

Use of preliminary Data from the Tax Authority in EU-SILC

by Rihard Inglic





Use of Preliminary Data from the Tax Authority in EU-SILC

Rihard Inglič

SURS

Statistical Office of the Republic of Slovenia

29 November 2016



Content

- Background of EU-SILC
- Timeliness
- Preliminary data vs final data
- Legislation
- Results of analyses
- Conclusions



Background of EU-SILC

- EU-SILC frame regulation
- Output harmonised survey
- Covers different areas:
 - Living conditions
 - Housing conditions
 - Work, employment
 - Health
 - Child care
 - Material deprivation
 - Ad hoc modules
 - **Incomes**

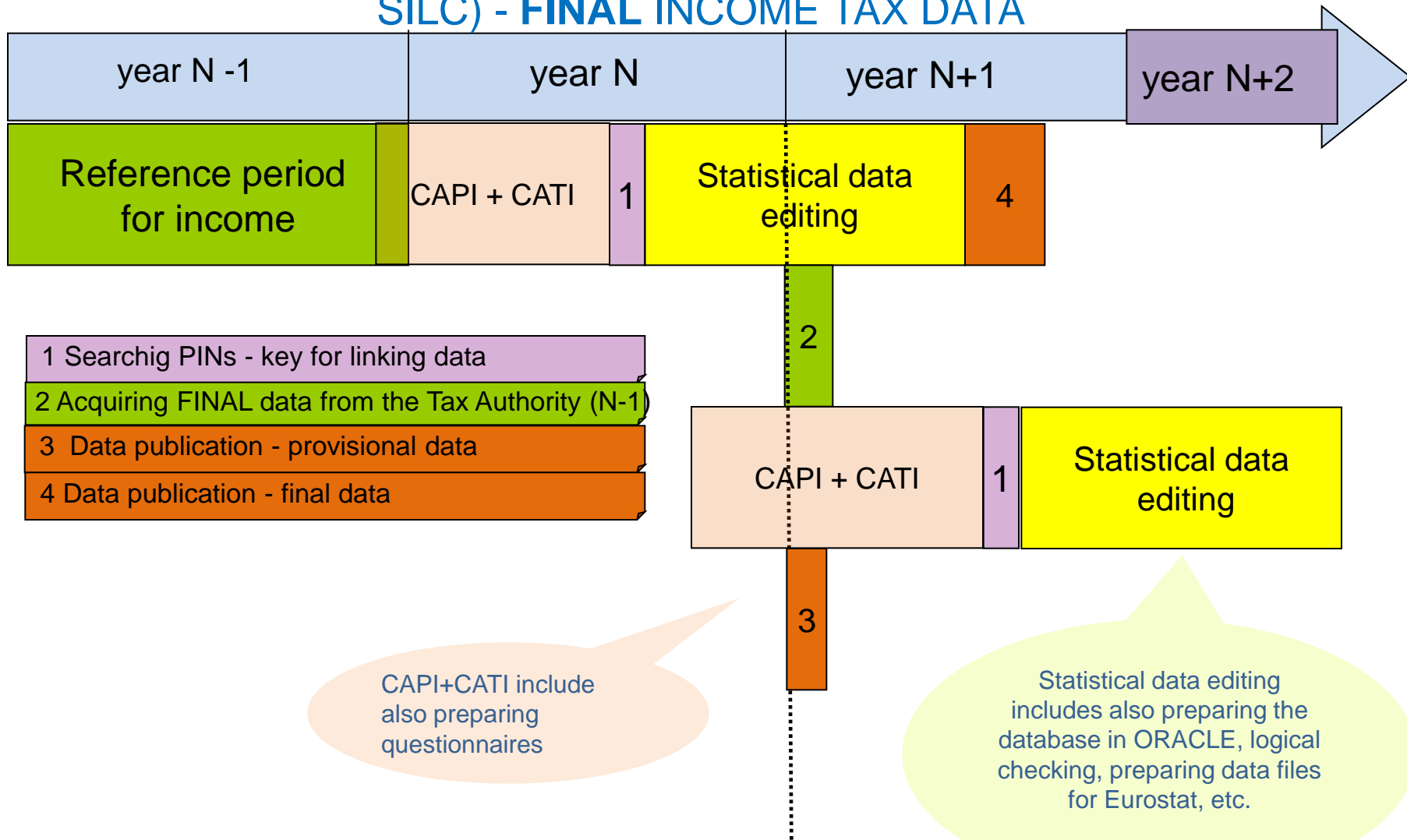


Timeliness

- Current situation (in Slovenia):
 - Collecting the data January-June N
 - Administrative sources – mostly in autumn N
 - Tax Authority: final data January N+1
 - reference period for income N-1
 - Data processing September N – June N+1
 - (at the same time collecting the data for year N+1)
 - Publishing the data:
 - Provisional: January N+1 (without income)
 - Final: July N+1



