United Nations

Economic and Social Council

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Population and Housing Censuses

Seventeenth Meeting
Geneva, 30 September to 2 October 2015
Item 6 of the provisional agenda
Expected issues in complying with the new CES Recommendations
(particularly by countries using registers or surveys) and results of work to address them

Deriving labour force characteristics from multiple sources in the Register-based Census of Slovenia

Note by the Statistical Office of the Republic of Slovenia[1]

Summary

The paper focuses on the labour force status in the census methodology in Slovenia where data are derived from nine administrative or statistical sources. The same methodology is now used also for the annual production of the economic characteristics of the whole population (full coverage) in demography statistics (Population Database). Besides the description of the differences of the concepts between the new CES recommendations and the register-based approach, the quantitative evaluation of the quality of input data is evaluated with the conclusion that census statistics on the economic characteristics is produced from trustworthy sources.

The register-based data on the labour force status were linked with the Labour Force Survey micro-data on the individual level using the unique identifier (PIN) to find out the extent of coherence between the field survey data and the data in registers. The results are encouraging, as 90% of the records were classified into the same category in both surveys, taking into account the main labour force status from the LFS (87% in case of the proposed census methodology from the new Recommendations for the 2020 Censuses of Population and Housing).

---

[1] Prepared by Danilo Dolenc.

I.   Introduction

1.   Producing of statistics on labour force characteristics from the administrative records has certain limitations (paragraph 488, Draft Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing (version 4 March 2015), hereinafter Draft Recommendations). The aim of the paper is to present the coherence, obstacles, differences between the new CES recommendations (in general more appropriate for traditional field enumeration) and the register-based approach that - in case of Slovenia - was focused predominantly on the labour force status. Namely, the labour force status is the essential economic characteristic having an implication on many other census topics.

2.   The aim of this paper is also to study the quality of the register-based outputs prepared according to the census methodology described below with the Labour Force Survey (hereinafter LFS) methodology (harmonized with concepts presented in chart 4: Classification of working age population by labour force status) of the Draft Recommendations on the individual base using the unique personal identifiers (PIN, personal identification number).

II.   The methodology of the register-based labour force characteristics in Slovenia

3.   In general the methodology follows the main principles of classification of persons by their labour force status giving priority to employment over unemployment and outside the labour force categories. The Statistical Office of the Republic of Slovenia uses nine sources of data on the labour force status using the source hierarchy methodology. We examined each provided source of the data on labour status in advance and assessed their suitability for use in terms of quality, reliability, timeliness, accessibility and comparability. The basic methodological principle is the hierarchy of the sources, which means that in case of availability of data on labour force status in several sources, the priority was given to the source indicated with higher priority (irrespective of whether data in different sources classified a person into the same category or not). In the next step, in case of sources which provided different labour force statuses (case of Health Insurance Institute data), the general priority rule according to the recommendations was applied.

4.   Annex provides information on the hierarchy of sources, the content of sources and the reference period of data used for register-based labour force characteristics. Eleven categories on the labour force status are used in our internal standardized classification. From 2011 onwards when the statistical process was established in the scope of the first register-based census in Slovenia, we derived the labour force status annually, using same methodology. The last column of Annex presents the latest data available (1 January 2014) as data on economic characteristics are produced in our Population Database approximately 11 months after the reference date.

III.   Comparing general concepts

5.   Paragraph 494 of the Draft Recommendations identifies five forms of work (own-use production; employment for pay or profit; unpaid trainee work; volunteer work; other work activities) where measurement of persons in employment is essential in census taking. Data for such measurement are available in administrative and other sources used in Slovenia. Other forms of work (with the exception of own-use production in less developed countries) should be the object of specialized household surveys.

6.   According to the labour force status presented in paragraph 503 of the Draft Recommendations, the classification of population is harmonized with register-based outputs with the exception of category 2.3 Performing unpaid household services which is less and less relevant in Slovenia. Nevertheless, also paragraph 504 of Draft Recommendations states that this category should be joined with category 2.4 Others.

