

# Creation of the income variable from administrative data sources

Urve Kask

Statistics Estonia



## The starting point - why we want to use the administrative data sources?

The transition to collecting data from registers entails both positive and negative side effects.

- The positive ones are considered:
  - 1) to be a smaller response burden,
  - 2) a shorter interview time,
  - 3) more accurate data,
  - 4) less missing answers and lower costs.

## ES The starting point - why we want to use the administrative data sources

- The negative side (at least in the first few years)
  - 1) the time spent on and problems with processing data is increased;
  - 2) and time spent on making survey data compatible with registers is increased;

## The starting point in Estonia

Statistics Estonia has transitioned to using administrative data to create the income variable in three stages:

- The first analysis was done in 2005 and it mostly focused on mapping the data of the Estonian Tax and Customs Board and on the option of using this dataset for creating the components of the income variable;
- In the second stage in 2012, all registers were analysed regarding the possibilities for obtaining the necessary components to create the income variable.

<http://www.stat.ee/metoodika>

## The starting point in Estonia

- In the 2013 survey on the income of 2012, no such questions were asked which regarded variables that could be created from registers and which proved to be very compatible in the analysis of the survey and register data: old-age pension, pension for incapacity for work, survivor's pension, unemployment insurance benefit, benefits related to family and children.

## The first analyses in 2005 and main results

- The analysis included the following income components: monetary income from waged labour; contributions to private insurance plans; old-age pension; income from rent; regular payments to other households – used only one data sources data of the Estonian Tax and Customs Board

## The main results:

- The database of the Estonian Tax and Customs Board **also provides information on those whose income data was not included in the survey** (helps to improve data quality in the data organisation process)
- If we consider the income of only such persons/households the data of which is in both the survey and the register, **then there are generally no statistical differences in the data**. In several cases, the differences in the indicators of the entire sample are also not statistically significant.

## ES The main results:

- In the survey, **the number of salaried employees is smaller and the average wage level was higher.** The reason for this is the higher share of non-response among persons with lower wages.
- Although the average magnitude of the income from pension recorded in the survey and the register do not differ significantly, **the discrepancy in the number of respondents is extremely large** and can have an impact on the total income of the survey – the Estonian Tax and Customs Board data are not a suitable source for data on old-age pension.



# The options for gathering income data from various registers and the main results of the analysis in 2012

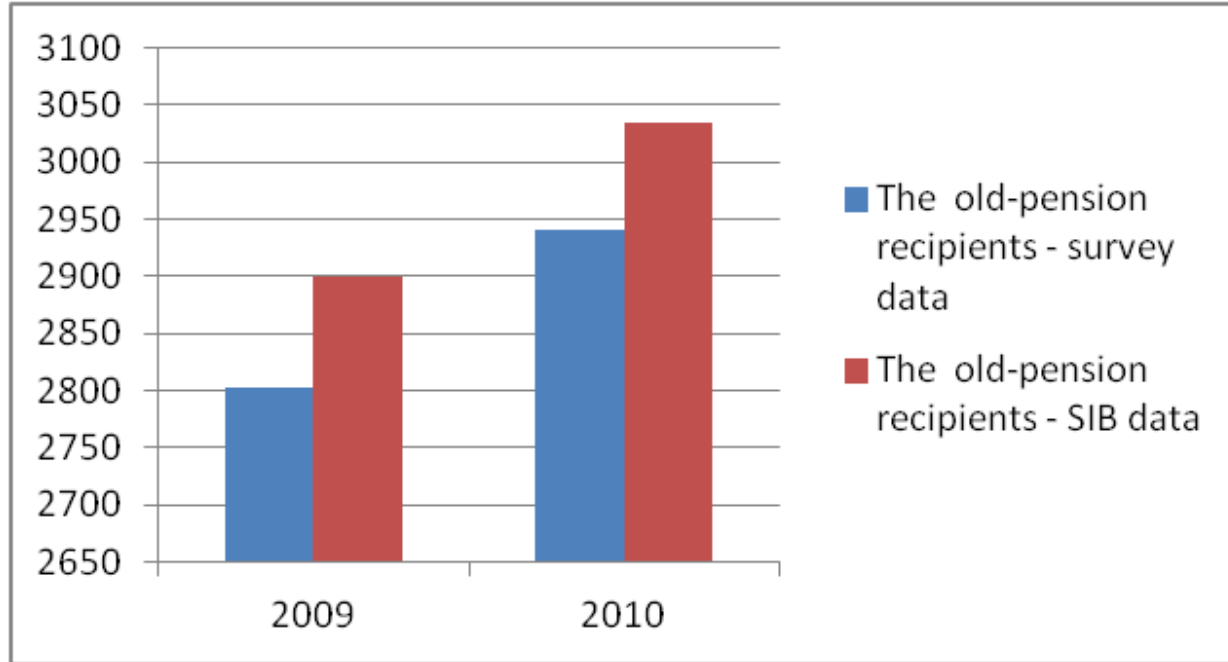
## Σ The coverage of income components by data sources

This compatibility analysis was based on three different register-based databases. With all three of them, Statistics Estonia has binding contracts, which guarantee the timely submission of data and the adequacy of individualised data.

