In accordance with Act no. 89/1995 concerning the state statistics service and with Act no. 101/2000 on individual data protection, the Czech Statistical Office carried out in February to April 2007 the third sample survey already of income and living conditions of households under the official title of "Living Conditions 2007."

The obligation to run this survey in the Czech Republic is embedded in the amended 1177/2003 framework Regulation and its implementing Commission regulations. The EU co-financed the survey.

The aim of the survey was to gather representative data on income distribution for the whole population and for various household types, data on housing – its quality and affordability, household durables, and labour, financial and health conditions of adults living in private households. Besides the unchanging annual part of the survey a module was included focusing on housing.

1. Organization of the survey

1.1 Sampling

As in the previous years, the survey was carried out on the whole territory of the Czech Republic. 4250 new dwellings entered the survey (1st wave) and 7461 dwellings were revisited (2nd and 3rd waves). The latter figure includes households that were traced down to their new homes (after their having moved).

The sample was obtained by utilizing two-stage probability sampling scheme independently for each of the 14 administrative regions (NUTS3 regions). The total number of dwellings selected in each region was chosen to be proportional to the region's size. At the first sampling stage small geographical areas (CEU's - census enumeration units or districts) were selected by probability sampling. These CEU's served as a basis for the second-stage selection (a sample of 10 dwellings was drawn from each CEU).

Before selecting the sample of dwellings, the sampling frame had to be adjusted to enable incorporation of small census enumeration units into the sampling process to reach the required full geographical coverage of the national territory. Small CEUs (with less then 20 inhabited dwellings) were merged with adjacent CEUs and the resulting larger CEUs entered the first stage of sampling. Consequently, in some cases, the 10 dwellings sampled in the second stage belong to two, in exceptional cases even more, real administrative CEUs.

The CZSO's regional fieldwork units (each covering one of the 14 NUTS3 administrative regions) received the list of selected dwellings (address + identification number of the flat in apartment buildings). Before the actual fieldwork, the regional fieldwork units' staff carried out identification of the selected dwellings and filled in the contact names on the list of selected dwellings for interviewers.

1.2 Fieldwork

Data collection in the field lasted from February 17th to April 29th 2007 and was coordinated by workers from regional departments responsible for fieldwork. Interviewers were hired and trained at the regional level. Their remuneration was based on number of visited and number of successfully interviewed households. Data collection in households was the most difficult part of the survey. Interviewers were facing general resistance to giving information (particularly on

income) and had to patiently explain the reasons for conducting such a survey and why the selected household should participate in it.

The sampling unit was the dwelling, i.e. all persons with usual residence in that dwelling (their only or main place of residence) were included in the survey. This includes also foreign nationals and sub-tenants living in the selected dwelling. Data was collected by interviews and interviewers filled in the answers into paper questionnaires (PAPI data collection).

The content of the survey was divided into four questionnaires with different units of reference:

Questionnaire A (dwelling unit questionnaire): contained the roster with the list of all persons with usual residence in the selected dwelling, their basic demographic characteristics, information on sharing of expenses to determine household units[1] and relationship of each person to the main user of the dwelling and to the head of household.

Questionnaire B (household questionnaire): filled in for each household, contained information on housing, childcare, financial situation of the household, consumer durables, inter-household transfers paid and received, consumption from household own production (i.e. small scale farming and similar activities), family social benefits, rental income and paid regular taxes on wealth (buildings and land).

Questionnaire C (personal questionnaire): filled in by each household member aged 16+ as of 31.12.2006 (i.e. persons born in 1990 and earlier). This questionnaire contained information on labour status and employment, personal income (from employment, private enterprise and social security schemes), participation in private pension plans, health and selected biographical information.

1.3 Processing of the questionnaires and collected data

Regional survey coordinators were responsible for collecting the questionnaires from interviewers, initial visual and systematic check of the collected data and the process of preparation of questionnaires for subsequent optical scan. Data were then captured using OCR technology. After the initial central checks (integrity of questionnaire identification numbers, completeness of the regional sets of questionnaires), the datasets with in-house developed software application and electronic images of the scanned questionnaires were sent to regional units for further logical checking and editing. The edited data were then delivered to the CZSO for further processing.