7.   At first sight, the essential differences between the register-based method and field enumeration when collecting data on labour force status according to the Draft Recommendations are:

• Short reference period to determine the labour force status (1 week);

• Work at least 1 hour for pay/profit in the reference period irrespective of self-declared main labour status as defined in the LFS methodology;

• Having a paid job/business from which the person was temporarily absent;

• Fulfilment of two additional conditions to be classified as unemployed.

8.   But in fact the only problematic methodological issue is the one hour criterion which could be applied only in field enumeration. The other three components can to a great extent be possibly achieved also in our register-based system.

9.   The main source for determining persons in employment (Statistical Register of Employment) consists in fact of all persons having a paid job or running their own business (with or without employees) as self-employed or farmers. All reasons for absence (paragraphs 510 and 511 of the Draft Recommendations) could apply for registered employed persons without knowing the reason which is in accordance with paragraph 512. Besides that, the employment for any limited period of time can be extracted from the Statistical Register of Employment. So the reference period to determine the employed labour status is the last 7 days before the critical date (1 January) which is very important due to the fact that many employed persons usually retire in the last days before the New Year (mostly because of the annual changes of legislation offering worse conditions in view of the amount of the pensions).

10.   Some of the specific groups of the employed persons listed in paragraph 513 of the Draft Recommendations cannot be detected with the register-based approach. This refers to all groups receiving pay in cash or in kind included in points (c) - Apprentices, interns and trainees; (d) - Participants in job training schemes or (e) - Having a seasonal job. Knowing our situation, categories (c), (d) and (e) are not numerous and statistically speaking are in the census context also unimportant. Categories (a) - Employed persons on training and (g) – Armed force members are included in the Statistical Register of Employment. Contributing family workers (b) are derived in the statistical process from the sources provided by the Health Insurance Institute of Slovenia for households with at least one registered farmer, i.e. from the Statistical Register of Employment; so their number is underestimated in the register-based census.

11.   The main difference in the context of defining the employed persons between field enumeration and the register-based approach is complying with the category (f) as stated: ''In accordance with the priority rules of the labour force framework, persons who, during the reference period, were mainly students, homemakers, pensioners, registered unemployed, etc., who at the same time were in employment as defined above, should be considered employed on the same basis as other categories of the employed persons.'' Simultaneously the time criterion of one hour work for pay/profit is a crucial methodological presumption included here.

12.   All three conditions for determining the unemployed persons from paragraph 516 of the Draft Recommendations (not in employment, seek employment, currently available) are incorporated also in the legislation regarding the right to be registered as an unemployed person, irrespective of whether this person is receiving social

transfers or not. We can say that the methodology of the Employment Service of Slovenia is harmonized with the stated paragraph 516. As for the employed persons, also for unemployed persons not registered at the responsible authority we cannot fulfill the recommendations presented in paragraph 518 of the Draft Recommendations with a quotation similar to paragraph 513 (f).

IV. The quality of the register-based data sources on the labour force status in Slovenia

13.   Data sources are divided into 3 categories according to the quality assessment (5 sources are indicated as being of high quality, 2 sources as low quality but still suitable, and 2 sources as outdated, i.e. one high and one low quality), see also Annex. Table 1 also includes imputations as the last source of data, as we decided to impute the labour force status for the whole working age population, not following the Finnish example (3) where   3.2% of the population did not meet any of the previous criteria and was categorized as ''Others outside the labour force''. In Slovenia, only 1.3% of the records for the working age population were not found in any source.

Table 1
Register-based labour force status by quality of the source, Slovenia, 1 January 2014

Register-based labour force status	*Total*	*Type of quality of source (% of labour force status)*			
		High	*Low*	*Outdated*	*Imputation*
Working age population	**100**	**85.5**	**12.8**	**0.4**	**1.3**
Labour force (1.0)	*100*	*97.0*	*1.2*	*0.6*	*1.2*
Employed (1.1)	100	97.1	1.0	0.7	1.2
Unemployed (1.2)	100	96.2	2.7	0.0	1.1
Outside labour force (2.0)	*100*	*72.7*	*25.7*	*0.2*	*1.4*
Schooling (2.1)	100	53.1	45.5	0.3	1.1
Pension recipients (2.2)	100	95.4	4.0	0.1	0.5
Others (2.4)	100	2.0	91.5	0.7	5.8

Source: Statistical Office of the Republic Of Slovenia, internal documentation.