Phases of work on data in case of the use of administrative and register data (SI EU SILC) - FINAL INCOME TAX DATA





Preliminary vs final data ⁽¹⁾

- Comparing preliminary and final data
 - complete data files from the Tax Authority
 - in database for EU-SILC
- Test aggregates, averages, number of receivers of specific income
- Test income's indicators (at-risk-of-poverty-rate, structure of incomes)



Preliminary vs final data (2)

- Analysis – EU-SILC 2012 and 2013
- Raw data vs edited data (preliminary, final)
- Imputations, editing
- Rate of imputations
- Repeating the whole data processing EU-SILC with the preliminary income tax data



Legislation ⁽¹⁾

- „Income tax“ legislation
- Timelines according to the legislation
- By 31st January payers of income report about all incomes to the Tax Authority for year N-1
- The Tax Authority sends informative calculations of personal income tax to recipients of income in two batches
 - in March and May of year N



Legislation (2)

- 1 month time to object to the informative calculation of personal income tax (in case of mistakes, the change of tax reliefs, additional income, for example from abroad)
- Tax Authority in additional process makes the final decision
- All these processes finished by the end of year N



Legislation ⁽³⁾

Preliminary income tax data
received by SURS in June N

- do not include all complaints
- do not include repayments/receipts for tax adjustment



Legislation (4)

- Legal basis for using registers in EU-SILC
 - National Statistics Act
 - SURS has the right to get all administrative sources in Slovenia and use them for statistical purposes



Data files from the Tax Authority

- Several datasets, in EU-SILC we use 3 of them
- Approximately 70 kinds of incomes
- Approximately 1,700,000 persons are included in the dataset (from 2,000,000 inhabitants)
- **Variables:** kind of income, gross value, tax paid in Slovenia, social contributions, tax paid abroad, costs
- **Derived variable:** net income
- Data processing is done before the data are used in EU-SILC



Wage – whole dataset

Wage, wage compensations, reimbursement of expenses	PRELIMINARY 2012					
	GROSS			NET		
	SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE
	12.768.100.000 €	801.174	15.937 €	8.317.119.704 €	801.174	10.381 €
	FINAL 2012					
	GROSS			NET		
	SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE
	12.833.327.822 €	803.639	15.969 €	8.355.604.313 €	803.639	10.397 €
	PRELIMINARY EU-SILC 2013 (INCOME 2012)					
	GROSS			NET		
SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE	
12.629.300.000 €	787.513	16.037 €	8.248.884.458 €	787.513	10.475 €	
FINAL EU-SILC 2013 (INCOME 2012)						
GROSS			NET			
SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE	
12.756.799.386 €	792.466	16.098 €	8.334.249.619 €	792.466	10.517 €	

Source: FURS



Pensions – whole data set

Pensions	PRELIMINARY 2012					
	GROSS			NET		
	SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE
	4.210.377.388 €	600.420	7.012 €	4.195.040.968 €	600.420	6.987 €
	FINAL 2012					
	GROSS			NET		
	SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE
	4.212.844.180 €	600.490	7.016 €	4.197.478.594 €	600.490	6.990 €
	PRELIMINARY 2013					
	GROSS			NET		
SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE	
4.208.139.139 €	609.691	6.902 €	4.194.596.285 €	609.691	6.880 €	
FINAL 2013						
GROSS			NET			
SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE	
4.210.320.855 €	609.874	6.904 €	4.196.729.012 €	609.874	6.881 €	

Source: FURS



Income from contract work

Income from contract work	PRELIMINARY 2012					
	GROSS			NET		
	SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE
	335.793.517 €	170.054	1.975 €	263.044.189 €	170.054	1.547 €
	FINAL 2012					
	GROSS			NET		
	SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE
	336.624.037 €	164.654	2.044 €	210.948.243 €	166.813	1.265 €
	PRELIMINARY 2013					
	GROSS			NET		
SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE	
308.458.006 €	166.676	1.851 €	210.606.274 €	166.676	1.264 €	
FINAL 2013						
GROSS			NET			
SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE	
311.557.489 €	166.813	1.868 €	210.948.243 €	166.813	1.265 €	