1.4 Successfully interviewed households and non-response

The fieldwork revealed that among the total of 11 924 dwellings in the sample there were 428 dwellings (3,6 %) unoccupied, unlocated or ineligible because the households had moved. Since there was no substitution for these ineligible units, the survey was conducted in 11 496 dwellings and 11 611 households (in some of the dwellings there is more than 1 household). The overview of the survey response can be summarised by the following table:

[1] Since the household definition is based on sharing of expenditures (housekeeping concept), there are dwelling units with more than one household. If this was the case, all households in selected dwellings were included as eligible for the survey.

	Households			Response (%)		
	Total	1st wave	2nd+3rd wave	Total	1st wave	2nd+3rd wave
Response, total	9675	2654	7021	83,3	64,9	93,4
Non-response, total	1936	1437	499	100,0	100,0	100,0
Refusals (unwillingness to give information)	1515	1138	377	78,3	79,2	75,5
household not contacted, temporarily absent	363	257	106	18,8	17,9	21,2
household unable to respond (health limitation)	47	31	16	2,4	2,2	3,2
Other reasons (linguistic etc.)	11	11	0	0,6	0,8	-

Refusals also include situations when the household did not refuse the survey as such, but did not accept to provide the information on income to the extent, which would qualify the household as successfully interviewed. The definition of successfully interviewed household allowed missing income data for only one person provided that the person is not the head of the household.

Non-contacts, temporarily absent category cover situations, when the interviewer did not establish contact with the selected household, despite the prescribed minimum number of three attempts of personal contact.

Region (NUTS3)	Total			1st wave			2nd + 3rd wave		
	HHs in survey	response		HHs in survey	response		HHs in survey	response	
		count	%		count	%		count	%
Hl. m. Praha	1237	864	69,8	538	254	47,2	699	610	87,3
Středočeský	1207	1006	83,3	462	326	70,6	745	680	91,3
Jihočeský	704	612	86,9	250	188	75,2	454	424	93,4
Plzeňský	669	562	84,0	210	129	61,4	459	433	94,3
Karlovarský	364	328	90,1	124	101	81,5	240	227	94,6
Ústecký	951	787	82,8	334	216	64,7	617	571	92,5
Liberecký	476	391	82,1	172	111	64,5	304	280	92,1
Královéhradecký	611	513	84,0	209	134	64,1	402	379	94,3
Pardubický	596	513	86,1	198	143	72,2	398	370	93,0
Vysočina	586	510	87,0	180	120	66,7	406	390	96,1
Jihomoravský	1176	948	80,6	433	252	58,2	743	696	93,7
Olomoucký	778	666	85,6	260	171	65,8	518	495	95,6
Zlínský	657	576	87,7	218	163	74.8	439	413	94,1
Moravskoslezský	1599	1399	87,5	503	346	68,8	1096	1053	96,1
CR total	11611	9675	83,3	4091	2654	64,9	7520	7021	93,4

Participation in this survey is voluntary, there is no duty imposed on households to provide the required information, like it is for example in the population census. The household must be informed about the content of the survey and that its participation is voluntary and left to its decision. The main reasons for refusal reported from the field are privacy reasons (objections against giving personal information and fear of abuse of the personal data), fear of contact with interviewers as strangers. There is a considerable group of persons, who as a matter of principle strictly refuse to give any information.

1.5 Grossing up and weighting

When compared with data from other statistics and registers, selected characteristics of our sample showed that a phenomenon typical of household surveys had occurred - high level of non-response (influenced with a rotational panel by prior response) had biased the proportions in the final data file from which results are obtained.

The deformation of demographic characteristics and social structure of the sample did not allow us to use simple techniques of grossing up (post-stratification). Sufficient level of bias elimination is a necessary pre-condition for obtaining good estimates. In practice, standard methods of arriving at optimum grossed-up results are used - in our case, iteration method of weight calibration was utilized minimizing the difference between the known and the grossed up values of selected characteristics.

Because of the fact that results are required for both households and individual persons, the only satisfactory solution is the system of integrated weights, i.e. a single set of grossing-up coefficients. With the rotational panel where the final dataset is composed of several subsamples (apart from the new households in the sample also the revisited households from the previous years) various ways to tackle this problem exist. Generally, it is recommended that the sample be divided into subsamples according to waves and that richer information for revisited households be utilized. In practice, however, this procedure did not result in greater consistency of the estimates from the point of view of the panel, quite to the contrary, it lead to unnecessarily high variability of weights and thus to lower stability of the estimates. That is why a simpler calibration method was applied not differentiating between the waves.