1	<b>Estonian Tax and Customs Board (ETCB):</b>
1.1.	Tax from income
1.2.	Income from the employer
1.2.1.	Income tax refunds and additional income
1.3.	Unemployment insurance payment
1.4.	Redundancy payment
1.5.	<b>Old-age pension</b>
1.6.	<b>Parental benefit</b>
2	<b>Social Insurance Board (SIB):</b>
2.1.	Pension for incapacity for work
2.2.	Survivor's pension
2.3.	Child benefit
2.4.	<b>Old-age pension</b>
2.5.	<b>Parental benefit</b>
3	<b>Estonian Health Insurance Fund (EHIF)</b>
3.1.	Sickness benefit
3.2.	Maternity benefit

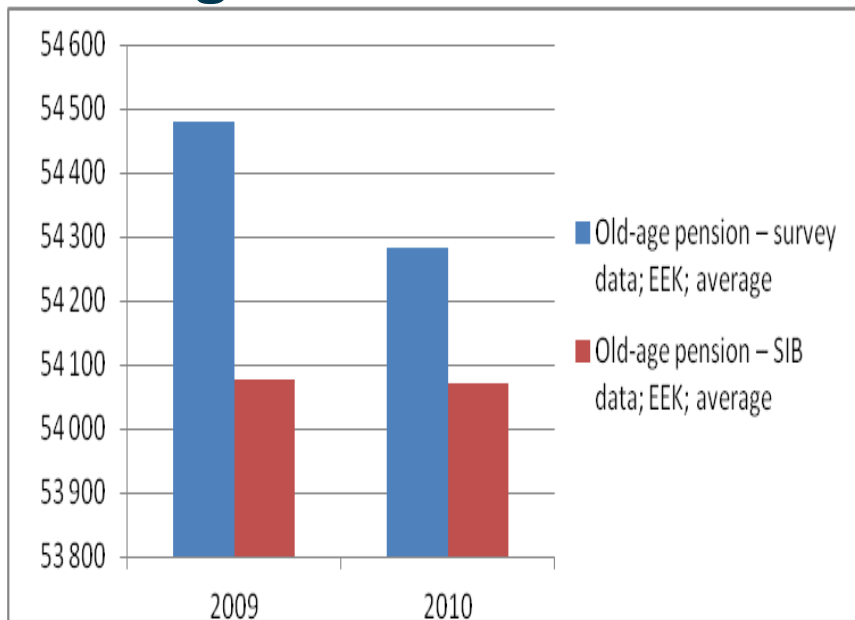
# THE MAIN RESULTS (SOME EXAMPLES)

