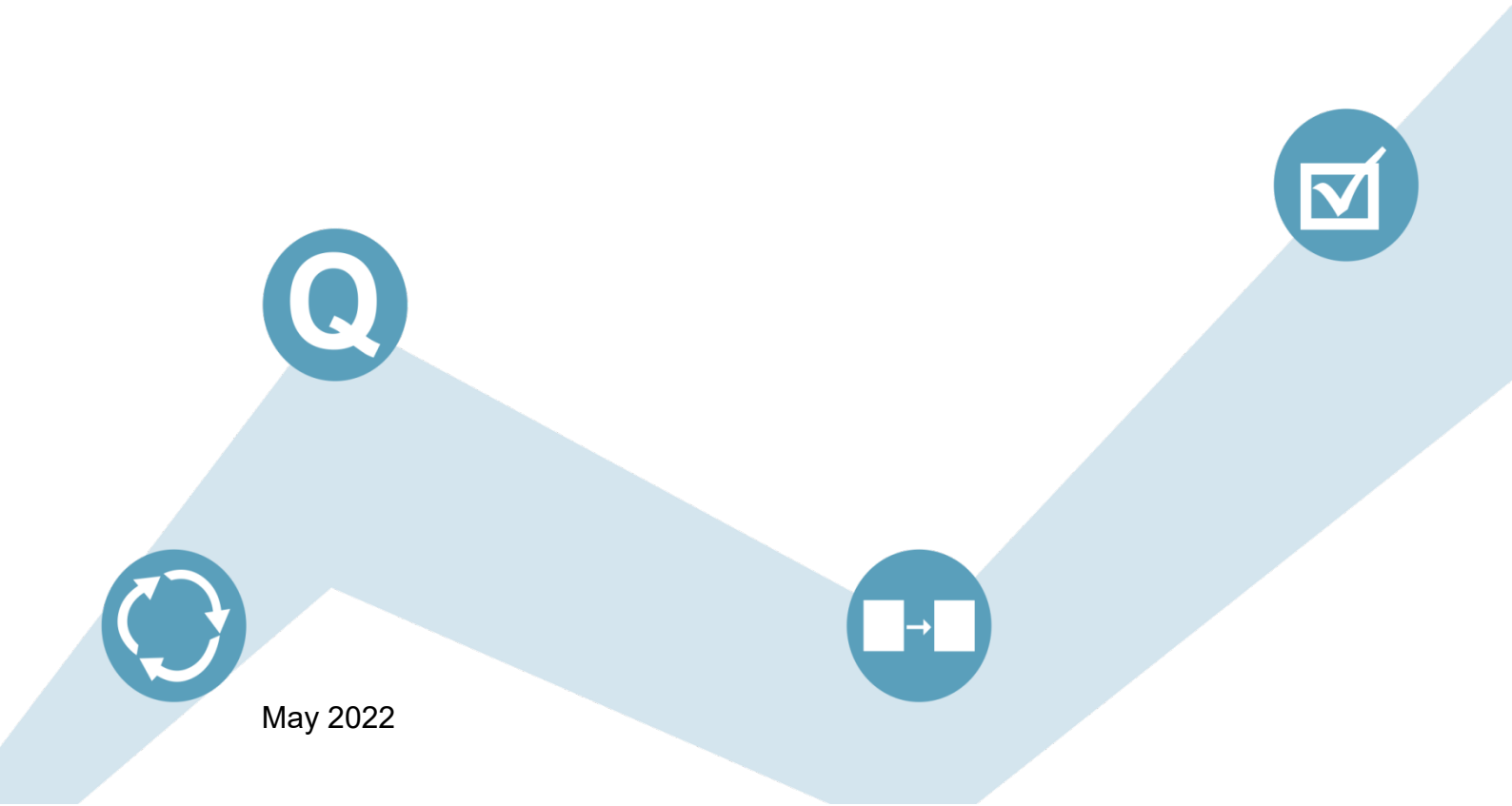
REPUBLIC OF SLOVENIA
STATISTICAL OFFICE

QUALITY REPORT

Tilen Mlakar

USAGE OF INFORMATION-COMMUNICATION TECHNOLOGIES IN HOUSEHOLDS AND BY INDIVIDUALS (IKT-GOSP)