As the basis for calculations the following traditional calibration variables were used:

- Number of inhabited dwellings in each NUTS3 region, subdivided into family houses (detached and semi-detached houses) and apartments, based on the 2001 Census continuously updated from administrative sources of construction authorities
- Population characteristics:
  - Population totals in each NUTS 3 region (from demographic statistics)
  - Economic activity characteristics in each NUTS3 region:
    - Number of pensioners (excl. pensions for orphans), based on the administrative data from social security administration
    - Number of unemployed (registered unemployed from the administrative source of the Ministry of Labour and Social Affairs, corrected for unregistered unemployment using the Labour Force Survey data)
    - Number of self-employed (estimate based on the Labour Force Survey)
    - Number of children aged 0-15 (from demographic statistics)

- Demographic characteristics at the national level (based on the demographic statistics):
  - Age groups (0-15, 16-24, 25-34, 35-44, 45-54, 55-64, 65+)
  - Sex
  - Municipality size (below 2 000 inhabitants, 2 000 - 9 999, 10 000-49 999, 50 000+ inhabitants)

Since the target population of the survey were persons living in private households, the data from demographic statistics were adjusted by subtracting institutionalised population (from social security administrative data) and persons in prisons.

As the sampling unit is the dwelling, all weight coefficients were calculated for dwellings and only subsequently assigned to all persons and households in them.

The method described above deals with non-response successfully, i.e. it corrects the bias due to similar composition of households that failed to respond. First of all, it improves demographic and social structure but, as a by-product, it also eliminates deformation of income indicators related to these structures.

Another source of bias, which needs to be taken into account, stems from the interviewing. Data on income obtained during interviews with household members have the tendency to underestimate certain income sources or data on some income components can be altogether missing (item non-response). So as not to reduce the size of the processed dataset pointlessly the missing income was imputed using correct statistical methods.

In Living Conditions 2007 the number of cases in which the interviewer obtained most of the demographic characteristics but failed to receive a person's income data was very low. The missing income of such individuals (only 7 in the whole dataset) was conjectured by the simple hot-deck method, i.e. it was replaced with income of another, randomly selected person with the same characteristics.

Underestimation of income is a natural consequence of the fact that respondents either tends to give lower than actual values or simply do not recall having had certain irregular or small incomes at all. It is, more or less, a non-sampling error, affected substantially by the incomes themselves and by their source. The possibilities to eliminate this underestimation of the survey data are limited. In the presented survey, only such adjustments were made where there was sufficiently reliable external statistical source or where the conjectures could be based on legislation.

Data on gross income from employment were compared with corresponding data from wage statistics broken into sectors of activity. In accordance with experience from other income surveys, income from work was underestimated (roughly by 4 %). Primarily, this underestimation concerned those incomes that were recorded as yearly lump sums. Such incomes were moderately raised so that the average monthly gross pay by sectors approached the data from wage statistics. With income from private enterprise, there was no need for corrections.

In the case of social benefits for which there is legal entitlement (parental leave benefit, child birth benefit, death grant provided to families of the deceased, to some extent also maternity leave benefit), a check on their receipt by eligible households was applied and amounts provided were corrected according to the amounts set by the legislation. CZK 1000 grant for books and other school equipment to children entering primary education was added based on children's age. With old age benefits (pensions from the social security system) the tendency to underestimation is negligible but provided there were falls in this kind of income without any outward reason, the amounts were corrected by the last year's values.

Amounts declared by the unemployed as unemployment benefits were overestimated. Unemployed respondents tend to report their income from the social security system as unemployment benefits and do not distinguish them from the minimum income support benefits (claimed on the basis of legal minimum subsistence amounts). In cases where the duration of unemployment and the reported amounts did not match the rules of the unemployment benefits provision, the reported amounts were re-classified as minimum income support benefits or other social benefits.

It was not possible to correct the underestimation of sickness benefits (where omissions related to short-term illnesses could not be identified on the existing data), means-tested social benefits whose claims depend on the previous income (prior to the income reference period), capital income, and non-monetary income generated by own-consumption.

Comparison of the aggregated income from this survey with the household sector aggregates of the national accounts (even after their modification taking into account the items, which are not covered by household income surveys) is relatively difficult. Concerning its aggregated value the income obtained by direct questioning in the households will always be lower. The more important fact for evaluation of their credibility is that the trend in development of household income is in line with the trends in the national accounts. From this viewpoint, the presented results of LC 2007 are in full agreement with data from the previous year and with related statistics from developed nations of the European Union.