# The old-pension recipients

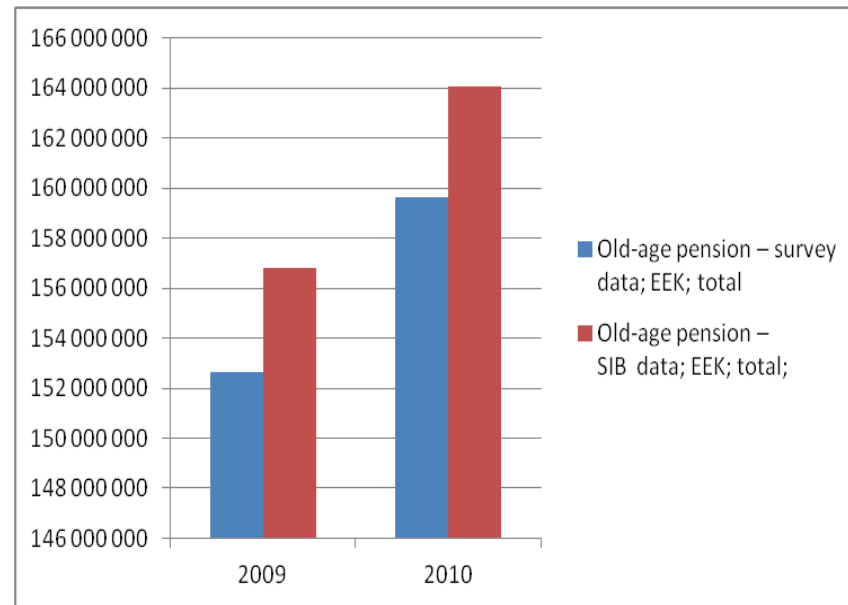


# Old –age pension

## Average



## Total sum



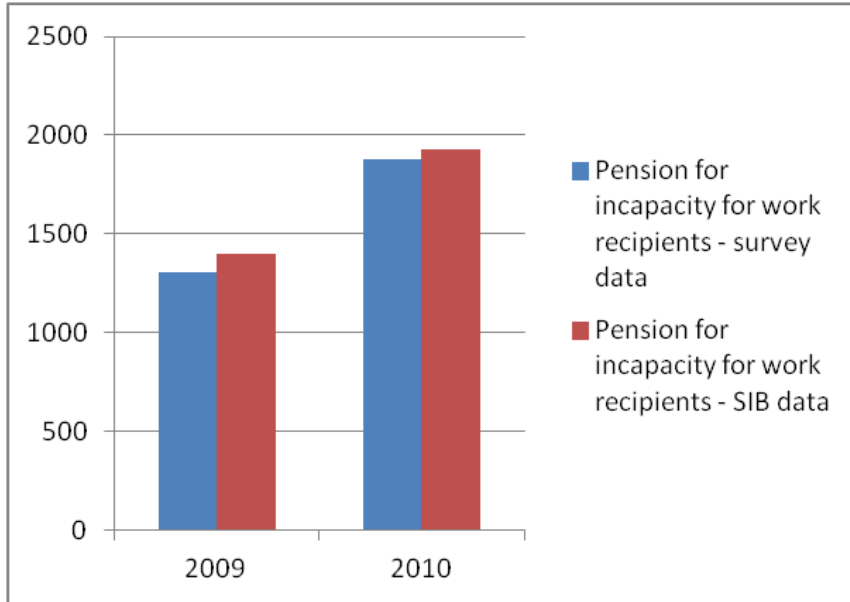
## Old-age pension

- Decision: To adopt the old-age pension data from the Estonian National Social Insurance Board; to expect the total number of pension recipients and a slight increase in the total sum; to expect a slight decrease in average values.