14.   The main outcomes revealed that most records on the labour force population and pension recipients exist in the most trustworthy sources. The main problem regarding the schooling population is the unavailability of data on secondary school enrolment (but this is to be solved in near future as a new source for monitoring all participants in educational institutions from kindergarten till 3rd Bologna cycle is now being established). Not surprisingly, most of the records for other persons outside the labour force are derived from the compulsory health insurance data.

15.   In the next step we also analyzed the number of sources for each record and the adequacy of the relation between the sources in case of two or more records being found in different data sources for the same person. The results are given in the following tables. Namely, almost half of the records were available from one source only (of which every third out of four defines an employed person from the Statistical Register of Employment). On the other hand, 95% of data on pensioners were available in two sources, of which 96% in the Pension and Disability Insurance Institute data and the Health Insurance Institute data (Persons in health insurance under specific schemes). Data on the unemployed persons and the persons attending an educational institution existed in numerous sources.

Table 2
Register-based labour force status by number of data sources, Slovenia, 1 January 2014[2]

Register-based labour force status	Total	Number of sources		
		1	*2*	*3+*
Working age population	**1,729,427**	**775,631**	**890,318**	**63,478**
Employed (1.1)	784,611	562,361	213,973	8,277
Unemployed (1.2)	124,486	18,306	93,017	13,163
Schooling (2.1)	170,327	65,608	62,833	41,886
Pension recipients (2.2)	533,578	25,480	507,959	139
Others (2.4)	116,425	103,876	12,536	13

Source: Statistical Office of the Republic Of Slovenia, internal documentation.

16.   Finally, the combinations of the labour force status and the data sources in case of at least a duplicate record for the same person were divided into three major groups, based on 30 detailed categories:

- All sources indicated the same labour force status: 75% of cases (type 1 in table 3);

- Sources indicated at least two different labour statuses but could be acceptable, as persons may be engaged in multiple activities at the same time (e. g. working and studying; part-time employed and part-time retired, etc.): 5% of cases (type 2 in table 3);

- Sources indicated at least two different labour statuses which could be treated as errors (e. g. employed and unemployed): 20% of cases (type 3 in table 3).

17.   Taking into account the hierarchy of sources as the primary methodological solution combined also by some other variables (age, type of income, base for health insurance, type of household) and the determined labour force status in case of multiple records, we can conclude that the quality of the register-based statistics is without any doubt of high quality.

18.   An evidence of the quality of input data is also the number of duplicate records in the two most excellent sources (Statistical Register of Employment and Registered unemployed persons database) which was only 1,867 or 1.5% of all records from unemployment database (0.24% of all records from the employment register).

---

[2] High and low quality sources only.

Table 3
Derived register-based labour force status (LFS) in case of multiple records from at least two sources, Slovenia, 1 January 2014[3]

Ty-pe	Description	Total	Labour force status				
			Emp-loyed	Unem-ployed	Schoo-ling	Retired	Others
	TOTAL	**953,796**	**222,250**	**106,180**	**104,719**	**508,098**	**12,549**
1	Same LFS in 2+ sources - employed	82,641	82,641	-[4]	-	-	-
1	Same LFS in 2+ sources - unemployed	13,762	-	13,762	-	-	-
1	Same LFS in 2+ sources - schooling	86,551	-	-	86,532	-	19
1	Same LFS in 2+ sources - retired	488,751	-	-	-	488,751	-
1	Same LFS in 2+ sources - others	10,617	262	-	11	45	10,299
1	Same LFS in 2+ sources - outside labour force	33,669	-	-	17,098	16,264	307
2	Different LFS in 2+ sources - employed + schooling	17,421	17,153	-	268	-	-
2	Different LFS in 2+ sources - employed + retired	6,053	5,907	-	-	146	-
2	Different LFS in 2+ sources - employed + social transfers recipient	9,966	9,923	-	-	-	43
2	Different LFS in 2+ sources - unemployed + schooling	2,272	-	2,233	39	-	-
2	Different LFS in 2+ sources - unemployed + retired	1,569	-	1,496	-	73	-
2	Different LFS in 2+ sources - unemployed + social transfers recipient	5,670	-	5,655	-	-	15
2	Different LFS in 2+ sources - schooling + social transfers recipient	1,834	-	-	597	-	1,237
2	Different LFS in 2+ sources - retired + social transfers recipient	2,448	15	-	-	1,808	625
3	Different LFS in 2+ sources - employed + unemployed	26,063	21,462	4,601	-	-	-
3	Different LFS in 2+ sources - employed + other inactive person	83,967	83,798	-	162	3	4
3	Different LFS in 2+ sources - unemployed + other inactive person	77,415	-	76,395	12	1,008	-
3	Different LFS in 2+ sources - employed + unemployed + other inactive person	3,127	1,089	2,038	-	-	-