2. Methodological notes to published tables

2.1 Basic definitions

The publication contains the results for households and for individuals aged 16 and older. The household definition is based on the sharing of expenditures concept, in line with the definition of Paragraph 115 of the Civil Code - based on the declaration of the persons in sampled dwelling unit that they permanently live together and finance together expenditures to cover their needs. As the 16-year olds such persons were regarded who had reached this age by 31 December 2006.

Reference periods:

- Age: December 31st, 2006
- Other demographic variables - marital status, education, housing, financial situation: on the date of the interview
- Economic activity was collected pertaining to individual months of 2006 and 2007 (up to the date of the interview). What was gathered was the self-definition of the respondent. Subsequently, other data was collected related to the respondent's work activity (status in employment, profession). At the same time, and also pertaining to individual months, parallel activities were surveyed (second job, study) together with data on receipt of pensions and social benefits.
- Economic activity pertaining to the whole 2006 was not collected but derived from the monthly data (the activity with the highest incidence was coded as the yearly value).
- Current employment variables (current employment status, occupation): on the date of the interview
- Income data (both monetary and in kind): calendar year 2006
- Subjective questions focused on housing and financial problems: on the date of the interview. Health problems: last twelve months.
- Housing, consumer durables, financial and social situation of household: on the date of the interview, unless the question specifically refers to some other period

2.2 Description of variables

2.2.1 Household composition

Size of the household - number of household members on the date of the interview, including persons temporarily away if the period of actual or foreseen absence is shorter than 6 months and the person has no other private address. For persons studying away from home, the period of absence may be longer than 6 months, provided that the person has no private address and retains financial ties to other household members. Persons with a period of absence longer than 6 months, persons without financial ties to the household and persons temporarily present at the time of the interview who have their private address elsewhere are excluded.

Employed - during 2006 the prevailing economic activity status of these persons was employed (employees, self-employed, members of production cooperatives, unpaid family workers in family businesses). Persons drawing sickness benefits, students who apart from their study worked (in employment, private enterprise), pensioners or persons on maternity leave with regular income from work were also included.

Dependent children - national definition in line with the Act 117/1995 on state social support; maximum age is 25, provided the person is still in education.

Pensioners (without economic activity) - persons receiving pension from social security system (old-age, disability, survivor's) without regular income from employment.

Unemployed - persons who did not have a job at the moment but who wished to have one. Such persons did not have to comply with the strict ILO definition about actively seeking a job and readiness to enter one.

Persons on parental leave - persons with prevailing status receiving parental benefit, without regular income from employment.

Other persons - inactive persons caring for household or household members in the need of care, persons living on property income and others.

Incomes are presented as household incomes, per capita incomes or equivalised incomes (using the standard and modified OECD equivalent scale) - see the headings of individual tables.

2.2.2 Household characteristics

Head of household - for couples with or without children it is always the male, regardless of his economic activity. In lone-parent families (one parent with children) and in non-family households (persons not related by marriage or partnership, nor parent - child relationship) the first criterion for determining of head of household was economic activity and the secondary criterion was income of household members. This rule was also applied in more complicated household types (for example in the case of sharing expenditures among more two-parent families).

Household type - is based on household composition. Two-parent families are based on a couple (married or cohabitating), with or without children. Lone-parent family’s category contains households with one parent and at least one child. These households may in addition to these basic structures contain other household members. The households where all children are dependent and there are no other members but one or both parents are labelled as nuclear families.

Household type (EU definition) - in contrast to the previous definition, this typology does not depend on family structures and is based on more “economical” concept of simply number of adults and number of dependent children. Subsequently to changes in calculation of Laeken indicators a change was introduced concerning age of dependent children. Dependent children are now all

persons 0-17 and, further, persons 18-24 who are economically inactive and live with at least one parent. Households of individuals and two adults were further divided into age groups: individuals with age below 65, individuals with age 65+, two adults both aged below 65 and two adults with at least one person aged 65+.

Duration of marriage - only for married couples living together.

Education - 4 categories (primary, secondary-vocational, complete secondary, tertiary). Complete secondary includes also vocational education with secondary school-leaving exam and post-secondary non-university education. Tertiary includes all tertiary programs - baccalaureate, graduate and post-graduate level.