FOR 2021



May 2022

CONTENT

METHODOLOGICAL EXPLANATIONS ON THE STATISTICAL SURVEY

Purpose of the survey

The purpose of the survey Usage of information-communication technologies in households and by individuals is to collect the data on developed the digital society in Slovenia is: how many individuals aged 16–74 years use the Internet and for which purposes, how many of them have digital skills and how many of their households have access to the Internet from home.

Selection of observation units

Observation unit is an individual aged 16–74 years and their households. Individuals are observed by age groups and sex, by education and sex, by status of activity and by degree of urbanisation, and their households by type of the household and by degree of urbanisation.

The survey is based on a sample, which includes only a part of the target population, on the basis of which we conclude about the characteristics of the entire population.

The basis for the sampling frame is the Central Population Register. The sample is stratified systematic. Strata are defined by statistical region (12 regions) and type of settlement within the statistical region (6 types). The sample size is 4,000 persons. The number of persons in an individual stratum is proportional to the share of persons aged 16–74 years living in an individual statistical region and in an individual type of settlement.

Sources and methods of data collection

Data are collected with the annual survey "Usage of information-communication technologies (ICT) in households and by individuals" (IKT-GOSP). The reference periods are the last three months and the last 12 months before the interviewing. Data are collected with the questionnaire "Usage of information-communication technologies (ICT) in households and by individuals" (IKT-GOSP). Every year the content of the survey and the questionnaire has an emphasis on a specific topic from the field of digital society.

Year	Content of specific module
2007	E-Skills
2008	Use of Advanced Services
2009	E-Commerce and trust
2010	Internet security
2011	E-Skills
2012	Mobile use of the Internet
2013	Usage of the websites of public authorities
2014	Usage of cloud computing services
2015	E-Commerce

2016	Provision and protection of data on the Internet
2017	E-Commerce
2018	ICT usage at work and digital skills
2019	E-skills and Internet security
2020	Usage of smart devices or systems
2021	E-skills and e-commerce

Since 2018 data are collected with a combination of online questionnaire (WEB) and with face to face interviews (CAPI). In 2020 and in 2021, we collected the data due to declaration of epidemic and social distancing with a combination of online questionnaire (WEB) and telephone interviews (CATI).

The selected person answers the questions. The majority of questions (questionnaire "Usage of information-communication technologies (ICT) in households and by individuals" – IKT-GOSP) refer to the selected person. A minor part of the questions refers to the household of the selected person (equipment of the household with ICT).

Data for the survey are not obtained from administrative sources.

Key variables

- Individuals who have used the Internet in the last 3 months,
- individuals who have used the websites of public authorities in the last month period,
- individuals who made purchases over the Internet in the last 12 month period.

Key statistics

- The number of individuals who have used the Internet in the last 3 months,
- the number of individuals who have used the websites of public authorities in the last 12 month period,
- the number of individuals who made purchases over the Internet in the last 12 month period.

Questionnaire

- Questionnaire:
    - Uporaba informacijsko-komunikacijske tehnologije v gospodinjstvih in pri posameznikih (IKT-GOSP))

Theme: Development and Technology, Subtheme: Digital Society

Questionnaires on SURSs website are available at https://www.stat.si/statweb/en/Methods/QuestionnairesMethodologicalExplanationsQualityReports.

Methodological explanations

- Methodological explanations:
    - Uporaba informacijsko-komunikacijske tehnologije v gospodinjstvih in pri posameznikih

Theme: Development and Technology, Subtheme: Digital Society

Methodological explanations on SURSs website are available at https://www.stat.si/statweb/en/Methods/QuestionnairesMethodologicalExplanationsQualityReports.