Source: FURS



Profit from self-employment

Profit from self-employment	PRELIMINARY 2012					
	GROSS			NET		
	SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE
	437.091.673 €	61.942	7.056 €	367.392.668 €	61.942	5.931 €
	FINAL 2012					
	GROSS			NET		
	SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE
	437.268.659 €	61.976	7.055 €	367.555.510 €	61.976	5.931 €
	PRELIMINARY 2013					
	GROSS			NET		
SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE	
410.766.645 €	60.675	6.770 €	348.647.667 €	60.675	5.746 €	
FINAL 2013						
GROSS			NET			
SUM	RECEIVERS	AVERAGE	SUM	RECEIVERS	AVERAGE	
413.237.419 €	61.266	6.745 €	350.773.796 €	61.266	5.725 €	

Source: FURS



Difference by wages

	Ratio	Difference in percentage points
Ratio 2012 sum gross final/sum gross preliminary	100.51	0.51
Ratio 2012 sum net final/sum net preliminary	100.46	0.46
Ratio 2012 receivers final/receivers preliminary	100.31	0.31
Ratio 2013 sum gross final/sum gross preliminary	101.01	1.01
Ratio 2013 sum net final/sum net preliminary	101.03	1.03
Ratio 2013 receivers final/receivers preliminary	100.63	0.63

Source: SURS, FURS



Difference by pensions

	Ratio	Difference in percentage points
Ratio 2012 sum gross final/sum gross preliminary	100.06	0.06
Ratio 2012 sum net final/sum net preliminary	100.06	0.06
Ratio 2012 receivers final/receivers preliminary	100.01	0.01
Ratio 2013 sum gross final/sum gross preliminary	100.05	0.05
Ratio 2013 sum net final/sum net preliminary	100.05	0.05
Ratio 2013 receivers final/receivers preliminary	100.03	0.03

Source: SURS, FURS



Difference by incomes from contract work

	Ratio	Difference in percentage points
Ratio 2012 sum gross final/sum gross preliminary	100.25	0.25
Ratio 2012 sum net final/sum net preliminary	80.19	-19.81
Ratio 2012 receivers final/receivers preliminary	98.09	-1.91
Ratio 2013 sum gross final/sum gross preliminary	101.00	1.00
Ratio 2013 sum net final/sum net preliminary	100.16	0.16
Ratio 2013 receivers final/receivers preliminary	100.08	0.08

Source: SURS, FURS

Higher probability for mistakes in preliminary datasets.

A mistake in the 2012 preliminary dataset – the costs are not included in the data files.



HY145 - repayments/receipts for tax adjustment

- In preliminary data these data are **not included**
- Is it possible to use the data from the previous year (N-1)?
- Analysis of the FINAL income tax data - difference between N-1 and N data at the household level



HY145N – 2012

	Difference	Frequency	Percent
Current year lower than previous year for 5 000€ - 10 000€		5	0.07
Current year lower than previous year for 1 000€ - 5 000€		203	2.71
Current year lower than previous year for 100€ -1 000€		2007	26.84
Difference between years lower than 100€		3031	40.53
Current year higher for 100€ -1 000€		2039	27.26
Current year higher for 1 000€ - 5 000€		189	2.53
Current year higher for 5 000€ - 10 000€		4	0.05
Current year higher for more than 10 000€		1	0.01

Source: SURS, FURS

1,726 households have HY145N=0 in both years



HY145N – 2013

Difference	Frequency	Percent
Current year lower than previous year for 5 000-10 000€	6	0.08
Current year lower than previous year for 1 000-5 000€	264	3.65
Current year lower than previous year for 100-1 000€	2498	34.50
Difference lower than 100€	503	6.95
Current year higher for 100-1 000€	3667	50.64
Current year higher for 1 000-5 000€	294	4.06
Current year higher for 5 000-10 000€	9	0.12
Current year higher for more than 10 000€	0	0.00