Occupation - 9 main classes of national classification KZAM. Households are classified according to occupation of the head of household. Soldiers were coded as 1 - Legislators, senior officials and managers.

Household group (former social group) was based on the status of the head of household. Status of all other household members did not play any role. Only for households of pensioners, economic activity of other household members was used as a secondary classification criterion.

- ***Households, total***: represent the average household in the Czech Republic
- ***Households of employees***: household head's prevailing activity status is employee
  - Households of lower employees: education of the head of household is primary or secondary-vocational
  - Households of higher employees: education of the head of household is complete secondary or tertiary
- ***Households of self-employed***: household head's prevailing activity status is self-employed (in whatever field it may be, including agriculture)
- ***Household of pensioners***: household head's prevailing activity status is pensioner without economic activity, this group is further divided into two groups based on whether there is anybody in the household who works
- ***Households of unemployed***: household head's prevailing activity status is unemployed (at the same time, in complete families the female partner or a grown-up child can by employed)
- ***Other households***: household head's prevailing activity status is other than one the four previous categories (for example a person on parental leave benefit, student, person living on property income)

Subsistence minimum is based on the amounts of national subsistence minimum applicable in 2006. The amounts were the following:

Individual amounts	CZK	Household amounts	CZK
Children 0-5	1 750	1 person	2 020
Children 6-9	1 950	2 persons	2 630
Children 10-14	2 310	3-4 persons	3 260
Children 15-25	2 530	5 and more persons	3 660
Adults	2 400		

Monthly amount of subsistence minimum is then the sum of amounts for individual household members and the household amount.

2.2.3 **Monetary and non-monetary income**

Incomes related to household as a whole were collected at the household level: social benefits targeted at households, rental income and value of goods produced directly by the household through either a private or a professional activity (e.g. own production of food from farming).

Collected at individual level: income from employment (main job, secondary jobs), sickness benefits, old-age benefits, unemployment benefits, social benefits attributable at individual level (such as parental leave benefit or disability benefits) and other incomes (capital income, sales of property, insurance claims).

Income from employment (both main job and possible secondary jobs) was collected both either gross of tax and social insurance or net, incomes from non-employment contracted work only gross. Self-employed persons could choose from several ways to record the result of their enterprise. They could state the gross profit/loss according to their tax declaration, they could give the sum which served as the yearly basis for calculating their monthly health and social security contributions or could make their own estimate of their gross or net profit/loss. With family members co-operating in private enterprise run by another member of the family only proportionate part of the income from the business was entered. Rental income was collected either gross or net, based on what information respondents were able to provide. All other kinds of income were collected net and subsequently appropriate rules of the tax system were applied to estimate the gross amounts. In addition, the information was collected on claimed tax deductibles to enable calculation of taxes and social insurance contributions. Sum of individual net incomes then forms the main national indicator – net monetary income of household, fully comparable with previous national Microcensus income surveys.

Besides this national indicator of household income, it was necessary to construct internationally comparable household income indicator, which is based on Eurostat methodology for EU-SILC surveys. This indicator is named "disposable household income". The difference between these two definitions of household income is in inclusion/exclusion of certain components of income (received lump sum and irregular inter-household transfers, non-cash employment income, income from life insurance, regular taxes on immovables).

The value of income in kind was an estimate of the household based on the amount of consumed food and other goods, own production and goods from own business during the year 2006 (for example food and animals from own small-scale non-commercial farming activity, value of meals from own restaurant, bread from own bakery and the like). Also included is the value of company car for private use (as non-monetary income of employees). In this case, the lowest possible amount applicable for taxation in the tax law is added to the non-monetary income of the employee (CZK 1000/month).

Detailed income components are presented in table 1. Many of the income components values are quite low. Therefore, the breakdowns in other tables are less detailed. Somewhat more detailed breakdowns are provided for gross income.

Selected income components:

- Income from employment: defined in line with the national tax law. Includes income from employment contract or similar arrangement between employer and employee. Also includes incomes of owners of the incorporated business from work for their company, income of members of statutory boards and other governing bodies of corporations, remuneration based on holding of elected public posts, income of apprentices in vocational schooling for their work undertaken as part of their practical training and income from flexible short-term contracts under special regime set in the Labour Code.

- Income from self-employment: includes also income from farming activities, if these are the professional activity, income from independent professional practices (lawyers, doctors) and income from intangible assets (copyrights).

  Income from main employment: includes income of employees from their main job. In case of multiple coincident jobs, the declaration of the main job was left to the respondent.