1 RELEVANCE

Relevance describes to what extent statistical data satisfy the needs of users. The rate is determined by whether all statistical data the users need are available and to what extent the published data and concepts used (definitions, classifications) meet the needs of users.

1.1 DESCRIPTION OF ADMINISTRATIVE SOURCES USED

1.1.1 Origin and original purpose of the data

Data for the survey are not obtained from administrative sources.

1.1.2 Method of takeover

Data for the survey are not obtained from administrative sources.

1.2 USERS OF SURVEY DATA

1.2.1 Key users of survey data

Table 1.1: Key users of survey data

Public sector	Ministry of Public Administration, Information Society and Informatics Directorate, Institute of Macroeconomic Analysis and Development
Business entities	
Science, research and education	University of Ljubljana: School of Economics and Business, Faculty of Public Administration, Faculty of Social Sciences
General public	no
Media	STA and other media
Foreign users	Eurostat, OECD, ITU
Internal users	

1.2.2 User needs and satisfaction

Direct communication with users takes place at the meetings of the Working Group on Information Society Statistics at Eurostat and the OECD and at the meetings of the statistical advisory committees group - the Statistical advisory committee for Digital Society Statistics. The members are representatives of key data users: from the public sector (Ministry of public administration), scientific sphere (University of Maribor) and they meet every 18 months. The last meeting of the panel was on 14 April 2022. More information on the content of committee is available at:

http://www.stat.si/statweb/NationalStatistics/AdvCommitteesDescription/91.

Communication with individual data users takes place through the Editorial board section or directly via the methodologies of the survey.

SURS measured general user satisfaction for the last time in 2021. Respondents assessed general satisfaction with SURS with the average score of 8.3 (on a scale from 1 – disagree completely to 10 – agree completely).

1.3 COMPLETENESS OF STATISTICAL RESULTS

1.3.1 Completeness of statistical results

Completeness of statistical results is the ratio between the number of statistical results that were disseminated (within a specific field) and the number of statistical results that were demanded (e.g. with regulations, in agreements). Statistical results that are not appropriate for a Member State or derogation is in force for them are not taken into account in the calculation.

1.3.2 Explanations

The index calculation is not applicable.

2 ACCURACY

Accuracy is defined as the degree to which the value at the end of statistical processing matches the true but unknown population value.

2.1 STANDARD ERRORS

2.1.1 Procedure for calculating standard errors

Standard errors were estimated using an analytical approach, using the Taylor linearization method.

2.1.2 Standard error

Standard error is the square root of the variance of the statistical estimate. To a large extent it is determined by the sampling error. It is also affected by other random errors in the survey implementation. The standard error can be shown also in the form of the coefficient of variation or the confidence interval.

Table 2.1: Standard error

Reference period	Domain name	Domain value	Statistics	Standard error
2020	Total	-	Internet access	0.01
2021	Total	-	Internet access	0.01
2019	Total	-	E-buyers - in the last 12 months	0.01

Reference period	Domain name	Domain value	Statistics	Standard error
2020	Total	-	E-buyers - in the last 12 months	0.02
2021	Total	-	E-buyers - in the last 12 months	0.01
2019	Total	-	Internet users - in the last 3 months	0.01
2020	Total	-	Internet users - in the last 3 months	0.01
2021	Total	-	Internet users - in the last 3 months	0.01
2019	Total	-	Users of websites of public authorities (e-government) in the last 12 months	0.01
2020	Total	-	Users of websites of public authorities (e-government) in the last 12 months	0.02
2021	Total	-	Users of websites of public authorities (e-government) in the last 12 months	0.01

