

# THE USE OF ADMINISTRATIVE DATA SOURCES FOR THE UPDATING OF THE POLISH STATISTICAL REGISTER

*CENTRAL STATISTICAL OFFICE OF POLAND*

Prepared by **Dorosława Kurek**,  
Deputy Director of Co-ordination of Statistical Surveys Division,  
in cooperation with:

- 1. Administrative Data Sources Section*
- 2. Sample Frames and Sample Surveys Section*
- 3. Integration of Statistical Information System Section*

Presented by **Agnieszka Prochot**,  
*specialist in Administrative Data Sources Section  
of Co-ordination of Statistical Surveys Division*

# ***THE NATIONAL OFFICIAL REGISTER OF THE NATIONAL ECONOMY UNITS (REGON)***

## **TASKS**

- to allocate identity numbers to the REGON register on the basis of the applications that have been submitted,
- to deliver certificates on the allocated REGON identity numbers,
- to run the REGON register as the computerised database,
- to provide public administration bodies and others interested parties with data from the register.

# ***THE BJS STATISTICAL REGISTER***

***sampling frame for all surveys conducted  
by the Central Statistical Office (GUS)  
on the population of Polish national  
economic units***

## **TASKS**

- to create frames for statistical surveys,
- to create lists for surveys,
- the description of units, which are subjects of statistical observation, for full and correct development of results

***ARTICLE 10 OF THE LAW  
ON OFFICIAL STATISTICS***  
**the legal basis for the creation  
of the statistical register**

The collected and gathered in the statistical surveys of official statistics individual and personal data shall be confidential and subject to particular protection; the data shall be used exclusively for statistical calculations, compilations and analyses and for the creation by the statistical services of official statistics sampling frames for statistical surveys conducted by those services; providing or use of individual and personal data for other than specified above purposes shall be prohibited (statistical confidentiality).

# A comparison between subjective scopes of REGON official register and BJS statistical register

## REGON

**Legal persons**  
**Organisational units without legal status**  
**Natural persons carrying out economic activity**

**Local units of the legal units mentioned above**

**Farmers, which are agricultural farms owners**

## BJS

**Legal units including natural persons carrying out economic activity (excluding agricultural activity)**

**Institutional units (legal units which are classified according to the algorithm)**  
**Enterprises (acc. to algorithm)**

**Local units**

**Local kind of activity units (LKAU) (in selected domains)**

***THE SUBJECT UNITS  
IN THE STATISTICAL REGISTER  
are described through:***

- **characteristics included in the REGON official register;**
- **additional characteristics based on data from surveys or administrative systems;**
- **additional classifications based on the characteristics mentioned before.**

# ***THE SOURCES OF INFORMATION FOR THE STATISTICAL REGISTER***

**The major source - the REGON register**

**Other sources:**

- **statistical surveys,**
- **the administration systems of the Ministry of Finance,**
- **other registers and lists,**
- **the official journal “Court and Commercial Gazette”,**
- **information from the units,**
- **postal and press information.**

***The MINISTRY OF FINANCE can provide to the CENTRAL STATISTICAL OFFICE the data from the tax system under the following conditions:***

- **there is a legal base for the use of these data for statistical purposes;**
- **uniform computing system (the best would be the central system) for tax data exists;**
- **the information scope of the BJS statistical register and tax system are consistent.**

**The Central Statistical Office (GUS)  
is provided with data from the National  
Tax Payer Registration Database (KEP)  
of the Ministry of Finance on the basis of:**

- *the 1999 Decree of the Cabinet on the way and methodology of running and updating the register of national economic units.*

**The other systems supply the Central Statistical Office with information on tax and cost on revenues, financial results and income tax (both for legal units and for natural persons carrying out business activity) on the basis of:**

- *the Decree of the Cabinet on the Programme of Statistical Surveys of Official Statistics.*

# 1. Comparison of the data sets of KEP database and BJS statistical register concerning the subject units

<b>Year</b>	<b>Total number of sample units</b>	<b>Number of sample units, which were not found in KEP database</b>
<b>2000</b>	<b>114 000</b>	<b>44 000</b>
<b>2001</b>	<b>114 630</b>	<b>9 597</b>

**Year 2002**

<b>Specification</b>		<b>Number of units</b>	
		<b>total</b>	<b>consistent</b>
<b>Legal persons Organisational units without legal status</b>	<b>KEP</b>	<b>642 486</b>	<b>531 514</b>
	<b>REGON</b>	<b>744 032</b>	
<b>Natural persons carrying out business activity</b>	<b>KEP</b>	<b>3 097 478</b>	<b>2 301 673</b>
	<b>REGON</b>	<b>2 682 623</b>	