  Income from secondary employment: includes salaries from secondary jobs, conducted besides the main job or self-employment activity of the respondent and income from flexible short-term contracts under special regime set in the Labour Code.

  Income from secondary self-employment activity: analogous to the secondary employment income. It includes income from secondary self-employment activity undertaken in addition to the main job of the respondent (where respondent declared employment contract as his/her main job).

- Social income: is in principle net. Gross amounts were included only for rare cases of pensions above the tax-exempt limit. In these cases, tax was applied to the amount above this limit (CZK 198 thousand).

  Sickness benefits item includes all sorts of benefits from the social sickness insurance, i.e. also maternity leave benefit, reduced employment income compensation in pregnancy and motherhood, income support for persons caring for household member in the need of short-term care (mostly care for children during their illness).

  Other social support benefits include social benefits for foster parents taking care of adopted children, birth grants, death grants, and CZK 1000 grant for books and other equipment of children entering primary education.

  Other social benefits include certain benefits connected to the termination of employment in selected professions, various other social benefits like benefit for persons providing long-term homecare for their relative in need, support for care in spas and other social benefits for families with children, old and disabled citizens, which are mostly administered by the municipal authorities.

  Minimum income support benefits include regular and lump sum monetary benefits and benefits in-kind granted to the household according to the national law on social needs.

- Other income

  Income from capital contains interest from savings, bonds and various forms of deposits, dividends from shares, profits from incorporated businesses, income from investments abroad.

  Other income includes income from occasional property rentals, life and material insurance and income from organisations not elsewhere classified (scholarships and pocket money of apprentices, grants from charity and non-governmental organisations, lottery winnings, prizes, pay for occasional not contracted jobs. With alimonies there is no change as against the previous year because although they come in as money transfers (questionnaire B), they are classified as Other incomes.

2.2.4 Housing costs

In the case of more than one household in one dwelling unit, the costs were divided according to their actual contribution to their financing.

When the household reported its housing costs only in one item as the rent paid for accommodation, the partial amounts were estimated based on the data from households, which provided the detailed information on their housing costs. Estimates were modelled by regression

models taking into account the type of dwelling (family houses vs. other), type of rent (market rent vs. regulated rent contracts), number of household members and usual local level of housing costs (municipality, census enumeration unit).

2.3 Data tables - description and notes

The publication contains data tables for households (table 1 to 13) and for persons aged 16+ (tables 14 to 17). Table 18 presents at-risk-of-poverty rates broken down by various characteristics together with some other Laeken indicators.

Values in the tables were calculated from the weighted microdata and rounded. The total counts of households or persons may therefore not always exactly correspond to the sum of the counts for a given breakdown. For the same reasons, the sum of percentages may not always be equal to 100.

Whenever the term "children" is used in the table headings, it always means dependent children (in accordance with the state social security law), except of tables 9 and 18, where the EU definition of dependent children is used. Net household income was used for all classifications based on income.

The publication tables with data on households were designed in line with the publications from previous income surveys (Microcensus 1996 and 2002, Social situation of households survey 2001, and previous LC surveys) to enable comparisons over a longer span of time. For this reason, the income definition used in tables is the national net household income definition. The tables' legends clearly show for what population (or subpopulations) of households the results are calculated. The set of tables is divided into five parts labelled by name and letter a) to e), each part with its own legend:

a) Basic data on household composition and income. Income data are mainly per capita averages; in selected tables the presented income data are equivalised using the EU (modified OECD) equivalence scale. Presented averages of consumption units enable users to re-calculate the equivalised income based estimates also for other breakdowns. Table one contains more detailed breakdown of income, other tables are restricted to only main income components.

b) Income distribution of households and persons for fixed income groups. The income group's brackets remained the same as in previous LC surveys. Incomes are further related to minimum income level in national legislation and to the median of per capita income calculated from income distribution of persons in all households. This part is not included in tables, where income per capita was used as classification criterion.

c) Characteristics of households, which describe their structure according to various classifications and which supplement or explain data on income.

d) Characteristics of housing of given household groups, equipment with selected consumer durables and housing costs, which are presented as monthly averages per household.

e) Subjective opinions of households on their housing, financial problems - for example in connection with the housing costs, repayment of loans and ability to make ends meet.

2.3.1 Notes to selected tables with household data

Table 1 gives data for household groups, which are comparable in long time series. It offers a look at changes in household structures, their demographic characteristics and incomes.

