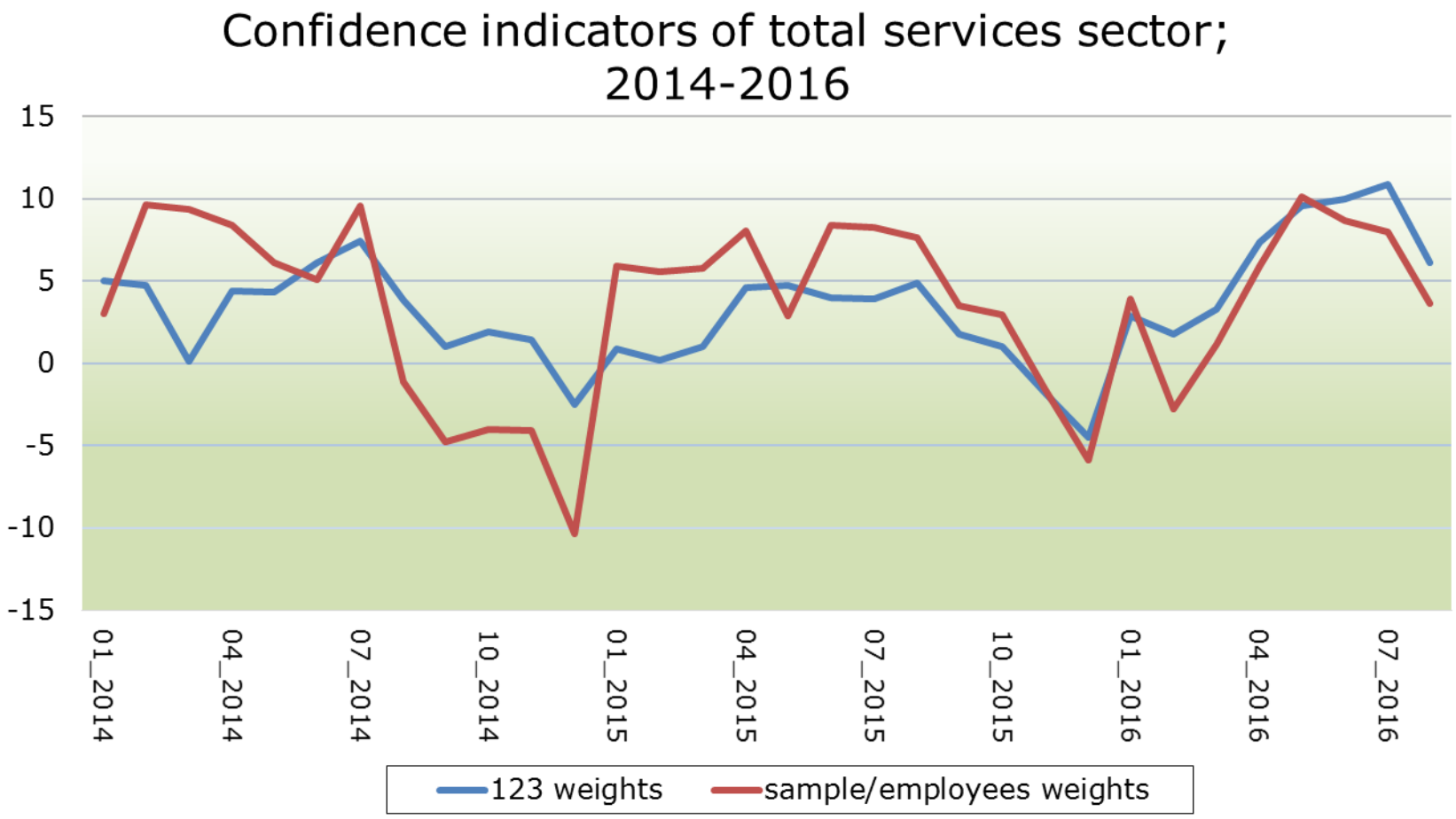
Frame of survey :

- ~26 000 enterprises (2016)
- stratified by NACE (30 groups at the 2-digit level)
- stratified by number of employees (3 size groups: 1-9; 10-49; 50+)

Only 600 frame enterprises (2.3%) belong to the group «50+»