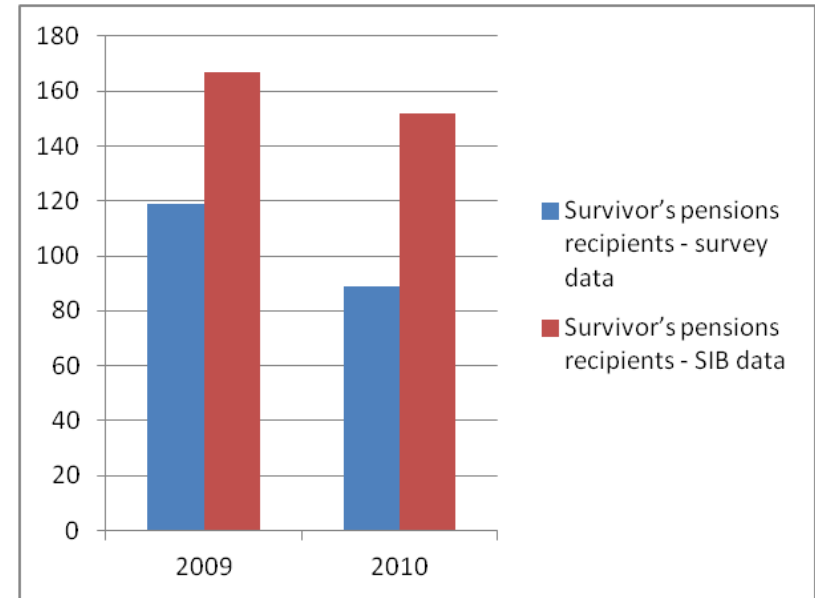


# Incapacity for work pension and survivor's pension

## Incapacity for work pensions recipients



## Survivor's pensions recipients



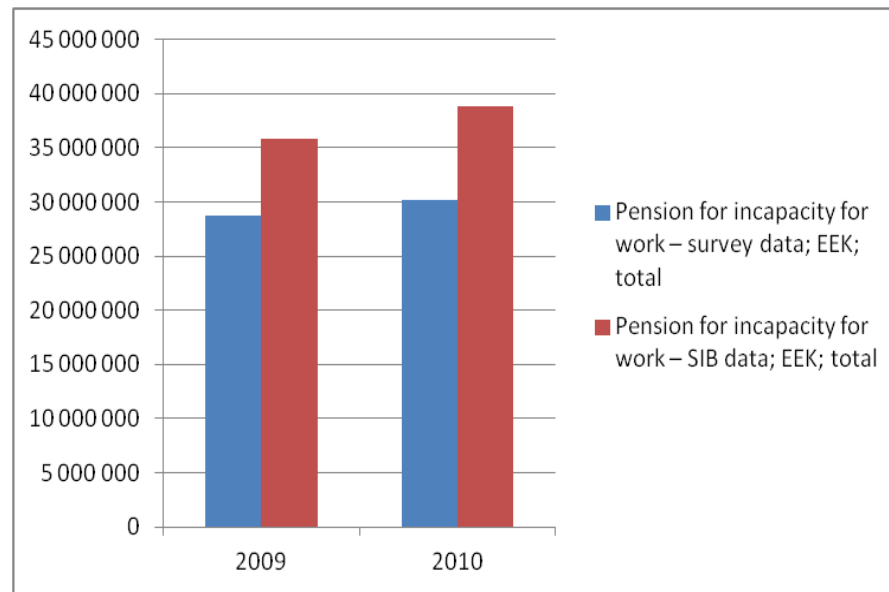


# Pension for incapacity for work

## Average



## Total sum

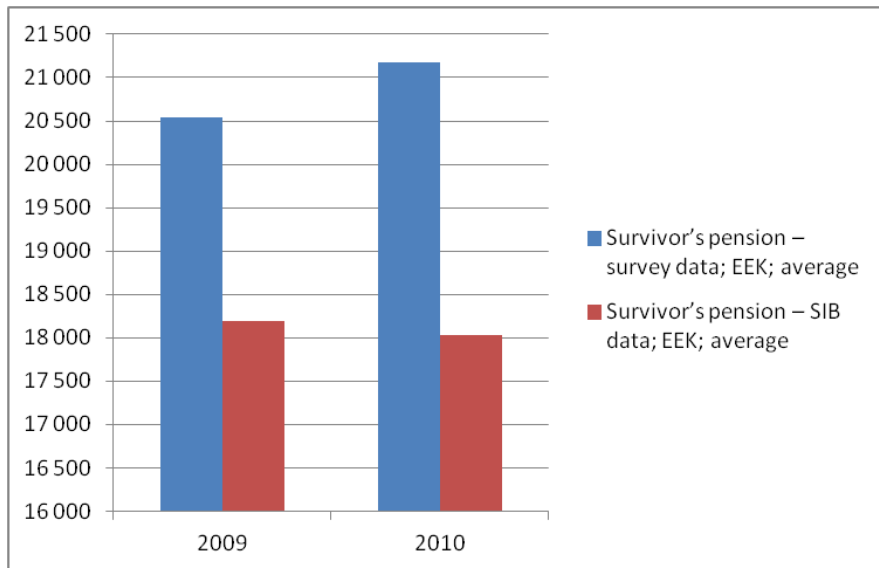




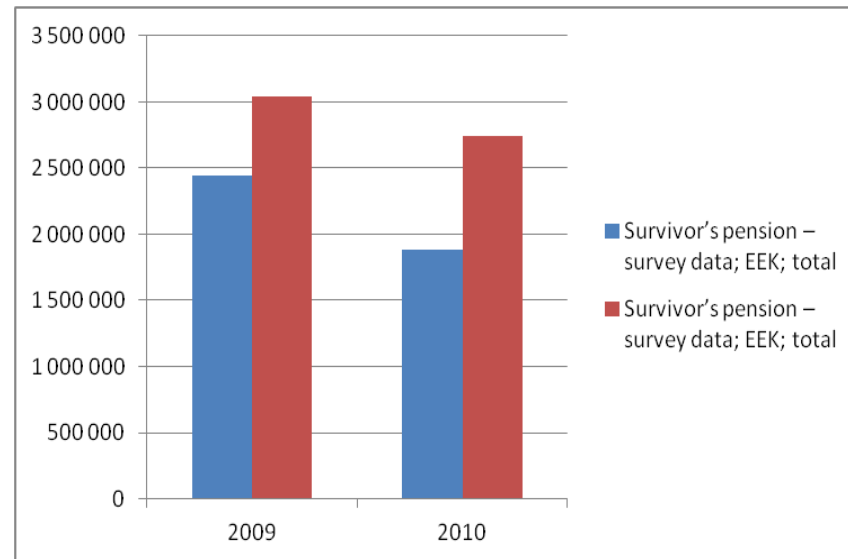


# Survivor's pension

## Average



## Total sum



## Pension for incapacity for work and survivor's pension

- The data from the Estonian National Social Insurance Board can also be used for both the pension for incapacity for work and the survivor's pension, but there are certain differences in the number of people receiving pensions and in average values.
- Decision: To use register data from the Estonian National Social Insurance Board in determining pensions for incapacity for work; to expect a slight decrease in the average values, higher total numbers of pension recipients and an increase in the total sums paid.