Source: Statistical Office of the Republic Of Slovenia, internal documentation.

---

[3] High and low quality sources only.
[4] No occurrence of event.

V. Linking of the register-based data on the labour force status with LFS data

19.   When analyzing data we have to take into account that the register-based labour force status refers to the reference date (1 January) while the LFS data are collected quarterly. For the purpose of the paper, the LFS data for the nearest quarters (fourth quarter of 2013 and first quarter of 2014) were linked with the data on the labour force status from the Population Database using the census methodology as of 1 January 2014 and as of 1 January 2013. Some methodological and technical particularities had to be applied in the preparatory phase of constructing the final analytics database:

- Only the working age population was the target.

- The same person could be interviewed in the LFS in both quarters according to the panel type of survey - in case of duplicates the record for Q1/2014 was obtained.

- The same person (PIN from LFS) had to belong to the population of Slovenia as of 1 January 2013 and 1 January 2014.

- PIN's for the household members (except for the sampled person) were derived automatically. The statistical procedure used names, surnames, addresses, sex and dates of birth collected in the field to determine the PIN with certain probability.

Table 4
Analytics database preparation

Description of operation	Deleted records	Nr. of records
Initial LFS database Q4/2013 and Q1/2014	-	31,379
1. Excluded records for those younger than 15 years on 1 January 2014	2,851	28,528
2. Excluded duplicate LFS records	8,763	19,792
3. Excluded unlinked records with the population database	177	19,615
4. Excluded linked records with the lowest probability	410	19,205

Source: Statistical Office of the Republic Of Slovenia, internal documentation.

20.   The final analytics database consisted of 19,205 records which represented 1.1% of the working age population as of 1 January 2014. As expected, the unweighted structure of the working age population in the analytics database over-estimated the retired population (in general they were more likely to want to participate in the field surveys), the weighted structure of the LFS working population underestimated the students (an important share of students living at the term-time address in the place of study were excluded from the sample), while the shares of the employed and unemployed persons were quite similar if we took into account separately both surveys, and in fact the differences between the surveys were above all the result of the methodology applied.

Table 5
Structure of working age population (%), LFS, Register-based statistics and Analytics database (different concepts of labour force statuses)

	Total	Labour force status				
		Emp-loyed	Unem-ployed	Schoo-ling	Re-tired	Ot-hers
LFS Q1/2014 - weighted	100	50.9	6.2	8.4	28.4	6.1
Register-based statistics, 1 January 2014	100	45.5	7.2	9.8	30.5	7.1
Analytics database (unweighted) - LFS						
Main labour status (self-declared)	100	45.7	9.8	11.0	31.0	2.6
Labour status during the reference week	100	49.7	9.3	9.3	29.4	2.4
Labour status during the reference week one year before	100	49.8	8.7	10.2	29.0	2.3
Analytics database (unweighted) - register						
Register-based labour force status	100	44.0	7.1	10.8	32.1	5.9
Register-based labour force status one year before	100	44.6	6.6	11.6	31.1	6.1

Source: Statistical Office of the Republic Of Slovenia, internal documentation.