2.1.3 Explanations

The table shows the standard error for key variables. The Statistical Office of the Republic of Slovenia draws attention to less precise estimates by flagging them with a special sign or by not publishing them at all. 1. If the table contains estimated population totals of (continuous) variables, publishing limitations are determined by the relative standard errors or the coefficients of variation (CV). In such cases it holds: If the coefficient of variation (CV) of the estimate is: 10% or below (CV <= 10%), the estimate is of acceptable precision and is published without limitations; in the interval from 10% and up to 30% (10% < CV <= 30%), the estimate is less precise and is flagged for caution with letter M; over 30% (CV > 30%), the estimate is too imprecise to be published and therefore suppressed for use by letter N. 2. If the table contains a number of units with certain characteristics, publishing limitations are determined by the standard errors of the estimates (SE) of the proportions. In such cases it holds: If the standard error (SE) of the estimate of a proportion is: 0.05 or below (SE <= 0.05), the estimate is of acceptable precision and is published without limitations; in the interval from 0.05 and up to 0.15 (0.05 < SE <= 0.15), the estimate is less precise and is flagged for caution with letter M; over 0.15 (SE > 0.15), the estimate is too imprecise to be published and therefore suppressed for use by letter N.

2.2 COVERAGE BIAS

2.2.1 Procedure for calculating the bias

2.2.2 Coverage bias

Coverage bias measures the error in a statistical estimate caused by the fact that a part of the population was left out of the observation on purpose. Bias due to other factors, such as non-response, measurement errors, processing errors, etc., is not estimated.

2.2.3 Explanations

The indicator is not calculated.

2.3 NON-SAMPLING ERRORS

2.3.1 Non-response errors

2.3.1.1 Unit non-response rate

The unit non-response rate is the proportion of eligible units for which we were not able to obtain any desired data or the obtained data were not useful. Unweighted and weighted values of the indicator can be calculated.

Table 2.2: Unit non-response rate

Reference period	Domain name	Domain value	Number of non-responses	Number of eligible units	Non-response rate (in %)
2019	Total	-	794	2446	32.46
2020	Total	-	1242	2494	49.80
2021	Total	-	2468	3997	61.75

Table 2.3: Weighted unit non-response rate

Reference period	Domain name	Domain value	Non-response rate (in %)
2019	Total	-	32.45
2020	Total	-	49.79
2021	Total	-	61.74

2.3.1.2 Explanations

To reduce the unit non-response rate, we use a notification letter and a reminder to explain the purpose of the survey to the respondents selected into the survey. To increase the understanding of the questionnaire we include notes and practical examples. We achieve a good response rate, which is obtained with a combined method of data collection. In 2021, the response rate of the unit was affected by the covid-19 epidemic. Due to the restriction of social contacts, we replaced the field survey with a telephone survey. The limited availability of telephone numbers has resulted in an increase of unit non-response.

2.3.1.3 Item non-response rate

The item non-response rate is the proportion of units for which we were not able to obtain data for a specific variable, even though the unit was eligible for this variable. In using administrative data the value of the indicator is the rate of unsuccessful linking of the variable showing the proportion of units for which we were not able to link units using one or more administrative sources and thus determine the variable value.

2.3.1.4 Explanations

The indicator is not calculated.

2.3.1.5 Imputation rate

The imputation rate is the ratio between the number of units for which the data for the key variable were imputed (for any reason) and the number of units for which we have at least a datum for the variable (after statistical processing). Unweighted and weighted values of the indicator can be calculated. The indicator is calculated only for key variables.

Table 2.4: Rate of unsuccessful integration of variables

Reference period	Domain name	Domain value	Variable	Imputation rate (in %)
2019	Total	-	Internet access in households	0.00
2020	Total	-	Internet access in households	0.00
2021	Total	-	Internet access in households	0.00
2019	Total	-	Education	0.24
2020	Total	-	Education	0.88
2021	Total	-	Education	0.46

Table 2.5: Weighted rate of unsuccessful integration of variables

Reference period	Domain name	Domain value	Variable	Imputation rate (in %)
2019	Total	-	Internet access in households	0.00
2020	Total	-	Internet access in households	0.00
2021	Total	-	Internet access in households	0.00
2019	Total	-	Education	0.29
2020	Total	-	Education	0.74
2021	Total	-	Education	0.29

2.3.1.6 Explanations

We do not impute data for the variable Internet access in households. Data for the variable education are imputed in cases when data are not available from the administrative source.