Tables 2 to 4 - households total by decile distribution base on net money income per capita and EU equivalence scale, households of employees and household of pensioners by quintile distribution based on net money income per capita. The households were ordered and divided into deciles/quintiles according to net per capita household income, or net equivalised household income using the EU (modified OECD) equivalence scale. The values of deciles and quintiles are incomes of the last household in that quantile group. While grossing up the survey data, it is not possible to maintain exactly the same number of households in each quantile group. Therefore, the household counts may slightly differ.

Table 5 is the result of comparison of the monthly net household incomes with their subsistence minimum from the national law on social need (as of 2006). The multiplying coefficients were chosen with respect to the entitlement to social benefits.

Tables 6 and 7 comprise information on households broken down by number of dependent children and number of household members at work

Table 8 presents a breakdown of childless households by at-work status of their members

Table 9 - the classification using EU-compatible typology enables international comparisons. Among others, households at risk of poverty are classified according to this typology. Because of the change in definition of the dependent child (the age until which the person is considered as a child unconditionallly rose from 16 to 17) the time series of some of the indicators might be broken.

Table 10 - size of municipality, as of December 31st, 2006 from demographic statistics.

Table 13 - type of household and education. Only households where the head of household is economically active were included. In two-parent family, the education of the head of household is combined with the education of his spouse. Some low frequency combinations are omitted. Primary education includes secondary-vocational education and persons with incomplete primary education.

2.3.2 Notes to tables with data on persons 16+

Tables 15 to 17- persons are classified according to the demographic characteristics and the size of the municipality where they live. In addition to the presented basic economic activity variables, the prevailing part of the table presents the data on subjective evaluation of health. This part does not include proxy respondents (respondents, for whom the questionnaire data was collected from another household member) - since proxy answers were not allowed for this part of the personal questionnaire. The percentages for reasons why there was an unmet need of medical care are calculated only for the subsets of respondents, where this situation occurred.

Table 18 - based on the EU harmonised methodology, the calculation of the at-risk-of poverty rate is based on the equivalised disposable income (see the definitions part for differences from the national net income definition). Disposable household income is equivalised by dividing by the number of consumption units (modified OECD equivalence scale). Calculated equivalised household income is then assigned to all household members (as a result, all household members have the same equivalised income value). Based on this income distribution of individuals, poverty line is defined as 60 percent of the median equivalised income. Alternative values of 40, 50 and 70 percent of the median are used for comparison as supplementary poverty lines. At-risk-of poverty rate is then expressed as the percentage of persons (in the total population or in a given subpopulation - by gender, age, economic activity) with their assigned equivalised income below the chosen poverty line. This harmonised methodology is the foundation for needed international comparisons between the EU countries.

Prevailing economic activity in this table corresponds to definition already mentioned in chapter 2.1. For inclusion in this part of the table, the activity must last at least 7 months of the reference year 2006. Persons not fulfilling this condition were not included in this calculation.

Since 2006 the changed Laeken indicators have necessitated a change in definition of dependent child. Dependent children are now all persons aged 0-17 and those persons 18-24 who are economically inactive and live with at least one parent. As a consequence, the time series of indicators affected by this change is broken.

The table is supplemented by selected other indicators of income poverty, which characterize in more detail the variability of income and give further, more detailed, information on poverty.

Quintile share ratio S80/S20 - is the ratio between the sum of equivalised income of the 20 percent of percent with the highest income (fifth quintile) and the sum of equivalised income of the 20 percent of persons with the lowest income (first quintile). Higher values of this ratio mean higher differentiation of incomes.

Relative at-risk-of poverty gap - is the difference between the median income of those under the poverty line and the value of the poverty line expressed as % of this value of the poverty line. Higher value of this indicator means deeper fall of persons under the poverty line. This indicator was calculated for the poverty line at 60 percent of the median.

Gini coefficient - is calculated from the ordered distribution of equivalised income. It reflects the relationship between the cumulative proportions of persons and cumulative proportions of income. Its values have the range of 0 to 1. The higher is the value, the higher is the income inequality. It is usually presented in publications in percents.

3. Results accuracy

When interpreting and analysing the results of the Living Conditions survey, one has to keep in mind the fact that the results are based on survey data only and subsequently inferred to the whole population. It means that all published data are but statistical estimates based on the survey sample comprising possible sampling and non-sampling errors.