## ***2. Equipping BJS register with the NIP numbers, which are identity numbers for tax system***

Year 2002

<b>Specification</b>	<b>Number of consistent units</b>	<b>% of units in BJS equipped with NIP identity number</b>
<b>Legal persons Organisational units without legal status</b>	<b>531 514</b>	<b>71,4%</b>
<b>Natural persons carrying out business activity</b>	<b>2 301 673</b>	<b>85,8%</b>

### ***3. Explaining differences in subject units between KEP and BJS records.***

**Some reasons of deficiencies in links between the REGON and NIP identify numbers:**

- units which have registered themselves in the REGON register (including natural persons carrying out economic activity) but which haven't started their activity,
- errors which occur in entries of the REGON and PESEL identity numbers,
- units which have been moved out from the REGON register but which still submit tax returns,
- cases where the personal company has been registered in the KEP database while its partners have been registered as natural persons carrying out economic activity in the REGON register.

## ***4. Preparations of adequate information (notes) in the Programme of Statistical Surveys of Official Statistics.***

The information on the use of the particular administrative source, which is located in the Programme, guarantees that the official statistics have a right to use this source for statistical purposes.

In the Programme of Statistical Surveys 2002, an obligation to provide the data from Ministry of Finance tax system for Central Statistical Office was introduced.

In 2002, the Central Statistical Office was provided with data:

- from 1999 and 2000 for two selected voivodships in order to conduct methodological work,
- from 2001 for the whole country.

***5. Comparative analyses of revenues in the tax system and statistical surveys together with a determination of reasons and range of differences.***

**The data from the tax system are derived from two databases:**

- **personal income tax database (PIT) - contains information taken from personal income tax return forms PIT,**
- **corporate income tax database (CIT) - contains information taken from corporate income tax return forms CIT.**

## **5. Comparative analyses of revenues in the tax system and statistical surveys together with a determination of reasons and range of differences (cont.)**

<b>The kind of tax return form</b>	<b>The number of units</b>		<b>% of units consistent with BJS</b>
	<b>Total in particular files</b>	<b>consistent with reference to BJS</b>	
<b>CIT</b>	<b>225 095</b>	<b>217 498</b>	<b>96,6%</b>
<b>PIT</b>	<b>1 258 231</b>	<b>951 345</b>	<b>75,6%</b>
	<b>among this, units, which are in KEP database marked as natural persons carrying out business activity on general rules</b>		
	<b>1 022 855</b>		<b>93,0%</b>

## **5. Comparative analyses of revenues in the tax system and statistical surveys together with a determination of reasons and range of differences (cont.)**

A comparative analysis of revenues of units concerned the enterprises which have the number of employees <10 and submitted reports in 2001

<b>% of the analysed units</b>	<b>The value of relative error</b>
<b>16</b>	<b>&gt; 10%</b>
<b>79</b>	<b>&lt; 3%</b>
<b>0,3</b>	<b>= 100 %</b>

In most cases, the analysis of distribution of relative error proves a very high quality of information on the revenues, taken from two accessible sources.

# ***The analysis of the activity of units from the statistical register***

**In order to define non-activity, the following rules were established:**

- **non-active units, for which a final cessation of activity is marked in the KEP database;**
- **potentially non-active units, which haven't been found in the KEP database and additionally which fulfil the following criteria:**
  - **the date of starting activity < 1997;**
  - **the date of any updating in the BJS statistical register < 2000;**
  - **the units didn't submit the reports in 2000 and 2001.**

# ***The REGON register - very important for quality of Polish official statistics surveys***

- it meets considerable part of statistics needs,
- the beginning of activity in the case of each legal unit follows a registration of this unit in REGON register,
- it ensures completeness of legal units in statistical register.

# ***The best way to improve the quality of surveys***

- to obtain the data from administrative sources, especially from the Ministry of Finance tax databases.

**The data sets obtained from the Ministry of Finance in 2002 made possible:**

- the updating of the economic activity of 141911 units (4,2% of the active units) in the statistical register;
- a change of rules of sample selection for the survey of enterprises;
- the beginning of work on the completion of non-responses in surveys, which are conducted on the basis of data from the tax system.

# ***Advantages of obtaining up-to-date information on units from the administrative sources***

- the elimination of overcoverage in the statistical register,
- the possibility of easier creation of unit stratification by size and better precision of surveys,
- in the case of non-responses – possibility of completion.

**For these reasons, the use of administrative data for updating data from statistical registers is a priority for the statistics.**