2.3.2 Coverage errors

2.3.2.1 Overcoverage rate

The overcoverage rate is the proportion of units in the sample frame that are not part of the target population (ineligible units). If the survey is conducted based on the sample, the overcoverage rate is estimated based on information from the sample, meaning that the weighted value of the indicator is calculated and disseminated.

Table 2.6: Overcoverage rate

Reference period	Domain name	Domain value	Number of ineligible units in the frame	Number of all units in the frame	Overcoverage rate (in %)
2019	Total	-	32815	1419752	2.31
2020	Total	-	5524	1388756	0.40
2021	Total	-	1000	1337592	0.07

2.3.2.2 Explanations

Over-coverage occurs due to the unavailability of data on the suitability of the selected unit at the time of the preparation of the sample, and due to the delay

between the design of the sample and the implementation of the survey.

2.3.2.3 Undercoverage errors

Undercoverage errors occur when certain units are not included in the sample, but they should be. The sample frame does not include dwellings and households in which no individuals has registered permanent or temporary residence. There are also no homeless individuals in the sample.

2.3.3 Measurement errors

2.3.3.1 Error detection controls

In the process of surveying and collecting data, we included logical controls (so-called light and hard controls) in the questionnaire, which warn the respondent that the answers are not correct or contradictory with already provided answers. Upon completion of the survey, the data is uploaded to a database, where data control is performed with the help of a software.

Individual and occasional systemic corrections are carried out. The type of correction depends on the content of the questionnaire in each year.

Error code	Comment
Light control	Warning! In question A4, you answered that you arranged a visit to the doctor at the health center, via an online form. This means that you have submitted a completed electronic form via a website of public authorities.
Hard control	Error! If you have carried out any of the listed educational activities, you cannot answer that you have not carried out any educational activities over the internet!

2.3.3.2 Reasons for measurement errors

Most errors occur in the process of collecting data. They are caused by the respondent's misunderstanding or not correct understanding of the questions. Mistakes can also occur due to the negative, disinterested attitude of the respondents or their ignorance. The participation of selected persons in the survey is voluntary, so they are not legally obliged to provide reliable data.

2.3.3.3 Editing rate

The editing rate is the ratio between the number of units for which we corrected the values of a certain variable in the editing phase and the number of all units that reported the data for the variable. Unweighted and weighted values of the indicator can be calculated. The indicator is calculated only for key variables.

Table 2.7: Editing rate

Reference period	Domain name	Domain value	Variable	Editing rate (in %)
2019	Total	-	Internet access in households	0.00
2020	Total	-	Internet access in households	0.00
2021	Total	-	Internet access in households	0.00
2019	Total	-	Education	0.00
2020	Total	-	Education	0.00
2021	Total	-	Education	0.00

Table 2.8: Weighted editing rate

Reference period	Domain name	Domain value	Variable	Editing rate (in %)
2019	Total	-	Internet access in households	0.00
2020	Total	-	Internet access in households	0.00
2021	Total	-	Internet access in households	0.00
2019	Total	-	Education	0.00
2020	Total	-	Education	0.00
2021	Total	-	Education	0.00

2.3.3.4 Explanations

We do not edit the data for the variable Internet access in the households. Other non-key variables are edited, which change annually with the content of the questionnaire.

2.3.3.5 Coherence of data sources

The indicator is calculated when we have the data for at least some units of the selected variable from two sources and both sources are used as data sources for this variable. The value of the indicator shows coherence of data from these two sources.

2.3.3.6 Explanations

The index calculation is not applicable.

3 TIMELINESS AND PUNCTUALITY

Timeliness of the release measures the time lag between the period to which the data refer and the release date.
Punctuality of the release measures the time lag between the actual and previously announced date of data release, which is determined in the release calendar. If the mentioned dates coincide, the release is punctual.