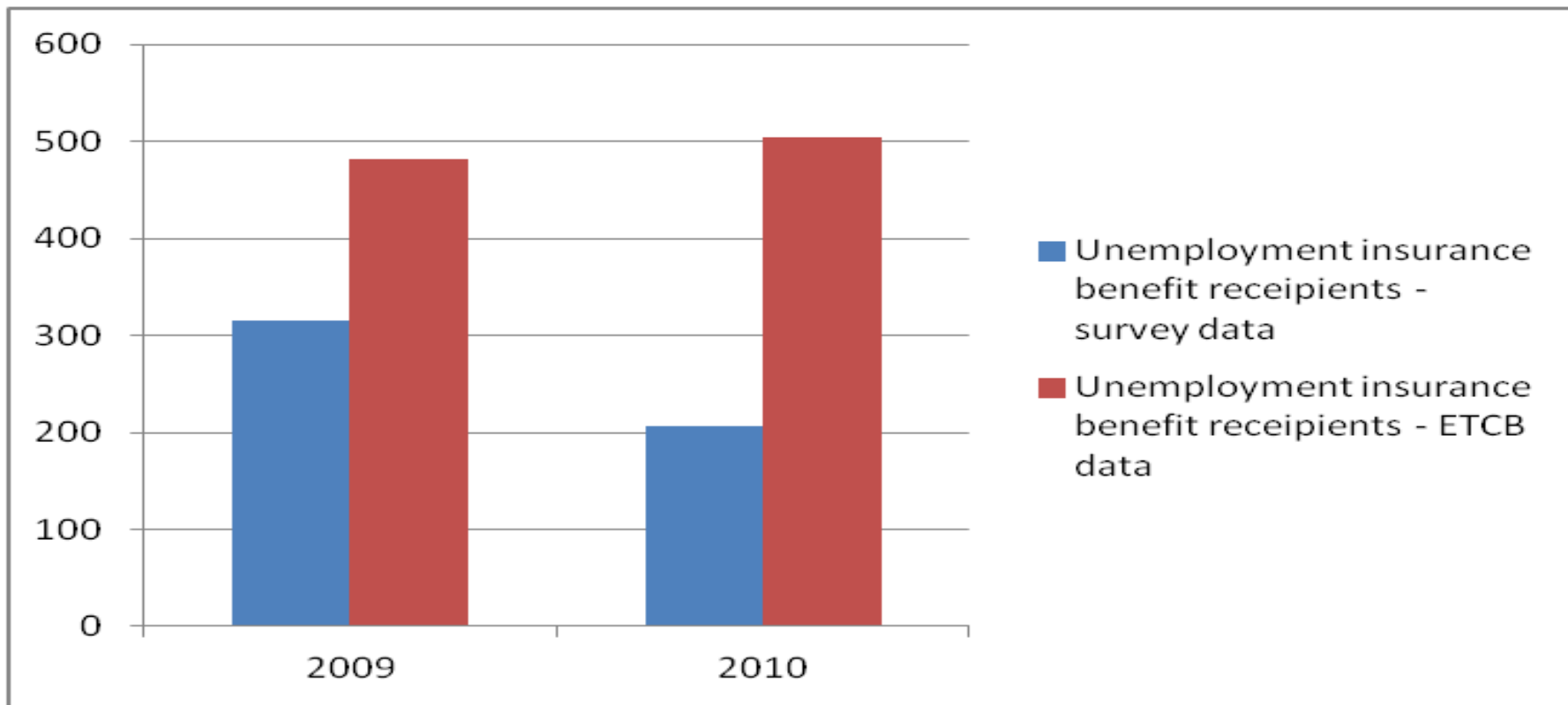


## Pension for incapacity for work and survivor's pension

- Decision: To replace the question about the survivor's pension in the personal questionnaire of the Social Survey with the register-based data from the Estonian National Social Insurance Board; to expect a slight increase in the number of pension recipients and a slight decrease in the average values