Central Statistical Bureau of Latvia

First general results



Correlation between series 0.65

Central Statistical Bureau of Latvia

Analysis of subsectors

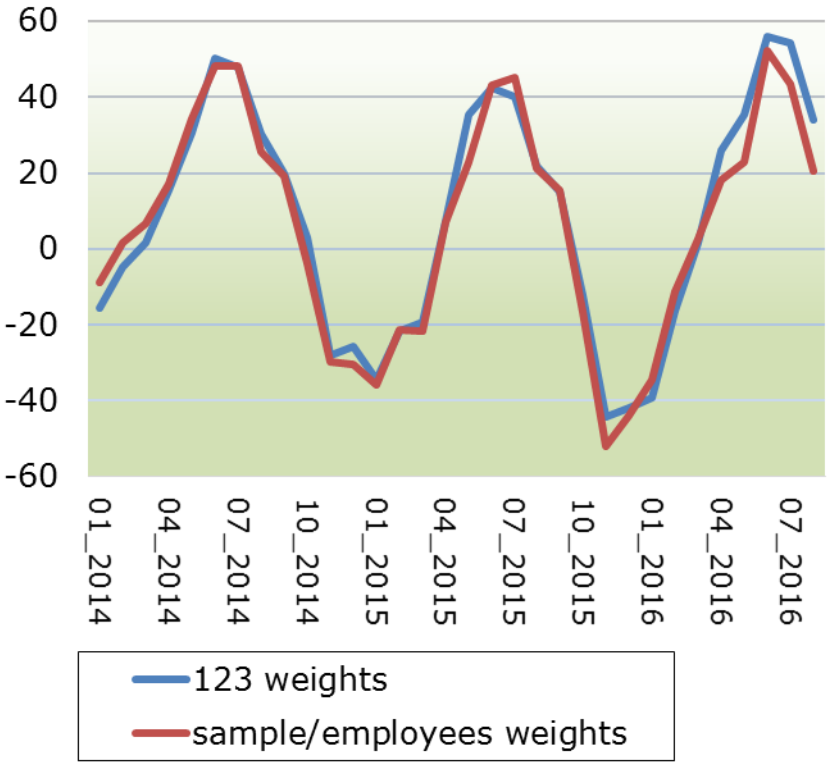
Coefficient of correlation (r)	Subsectors (NACE Rev 2)
High (r >= 0.7)	50, 51, 55, 56, 60, 61, 64, 66, 72, 77, 79, 80, 81, 82
Medium (0.5 <= r <0.7)	49, 58, 65, 68, 70, 75, 78
Low (r <0.5)	52, 53, 62, 63, 69, 71, 73, 74

We carry out a detailed analysis of every NACE group that shows low correlation as well as detect reasons for significant differences between results