3.1 TIMELINESS

3.1.1 Timeliness of the first release

Timeliness of the first release measures the time lag between the date of the first release of the data and the reference date of the observed (reference) period. The reference date is usually the last day of the (reference) period to which the published data refer, but it can also be another key date within this

period (e.g. start of the school year).

Table 3.1: Timeliness of the first results

Reference period	Reference date of the period	Date of publishing	Time lag (in months)
2019	1. 03. 2019	8. 10. 2019	7
2020	1. 03. 2020	8. 10. 2020	7
2021	1. 03. 2021	7. 10. 2021	7

3.1.2 Timeliness of the final results

Timeliness of final results measures the time lag between the date of the release of final data and the reference date of the observed (reference) period. The reference date is usually the last day of the (reference) period to which the published data refer, but it can also be another key date within this period (e.g. start of the school year).

3.1.3 Explanations

The period of publication of data does not change over the years.

The index calculation is not applicable (timeliness of the final results).

3.2 PUNCTUALITY

3.2.1 Punctuality of the first release

Punctuality of the first release measures the time lag between the announced and actual date of the first release.

Table 3.2: Punctuality of the first results

Reference period	Announced date	Date of publishing	Time lag (in days)
2019	8. 10. 2019	8. 10. 2019	0
2020	8. 10. 2020	8. 10. 2020	0
2021	7. 10. 2021	7. 10. 2021	0

3.2.2 Reasons for deviations in punctuality

The data were published according to date specified in the Release calendar.

4 ACCESSIBILITY AND CLARITY

Accessibility of statistical data describes the possibilities available to users for simple access to statistical data. It refers to physical circumstances in which the data are available to users: where and how the data can be obtained, when they will be available, how much does the service cost (clear price list of services), conditions for using the data (copyright), availability of microdata and metadata, availability in various formats. Clarity of statistical data describes how simple it is for users to understand the data. It refers to the information environment in which the data are presented: are the data equipped with

appropriate methodological explanations and are they properly presented with graphical presentations or other material, is information on punctuality of data and on limits to use available to the users, do the users have access to additional information should they need it.

4.1 ACCESSIBILITY

4.1.1 Frequency of publication

The data are published annually.

4.1.2 Means used for dissemination

Table 4.1: Means used for dissemination

Means	**Used**
Web release	yes
Tables in the SiStat Database	yes
Publication in interactive tools and applications	no
Printed publications	yes
Publication on social media	yes
International databases (Eurostat database, OECD database)	yes
Microdata of the statistical survey	yes
Metadata	yes

4.2 CLARITY

4.2.1 Disseminated results

The results are presented in absolute figures and in shares (press releases). Data are presented in graphical and tabular form.

Methods of seasonal adjustment are not used in this survey.

4.2.2 Level (detail) of dissemination

Data are published at the state level, by the degree of urbanization of the settlement where the inidvidual lives and data for his/her household at the state level, by degree of urbanization of the settlement and by NUTS 2, where the household resides.

Data are published by age classes and sex, education and gender, status of activity and his household by type of household.

5 COMPARABILITY

Comparability of statistics measures the differences due to the use of different statistical concepts (classifications, definitions, target population) or different statistical methods in calculating statistics in different geographical areas, other domains of the population or different time periods.

5.1 COMPARABILITY OVER TIME

5.1.1 Time coverage

Data are available from 2004 onwards.

5.1.2 Length of comparable time series

The indicator shows the length of the time series since the last break in the time series, i.e. the number of points in the time series since the last break.

Table 5.1: Length of comparable time series

Reference period	Statistics	First comparable time period	Length of comparable time series
2021	Internet access	2004	18
2021	E-buyers - in the last 12 months	2004	18
2021	Internet users - in the last 3 months	2004	18
2021	Users of websites of public authorities (e-government) in the last 12 months	2004	18

5.1.3 Explanations

The length of comparable time series for each statistics depends on the inclusion of statistics in the survey, data needs, changes in information technology and survey revisions.