Source: SURS, FURS

1,760 households have HY145N=0 in both years



HY145N

2012

HY145N	Min	Max	Mean	Median	Lower Quartile	Upper Quartile	Sum Wgts
Data used in final data	-6080.40	9520.85	-169.02	0	-373.70	0	794297.49
Data from previous year	-5397.73	14088.63	-161.86	0	-357.31	5.55	794297.49

2013

HY145N	Min	Max	Mean	Median	Lower Quartile	Upper Quartile	Sum Wgts
Data used in final data	-6062.02	20303.04	-201.06	0	-435.81	0	803840.00
Data from previous year	-6080.40	20841.23	-171.67	0	-378.20	2.03	803840.00

Source: SURS

Comparison of data in EU-SILC database

- Different combinations:

Preliminary raw data	Preliminary imputed and edited data
Final raw data	Final imputed and edited data

- The most important is difference among the data in the green fields



Wage – net (EU-SILC population)

- PY010 is composed of several variables from the administrative source and from 2 variables from the questionnaire
- In analysis we took into account only the most important variable from the administrative source
- (Unweighted) share of persons receiving wage - approximately 40%



Wage net – raw data - 2012

Difference preliminary raw vs final raw data	Freq	Percent
Final data lower than preliminary data for more than 10 000€	1	0.00
Final data lower than preliminary data for 1 000€-5 000€	30	0.11
Final data lower than preliminary data for 100€-1 000€	196	0.70
Difference lower than 100€	27783	99.00
Final data higher than preliminary data for 100€-1000€	12	0.04
Final data higher than preliminary data for 1 000€-5 000€	6	0.02
Final data higher than preliminary data for 5 000€-10 000€	8	0.03
Final data higher than preliminary data for more than 10 000€	28	0.10

Source: SURS, FURS

Income vs Employment status

- Income depends on employment status
- Employment status depends on income
- Some logical checks:
- if $\text{sum}(\text{PL073}, \text{PL074}) > 3$ and $\text{PY010G} = 0$ then „something should be done“
- if $\text{PL085} > 3$ and $\text{sum}(\text{PY100G}, \text{PY110G}, \text{PY130G}) = 0$ then „something should be done“



Imputations and editing

- All data from SILC were copied to another server; we added preliminary data in an additional table. The data were processed there.
- We used completely the same methods for data processing as in regular SILC data processing.
- For the data where we found mistakes in 2012 but not in further years, we took the final data also for calculating indicators from preliminary data



Effect size preliminary data vs final data – equvivalised income

	2012	2013
Number of units final data	27 265	28 064
Number of units preliminary data	27 265	28 064
Variance final data	43 265 545	40 505 242
Variance preliminary data	41 757 952	41 901 185
Standard deviation	6 520	6 419
Mean final data	12 639	13 485
Mean preliminary data	13 014	13 376
d	0.058	0.017
Regression	0.980297	0.998574

Source: SURS, FURS



Persons at risk of poverty (2012)

Unweighted data

Preliminary data	Final data		
	Poor	Not poor	Total
Poor	2804	184	2988
Not poor	106	24970	25076
Total	2910	25154	28064

Weighted data

Preliminary data	Final data		
	Poor	Not poor	Total
Poor	261282	13137	274419
Not poor	9406	1717100	1726506
Total	270688	1730237	2000925

Indicators 2012

- **At-risk-of-poverty threshold**
 - final data: 7,273€
 - preliminary data: 7,166€
- **At-risk-of-poverty rate**
 - final data: 13.5%
 - preliminary data: 13.7%
 - Difference is **0.2** of a percentage point

Persons at risk of poverty (2013)

Unweighted data

Preliminary data	Final data		
	Poor	Not poor	Total
Poor	2854	361	3215
Not poor	330	23270	24050
Total	3184	24081	27265

Weighted data

Preliminary data	Final data		
	Poor	Not poor	Total
Poor	265960	27777	293737
Not poor	24976	1691438	1716414
Total	290936	1719215	2010151