Central Statistical Bureau of Latvia

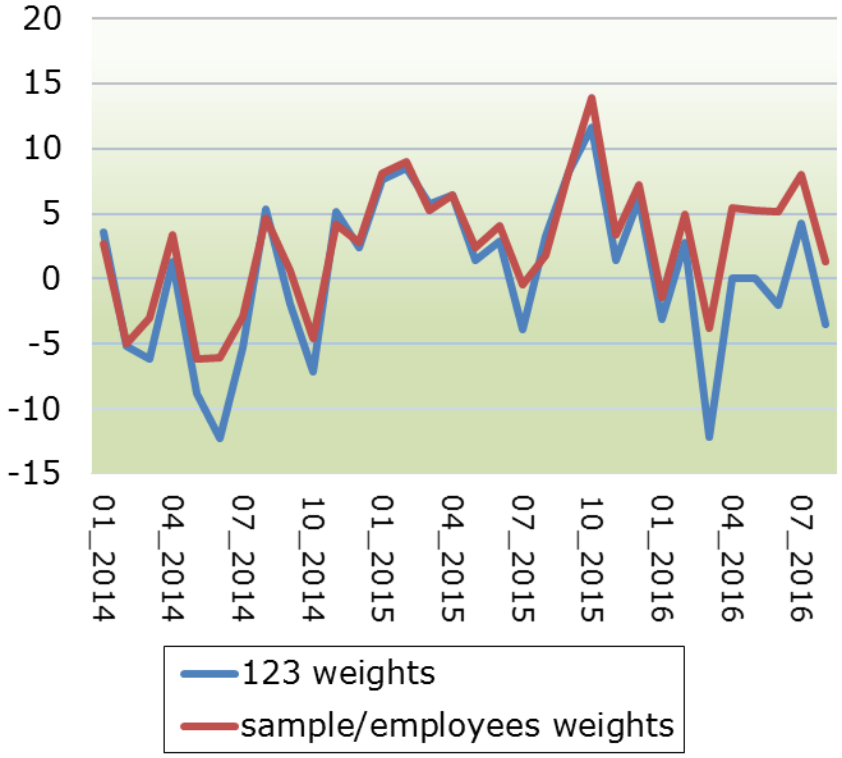
Subsectors 55 and 61

Confidence indicators NACE 55



Correlation 0.99

Confidence indicators NACE 61

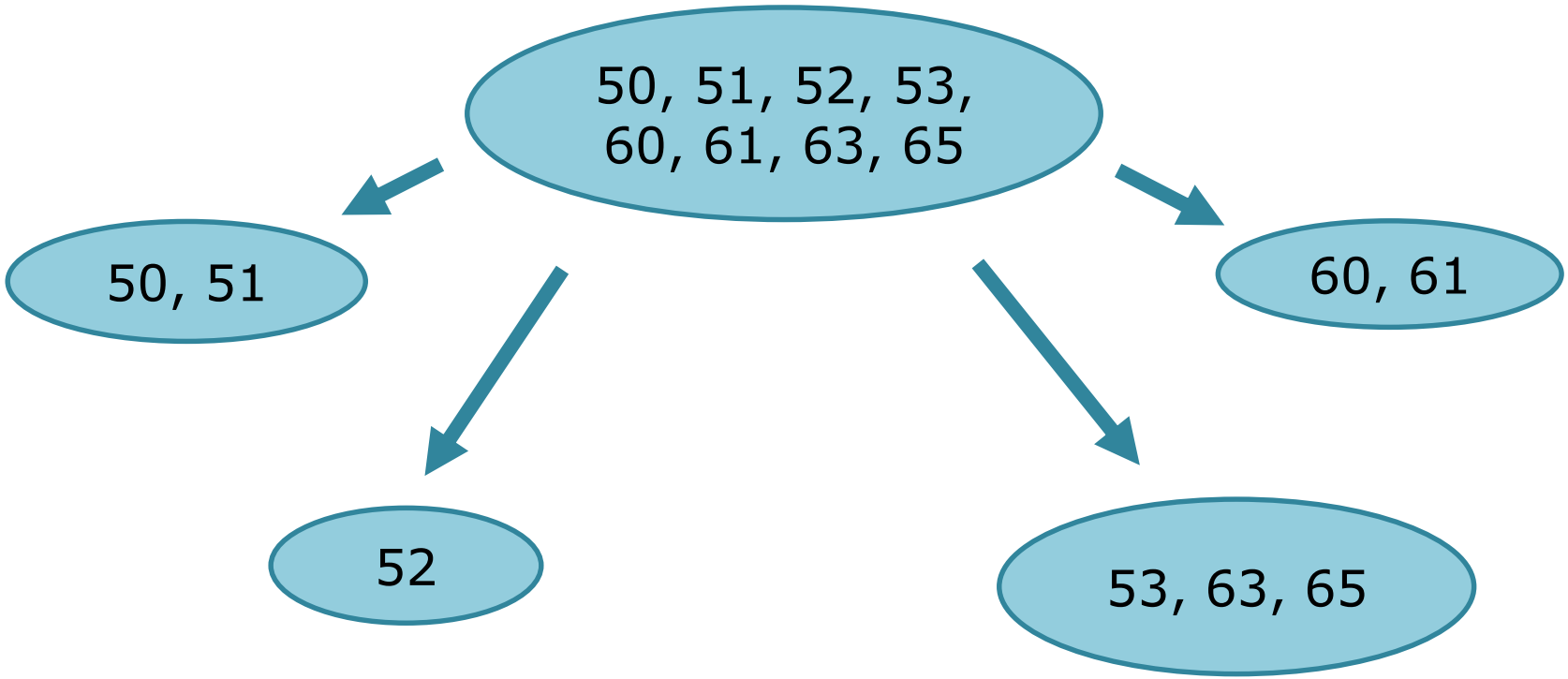


Correlation 0.97

Central Statistical Bureau of Latvia

Dominating enterprises – possible reason of low correlation

8 NACE groups with dominating enterprises –
4 different situations



Central Statistical Bureau of Latvia

Situation 1: low number of enterprises in sample frame

50 Water transport	51 Air transport
Correlation 0.87	Correlation 0.80
Low number of enterprises in sample frame (38 in 2016)	Low number of enterprises in sample frame (18 in 2016)
	
Values of confidence intervals differs significantly but in general by using of sample/employees weights we increase absolute value of confidence comparing with 123-weights method	

Central Statistical Bureau of Latvia