5.1.4 Factors influencing comparability over time

There are no breaks in time series, so all points in time are comparable.

Until 2017 data was published under the following titles:

- First Release (Development and Technology, Digital Society): »Usage of information-communication technologies in households and by individuals, Slovenia, annually«.
- Electronic Release (Development and Technology, Digital Society):): »Usage of information-communication technologies in households and by individuals, Slovenia, annually«.

5.1.5 Seasonal adjustment

Seasonal adjustment is not applicable.

5.2 GEOGRAPHICAL COMPARABILITY

5.2.1 Comparability with other international organisations

Data are collected on the basis of Commission Regulation (EU) No. 808/2004 on Community statistics on the information society, as well as the annual implementing regulations, which define the variables to be monitored in each year. The data are transmitted to Eurostat once a year and are comparable with the data from other members of the European Statistical System.

6 COHERENCE

Coherence in statistics is the adequacy of statistical data to be reliably combined in different ways and for various users. It describes limitations in combining statistics from different sources that are the result of using different statistical procedures.

6.1 COHERENCE BETWEEN PROVISIONAL AND FINAL DATA

6.1.1 Policy of releasing provisional data

Provisional data are not disseminated. Only final data are published.

6.1.2 Coherence between provisional and final data

Coherence between provisional and final data shows absolute or relative difference between the disseminated value at first release and the value at the release of final data. Revisions are only corrections that are part of the regular procedure of publishing statistical results. Corrections due to errors are not revisions and are not taken into account in calculating the indicator. Even though the revision policy stipulates several versions of (provisional) data for the same reference period, for the sake of simplicity only the difference between the values of provisional data in the first release and the values in the final data release is calculated.

6.1.3 Reasons for larger differences between provisional and final data

Provisional data are not disseminated. Only final data are released.

6.2 COHERENCE WITH THE RESULTS OF THE REFERENCE SOURCE

6.2.1 Brief description of the reference source

6.2.2 Coherence with reference sources

Coherence with reference sources shows absolute or relative difference in view of the results of the selected reference survey in which the same or at least a related and therefore a comparable phenomenon is observed but which can be observed with a different methodology or a different periodicity. Results of short-term surveys can be compared with results of structural surveys, and vice versa. Results of surveys can also be compared with results disseminated by national accounts or various institutions in the country or even abroad (e.g. mirror statistics on tourism).

6.2.3 Reasons for larger differences between sources

There is no known reference source for this statistical survey.

7 BURDENS

Burden of interviewed persons and business entities is not a separate quality component but an important factor in assessing the quality, since it usually has an impact on all other quality components.

7.1 BURDEN OF REPORTING UNITS

7.2 EXPLANATIONS

The indicator is not calculated.

8 EVALUATION OF SURVEY QUALITY

At the end of survey implementation, quality indicators and other important information resulting from the survey implementation are reviewed. On this basis survey implementation is evaluated, weak points regarding survey implementation are identified and improvements for the next implementation are planned.

8.1 GENERAL QUALITY ASSESSMENT

The research has a satisfactory quality. The biggest challenge is the formulation of questions. Due to technological advances and the content of the survey, these are often technical and difficult to understand for individuals in different age groups. Due to this reason, we cognitively test the new questions.

Due to COVID-19 limitations CAPI could not be done. SURS has no legal means to acquire all telephone numbers from the individuals. This year we had around 25% telephone numbers and this number keeps declining every year. Since we were dependant on WEB and CATI interviews there was a problem contacting enough of the individuals through telephone. For these reasons, the response rate may be lower than in years when it was possible to conduct a field survey (CAPI) in combination with an online survey (WEB).

8.2 MEASURES FOR IMPROVEMENT

Planned improvements:

- in order to reduce the non-response of units, efforts will be intensified to emphasise the importance of participation in the survey;
- to reduce measurement errors, the questions will be cognitively tested and, if necessary, additional explanations will be provided.