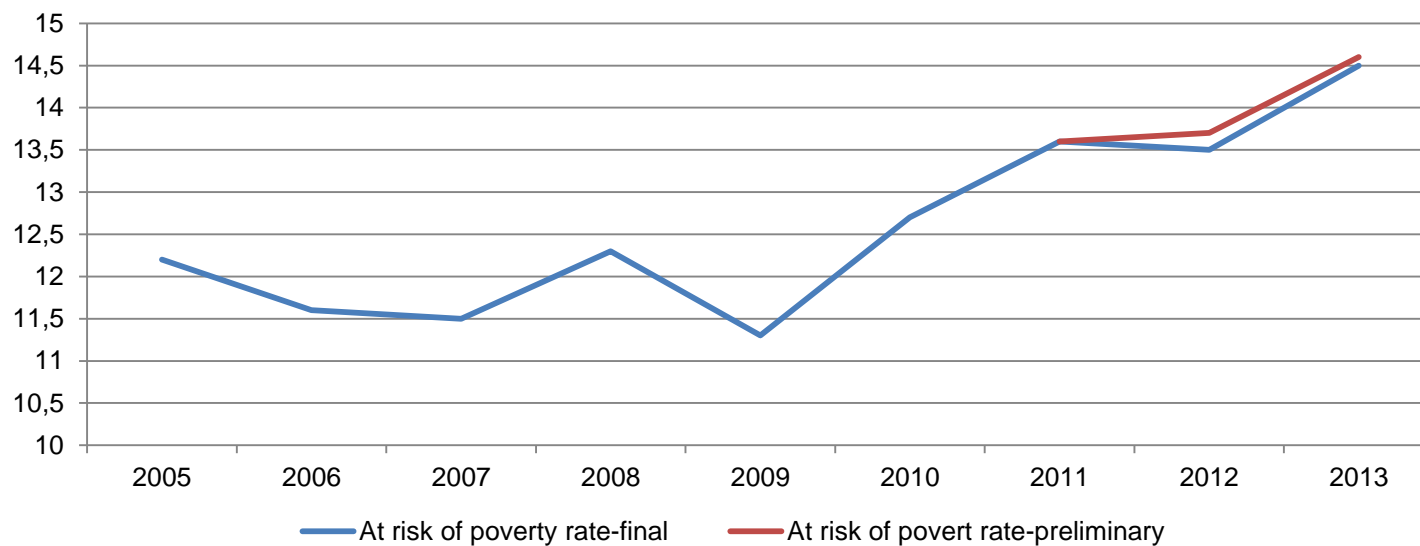


Indicators 2013

- **At-risk-of-poverty threshold**
 - final data: 7,111€
 - preliminary data: 7,379€
- **At-risk-of-poverty rate**
 - final data: 14.5%
 - preliminary data: 14.6%
 - Difference is **0.1** of a percentage point