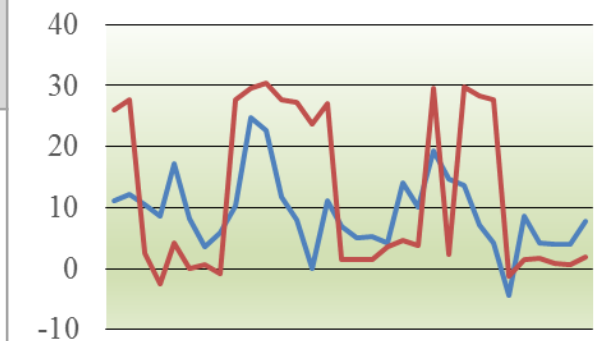
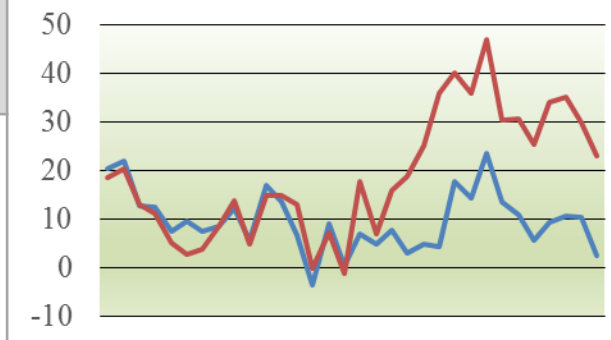
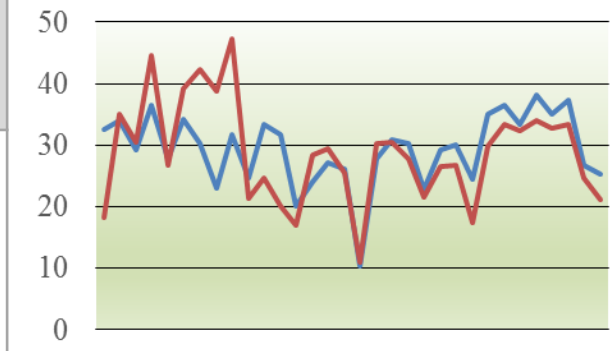
Situation 2: dominating enterprises respond only neutral answers

60 Programming and broadcasting activities	61 Telecommunications
Correlation 0.91	Correlation 0.92

- Dominating enterprises year after year submit only neutral answers
- It is the reason of very high correlation between results

Central Statistical Bureau of Latvia

Situation 3: low or medium correlation

53 Postal and courier activities (cor=0.46)	
• Small number of enterprises in frame • Variable answers of dominating(!) enterprise	
63 Information service activities (cor=0.47)	
• Significant changes in the list of large enterprises in the middle of research period	
65 Insurance and pension funding (cor=0.61)	
• Variable answers (2014) and neutral answers (2015 & 2016) of dominating enterprise	