21. The most similar concept, classification and consequently the most equivalent data should be obtained when comparing the register-based labour force status with the main labour status from the LFS which is based on the self-declared labour status (see also Eurostat's document Implementing core variables in the EU social surveys). The main difference was observed in the status ''Fulfilling domestic tasks'' from the LFS. Results were even better than expected: 90% of persons had the same labour force status in both surveys (excluding the LFS status homemakers). There was also a significant difference in terms of quality of register-based sources – for high quality sources the full matching reached 95.4% (the share of records was 87.3%), but for the low quality sources the full matching was only 54.5% (for 12.2% of records). The number of labour force statuses derived from outdated (57) and imputed (39) records was too small for reasonable analyses.

22. Age is quite an important factor generating differences between the field results and the register sources. For the age group 25-34 years, the exact match for the main labour status was the lowest (84.1%), while for the persons older than 64 years it was understandable that it was the highest (94.4%).

Table 6
Main labour status from LFS (Q4/2013 or Q1/2014) and register-based labour force status (1 January 2014) using census classification

LFS main labour status	Total	Register-based labour force status				
		Employed	Unem-ployed	Schoo-ling	Retired	Others
Total	**19,205**	**8,446**	**1,365**	**2,083**	**6,171**	**1,140**
Employed	8,776	8,170	171	51	162	222
Unemployed	1,878	184	1,087	49	22	536
Schooling	2,104	22	28	1,976	2	76
Pension recipients	5,944	33	26	2	5,837	46
Homemakers	302	28	16	0	98	160
Others	201	9	37	5	50	100

Source: Statistical Office of the Republic Of Slovenia, internal documentation.

23.   A slightly higher share of employed persons was observed in the field survey. More persons declared in the LFS that they considered themselves to be unemployed, but they mostly did not fulfill the conditions or they did not want to be registered at all. Most of them were classified as other inactive person in the registers. Very few persons still declared themselves as homemaker in the LFS, 85% of them were outside the labour force, so this group is irrelevant for the labour market in Slovenia.

24.   When using the ILO definitions in the LFS, the data asking about the labour status during the reference week (working at least one hour for pay/profit) should be less comparable. Anyway, 87% of the working age population were classified the same in both surveys. The share of the employed persons presented in table 5 increased in that case from 45.7% to 49.7% (compared to 44.0% which was the register-based result). The main contributors to the labour force were employed students working via a special scheme (the student employment office is a favorable option for both employers and students, due to fewer taxes) and the retired persons working at their own farm or as contributing family workers. It seems that most of the self-declared unemployed persons did not work during the reference week. Yet, the fewest persons that were recruited as employed were the others inactive (42), followed by the unemployed (98), the retired (305) and the students (318) of whom almost all (312) indicated they did irregular short-term work via the student employment office. Most of the employed persons were engaged on a legal basis (having a work contract, being registered as own-account workers or farmers), less than 4% of the participants in the LFS survey considered themselves as employed even if their legal status could be classified as informal (almost three quarters of them were contributing family workers). Only one third of them were also classified as employed in the register-based statistics.

Table 7
Labour status during the reference week from LFS (Q4/2013 or Q1/2014) and register-based labour force status (1 January 2014) using the census classification

LFS labour status during the reference week	Total	Register-based labour force status				
		Employed	Unem-ployed	Schoo-ling	Retired	Others
Total	**19,205**	**8,446**	**1,365**	**2,083**	**6,171**	**1,140**
Employed	9,539	8,205	234	337	465	298
Unemployed	1,780	171	1,033	48	21	507
Schooling	1,786	14	21	1,692	2	57
Pension recipients	5,639	28	25	2	5,542	42
Homemakers	267	20	15	0	93	139
Others	194	8	37	4	48	97

Source: Statistical Office of the Republic Of Slovenia, internal documentation.

VI.   Conclusion

25.   It is evident that full compliance with the new Recommendations for the 2020 Censuses of Population and Housing on the labour force status can be achieved only in the field survey, similar to the Labour Force Survey, which is in the EU Member States conducted quarterly and gives data on the current situation on the labour market. On the other hand, censuses are in most UNECE countries conducted only every tenth year and processing of data could last for a longer time, so there is a certain delay in view of the dissemination. According to the EU Regulation on 2011 Census, the deadline for transmitting the data to Eurostat via the Census Hub was March 2014, the formal opening of the new tool was on 8 December 2014. In case

of Slovenia that meant almost 4 years after the reference date. In the meantime several news releases with LFS data were released.