At-risk-of-poverty rate



Source: SURS

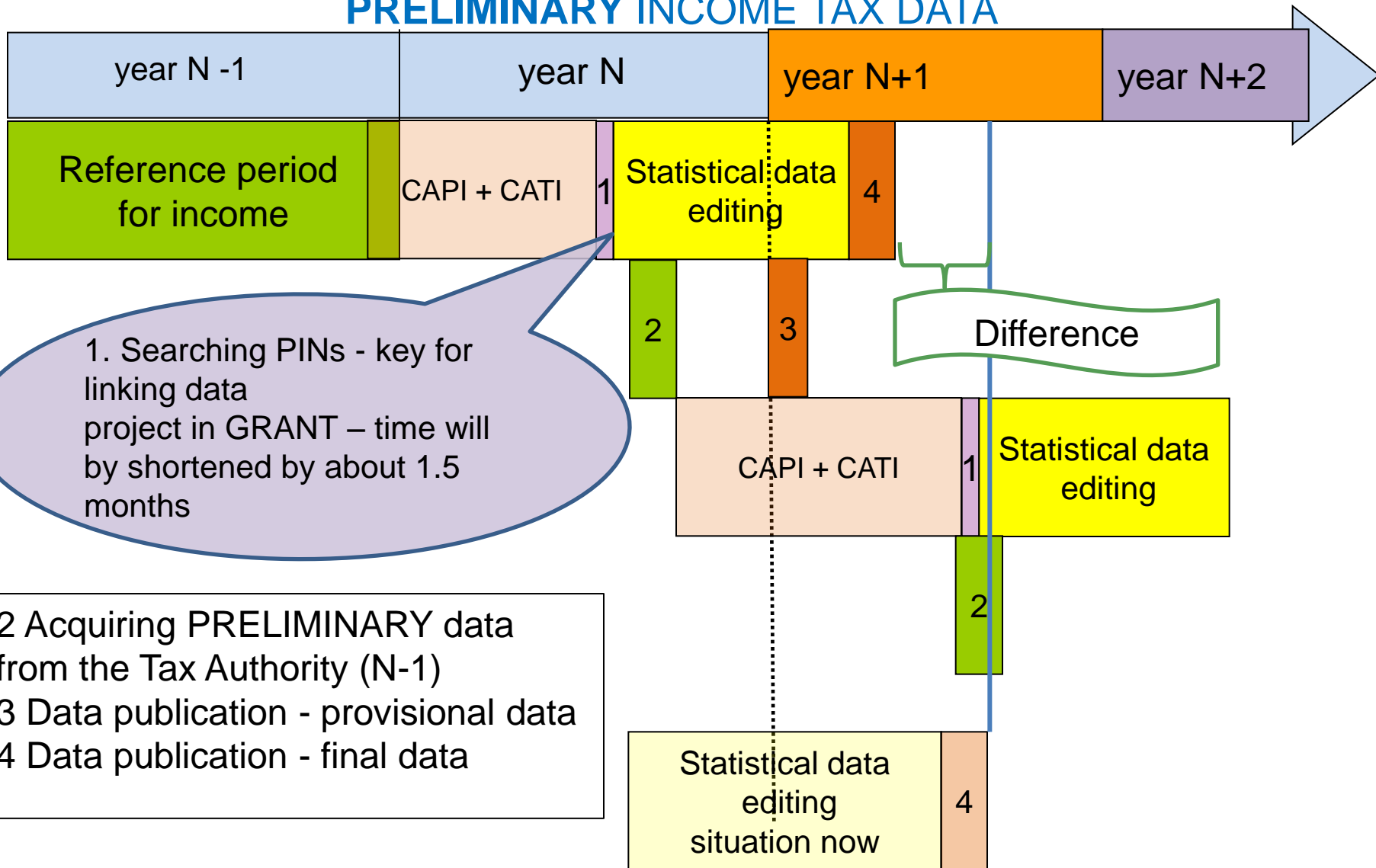
At-risk-of-poverty rate by age

		0-17	18-64	65+
2012	Below the at-risk-of-poverty threshold – preliminary	13.9	12.4	19.2
2012	Below the at-risk-of-poverty threshold – final	 13.5	12.2	 19.6
2013	Below the at-risk-of-poverty threshold – preliminary	13.2	13.1	22.7
2013	Below the at-risk-of-poverty threshold – final	 14.7	13.0	 20.5

Source: SURS



Phases of work on data in case of the use of administrative and register data (SI EU-SILC) - PRELIMINARY INCOME TAX DATA



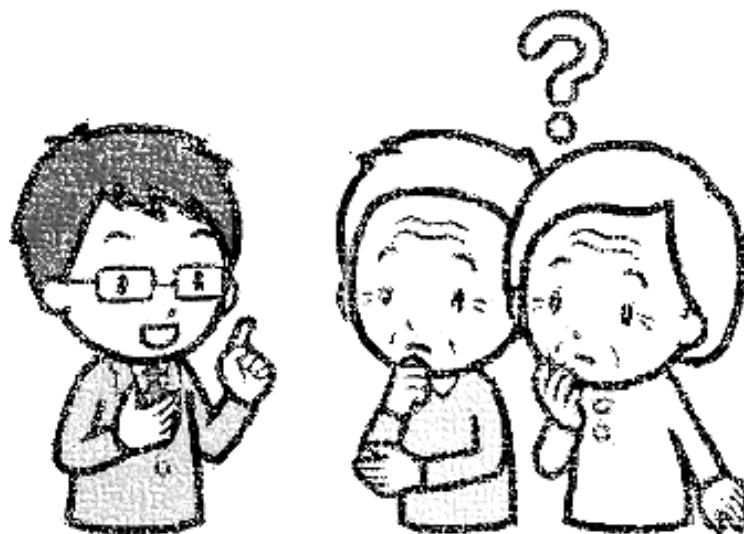


Conclusions

- Timeliness
- Quality of the data
- The main indicator does not differ a lot
- By categories the differences are much higher
- How many times should the data from SILC be published?
- PRELIMINARY income tax data = provisional or final SILC data?
- FINAL income tax data – to publish or not to publish as the final SILC data?
- What should be the explanation for the public in the case of differences?
- Policy makers – what would they do in the case of changing the data after some months?



Thank you for your attention



rihard.inglic@gov.si