Central Statistical Bureau of Latvia

Situation 4: NACE 52

52 Warehousing and support activities for transportation (cor=0.46)

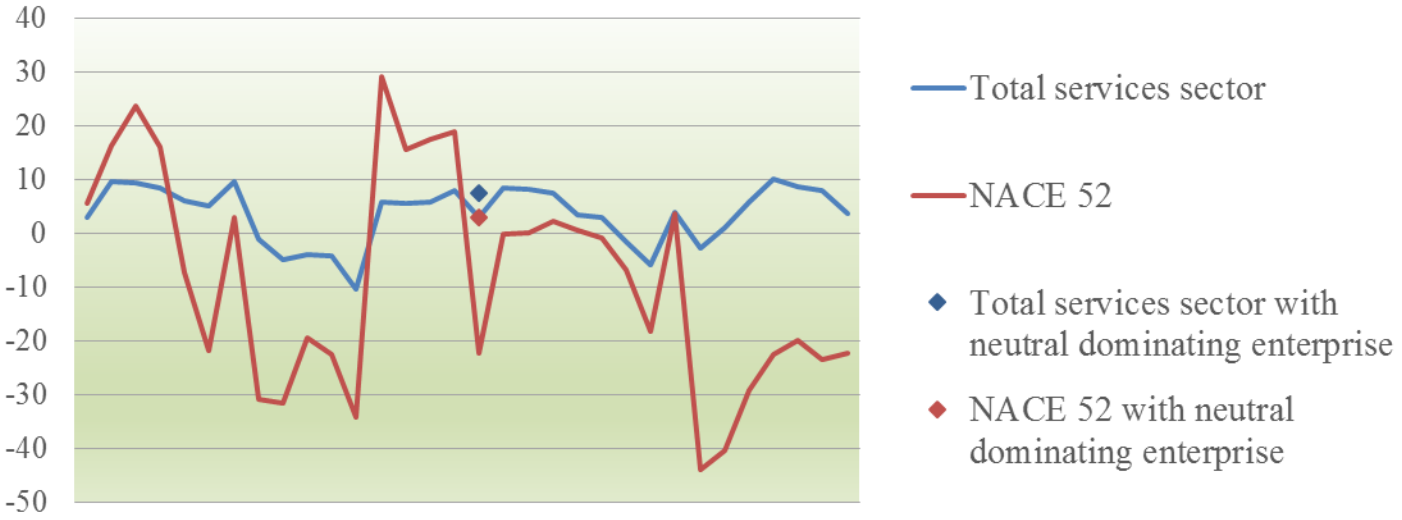
- Number of «50+» enterprises allows real sampling in this stratum
- Dominating enterprise has both very high number of employees and sample weight >1
- Results for NACE 52 (sample employees weights) are largely depending of answers of dominating enterprise

Central Statistical Bureau of Latvia

Situation 4: NACE 52

- NACE 52 is one of three subsectors with significant contribution in total service sector
- Using the sample /employees weights without any changes in stratification can give significant impact to results of total services sector