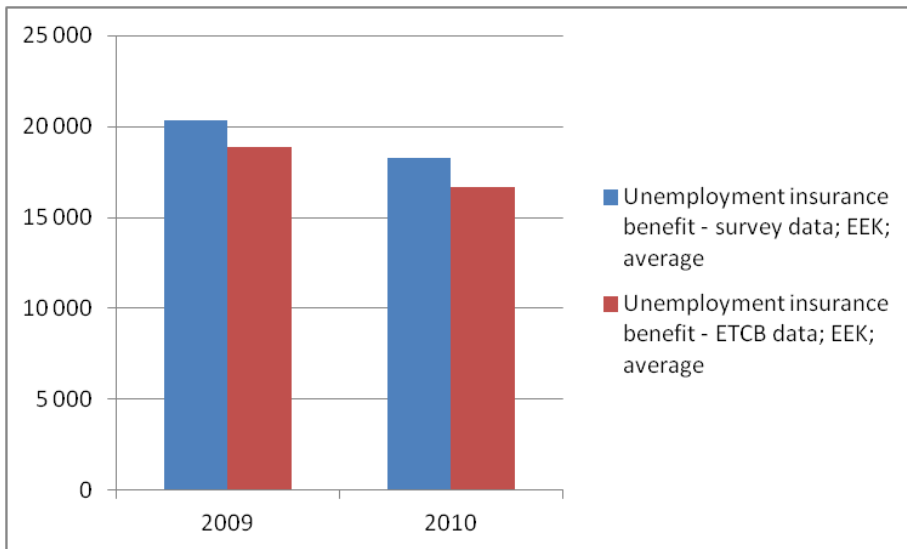
# Unemployment insurance benefit



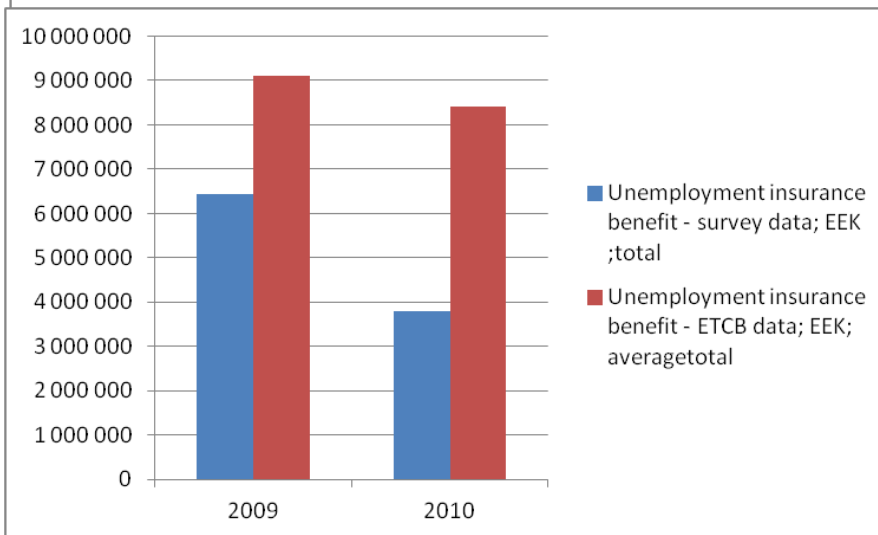


# Unemployment insurance benefit

## Average



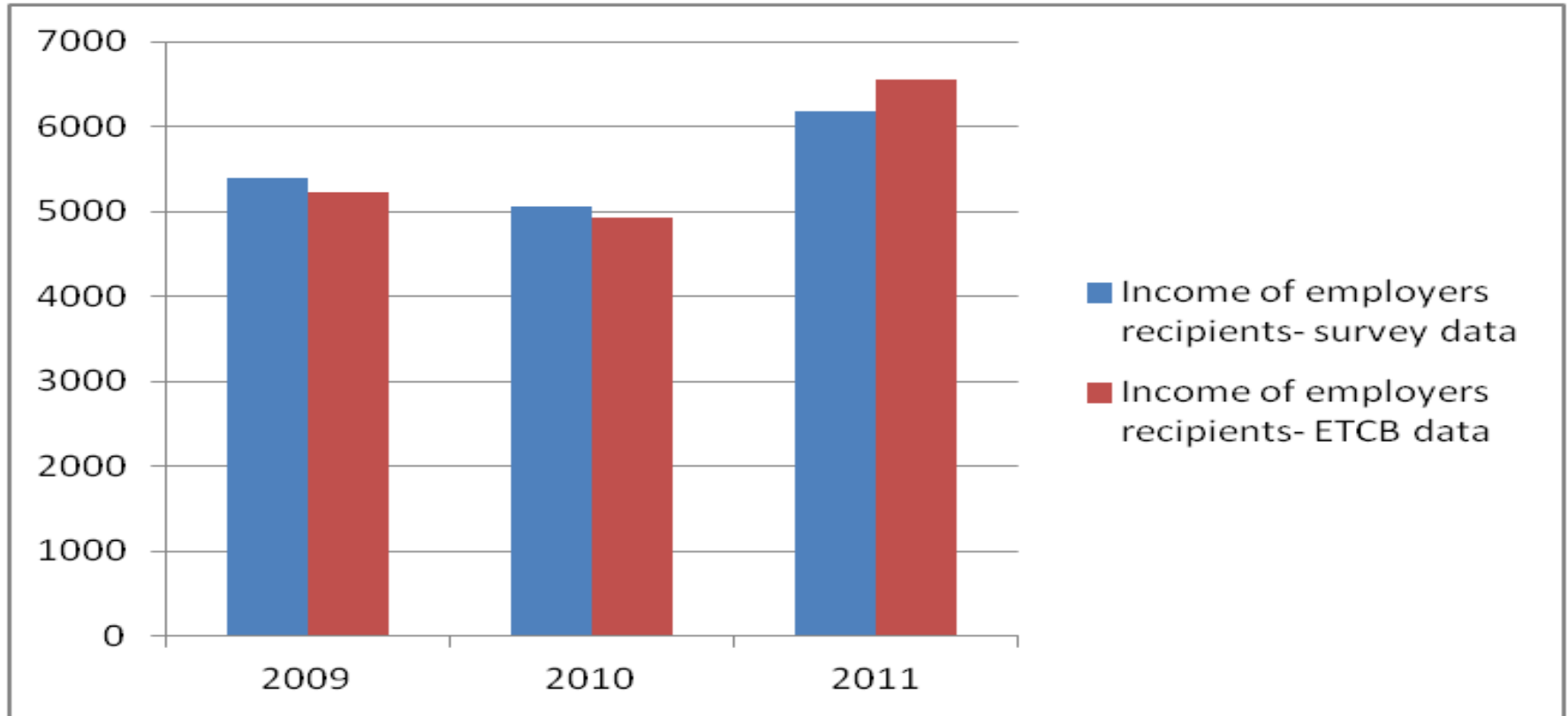
## Total sum



# Unemployment insurance benefit

- The question regarding unemployment insurance benefit can be replaced with data from the Estonian Tax and Customs Board (ETCB). Due to the tendency that people do not remember or indicate the relatively low sums of the received unemployment benefit, the register-based data from the Estonian Tax and Customs Board should be adopted in order to collect more accurate and reliable data. In that case, **the data will indicate a slight decrease in the average sums of unemployment benefit and a greater number of persons receiving the benefit.**