26.   Differences in the structure of the working age population according to the labour force status between register-based approaches and field surveys are therefore anticipated. In case of Slovenia, where the census methodology data and LFS data were linked for the first time, the results proved to be much better that foreseen before having performed the task. We can even explain the reasons for most of the differences presented in tables 6 and 7.

27.   Besides that we have once again thought about dealing with persons for whom no records on the labour force status can be found in any of the sources (1.3% of the working age population as of 1 January 2014). An important issue is that we are now also aware of the imperfection of some of the sources and also that there is still room for both the improvement of the quality of the administrative sources and also modification of the statistical processes.

References

The CES recommendations for the 2020 round of censuses - version for CES consultation, 63rd plenary session of the Conference of European Statisticians, download on 7 July 2015 from http://www.unece.org/index.php?id=38920#/

Prioritization of information from multiple register sources - case of main type of activity in the Finnish census, WP.12, Joint UNECE/Eurostat Expert Group Meeting on Register-Based Censuses, The Hague 2010

Use of Survey data (LFS) to evaluate the quality of register-based census in Finland (Finland), WP.11, Expert Group Meeting on Censuses Using Registers, Geneva 2012

Annex

Hierarchy of the sources used in the register-based statistical process on the labour force status, Slovenia

Prio-rity	Source	Qua-lity[5]	Content	Reference period	Labour force status[6]	Nr. of records[7]
1	Statistical Register of Employment run by the Statistical Office	HQ	Persons in paid employment Self-employed persons Farmers included in compulsory social security Detached workers Cross-border workers	Last seven days before reference date (1 January)	Employed Self employed Farmer	701,706 62,533 12,846
2	Employment Service of Slovenia	HQ	Registered unemployed persons	1 January	Unemployed, never worked before Unemployed, previously in employment	22,419 98,685
3	Statistical survey on student enrolment in tertiary education (full coverage)	HQ	Full-time and part-time students in vocational and professional higher education	Current academic year	Student	73,908
4	Statistical survey on awarding of scholarships (almost complete coverage)	HQ	Recipients of scholarships in upper secondary and tertiary education	1 January	Pupil Student	12,766 1,247
5	Pension and Disability Insurance Institute	HQ	Recipients of old-age pensions Recipients of disability pensions Recipients of survivor's pensions	1 January	Pupil Student Pension recipient Social transfers recipient	2,544 1,227 512,357 2,454
6	Health Insurance Institute of Slovenia	LQ	Persons in health insurance under specific schemes	1 January	Employed Self employed Contributing family worker	5,175 119 973

[5] HQ – High quality, LQ – Low quality, HQ-OD - High quality but outdated, LQ-OD - Low quality and outdated

[6] According to the internal standardized classification used in register-based census having 11 categories (Farmer / Contributing family worker joined as 1 category).

[7] Data from the Population Database as of 1 January 2014.

					Unemployed, previously in employment	3,382
					Pupil	11,904
					Student	3,483
					Pension recipient	19,012
					Social transfers recipient	16,761
					Other inactive person	69,126
7	Health Insurance Institute of Slovenia	LQ	Family members of insured persons	1 January	Contributing family worker	1,215
					Pupil	55,023
					Student	8,210
					Pension recipient	2,173
					Other inactive person	28,060
8	Ministry of Labour, Family and Social Affairs	LQ-OD	Recipients of social benefits	Previous calendar year	Social transfers recipient	774
9	Financial Administration of the Republic of Slovenia	HQ-OD	Income tax payers	Previous calendar year	Employed	4,955
					Self employed	892
					Farmer	3
					Pupil	138
					Student	433
					Pension recipient	645
					Other inactive person	42
10			Imputations		Employed	7,647
					Self employed	1,110
					Farmer / Contributing family worker	844
					Unemployed, never worked before	212
					Unemployed, previously in employment	1,202
					Pupil	1,099
					Student	816
					Pension recipient	2,709
					Social transfers recipient	655
					Other inactive person	6,548