Confidence indicators using sample/employees weights



Central Statistical Bureau of Latvia

Recalculation

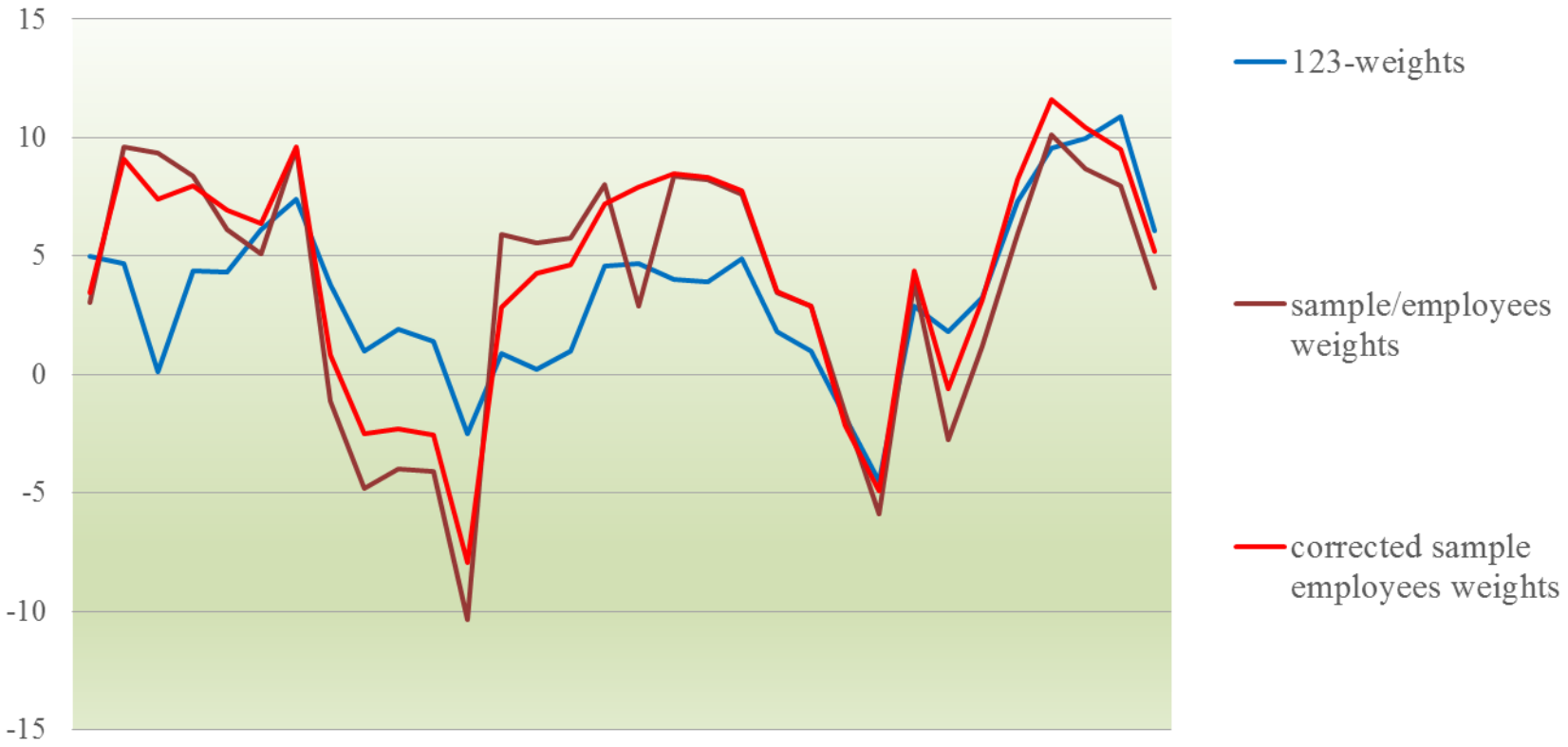
- To prevent the inadequate influence of dominating enterprise to total results we
  1) detected this enterprise (and also all dominating enterprises of other NACE groups) as an outlier with sample weight 1
  2) made recalculation with corrected sample/employees weights

- As result correlation between total services sector confidence indicators using 123-weights method and sample/employees weights increases from 0.65 to 0.80 after recalculation

Central Statistical Bureau of Latvia

Recalculation

Confidence indicators total services sector



Central Statistical Bureau of Latvia

Underrating of small enterprises– possible reason of low correlation

- 123 method currently underrates the significance of small enterprises (1-9 employees) in the NACE groups, where the role of small enterprises is the most important.

- NACE 62, 68, 69, 71, 73 and 74

NACE group	Number of enterprises within frame (2016) by employees		
	1-9	10-49	50+
62	1686	104	23
68	3496	323	42
69	2656	59	10
71	1256	147	21
73	1302	106	8
74	986	37	2

Central Statistical Bureau of Latvia

Next steps

- Comparing the results of both weighting methods with hard data

- Seasonal adjustment – available time series of sample/employees method at the moment are too short.

Central Statistical Bureau of Latvia

Conclusions

- BTS indicators calculated using 123-weights method are relatively flat. This weighting method can pass the significant changes in enterprises with high influence over.

- By yearly sampling we must be more careful looking at the distribution of enterprises by size classes in each NACE group to prevent underrating of small enterprises.

- Some explanations about BTS surveys and role of answers of each respondent is needed for enterprises, who year after year report only neutral answers, especially for large enterprises.

Central Statistical Bureau of Latvia

Conclusions / Questions

- BTS data of Latvia would be more qualitative and suitable for our users if we were going to change our current 123-weighting method to widely used sample-employees weighting

---

- How long back calculation would be accepted?

- Would be possible to realize the change of weighting method for all BTS during several years (not at the same time)?

Central Statistical
Bureau of Latvia

Thank you for attention!