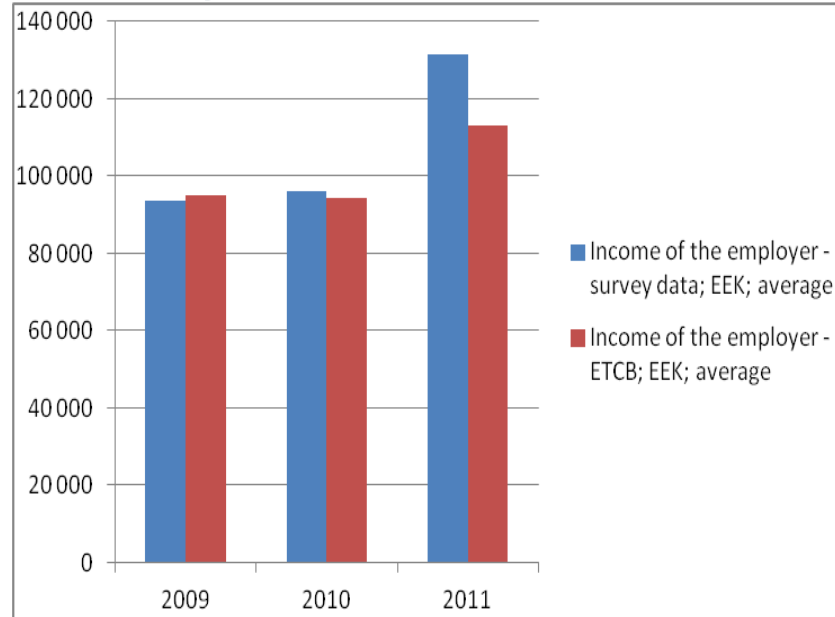
# Income of employers recipients



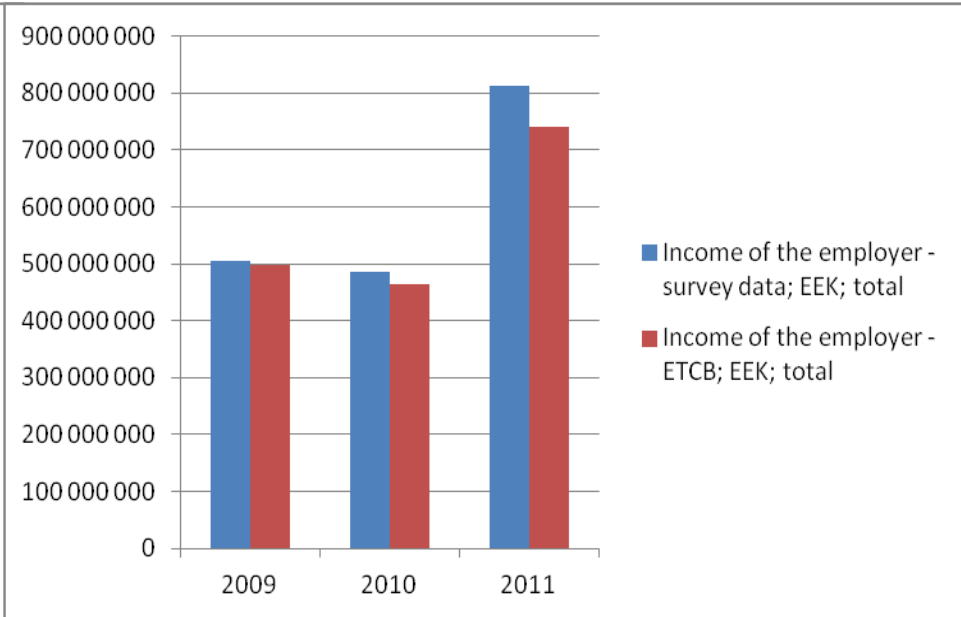


# Income of the employer

## Average



## Total sum





## The main results:

- Discrepancies in the data may be a result of "envelope pay" (it is presumed that, in the survey, a person will report also the part of income that was received as an "envelope pay") as well as answering a different understanding of the definition.

# The main results

- Decision: To keep the question regarding income from the employer in the ESS section regarding the income earned in 2013. At first, only missing values are to be replaced with data from the Estonian Tax and Customs Board. The transition is to be carried out in 2014 in order to avoid introducing great changes (benefits and income from the employer) in survey data all at once. In 2014, those income components from the employer are to be retained which cannot be obtained from registers (or which have not been analysed yet).

# The data of the Estonian National Social Insurance Board were highly compatible with the survey data in terms of household income:

## Types of family and child benefits and the potentially compatible data registers

<b>EU-SILC</b>	<b>Type of benefit</b>	<b>Register</b>
D11F	child benefit	SIB
D11D	pregnancy and maternity leave benefits	EHIF
D11H	childbirth allowance	SIB
D11B	parental benefit	SIB, ETCB
D11J	childcare allowance	SIB
D11K	conscript's child allowance	SIB
D11L	allowance for families with seven or more children	SIB
D11M	child benefit from local government	MISSING
D11N	adoption allowance	SIB
D11E	adoption benefit	SIB
D11O	foster care allowance	SIB
D11P	start in independent life allowance	SIB
DA3	state allowance	SIB

# The main results

- Among benefits related to family and children, the questions on most variables can be replaced with register data. Information regarding maternity benefits is compatible with the register of the Health Insurance Fund, parental benefit with the register data of the Estonian Tax and Customs Board and child benefits with the data of the Estonian National Social Insurance Board.
- A slight increase in average values and total sums is to be expected.



# The main results

- Questions regarding health-related benefits will not be replaced with register-based data from the Health Insurance Fund; in order to maintain the reliability and accuracy of the sickness benefit variable, the compatibility or combining of other databases with the dataset of the Health Insurance Fund should be analysed.

## The main results:

- In order to guarantee the validity of the analysis, for the moment, the question regarding redundancy payment should be retained in the questionnaire of the Estonian Social Survey (ESS). Next year, the ESS questionnaire should be re-examined for compatibility with either the data of the Estonian Tax and Customs Board or the Estonian Unemployment Insurance Fund

## The main results:

- In the 2015 survey on the income of 2014, there will still be some income components. For example: in terms of monetary income, travel expenses are asked about; the question about net earnings has also been retained (remained in control of question); additional questions are asked about non-monetary income from waged labour; questions regarding the following components were also retained: business income, property income, insurance, transfers from local governments, and also earnings, transfers and benefits received from a foreign country.



**Thank you!**



REPUBLIKA  
SRBIJA  
STATISTIKA