The non-sampling error occurs in all surveys and censuses as well. It might come to existence as a consequence of many reasons, mostly inaccurate methodological instructions, not respecting them by interviewers, wrong wording of questions, processing mistakes, non-readiness to participate in the survey or giving purposely biased answers. Due to meticulous care in all phases of data collection and processing one can reduce this component of total bias significantly. However it is difficult or nearly impossible to evaluate its influence on the results. Providing a good definition of auxiliary variables one can compare their distributions in the sample with the known distribution in the whole population (census).

The sampling error is the consequence of processing the results of not all units of the population, but of sample data only. One has to infer the figures for the whole population from the obtained survey results. It can be evaluated using the sampling survey theory. This error can be limited by choosing a sample, which is large enough and representative. Also other factors can influence the sampling error: sampling design, incidence of measured variable and its natural variance.

Due to increasing of the sample size in 2007 the majority of published estimates turned out to be more accurate. However, the relatively low readiness of households to participate remains a problem. This results (especially in the case of repeated visits in the panel) in narrower range of types of household in the primary data collected. These data are the input for the statistical processing and evaluation, and this increasing bias is corrected by calibration techniques described in chapter 1.5.

3.1 Estimates of sampling errors, confidence intervals

There are two ways to evaluate the sampling error: either by a point-estimate of variance or by a confidence interval for the observed variable. Mostly 95% confidence intervals are constructed by multiplying the standard error by quantile of normal distribution - 1,96. It says within which interval the measured variable is going to lie with the probability of 95%. In this publication the measured variables are either frequencies - both relative and absolute - of how many households carry a certain feature, or means/totals of incomes.

Sample survey theory distinguishes between two types of totals - population totals and sub-sample totals. The sub-samples are the results of applying various criteria to the whole population, like dividing the whole population into specific household groups according to head of household's economic activity.

The standard error is computed for each estimate separately. Computing standard errors of percentage totals or relative occurrence is the easiest. Relative occurrence can mean e.g. number of households of self-employed members and their percentage of all households. In the case of other estimates (e.g. income totals and their means per household or per capita) one must compute the standard error directly from the primary data and for each sub-sample separately. The tables illustrate the volatility of variability of various indicators, different sub-samples and several types of income.

3.2 Confidence intervals for frequencies

The following two formulas are simplified approximations of exact formulas and are applicable only to variables with binomial distribution. This is the case of estimating frequencies of household characteristics, like percentage of incomplete families. The deviations between the approximations and exact formulas in such cases are statistically significant. However the formula for sub-population totals (onward characteristic A) might give inexact results for small area estimates. Therefore the values in the upper left corner of Table II were omitted.

Both formulas can be used as a guide for computation of confidence interval of random variables with binomial distribution:

a) for the population total

95% confidence interval of estimate $Y_A = y_A \mp 1{,}96 . s_{yA}$, where

$$s_{yA} \cong N \cdot \sqrt{(1-f).\frac{\frac{y_A}{N}.(1-\frac{y_A}{N})}{f.N}} \qquad (1a)$$

and N is the population size,
f is the relative sample size (n / N),
y_A is the estimate of total Y_A of characteristic A in the population

Note: In the case of estimating confidence interval of relative frequency, one should substitute the ratio $\frac{y_A}{N}$ in the numerator with this relative frequency.

b) for the sub-population total (of observed characteristic B on the set of A)

95% confidence interval of estimate $Y_{AB} = y_{AB} \mp 1{,}96 . s_{yAB}$, where

$$s_{yAB} \cong y_A \cdot \sqrt{(1-f).\frac{\frac{y_{AB}}{y_A}.(1-\frac{y_{AB}}{y_A})}{f.y_A}} \qquad (1b)$$

and y_A is the estimate of total Y_A of characteristic A in the population,
f is the relative sample size (n / N),
y_{AB} is the estimate of total Y_{AB} of characteristic B on the set of A.

Note: One can substitute the ratio $\frac{y_{AB}}{y_A}$ in the numerator of the ratio again with respective relative frequency of characteristic B on the set of A.

How to use the attached tables for determining frequency confidence intervals

Table I Estimates of 95 % confidence intervals of population totals for households and persons in the Czech Republic

The table serves to determine an approximate 95 % confidence interval of population totals of frequencies from the set of households or from the set of persons on the level of the Czech Republic as a whole.
Let us take an example. In "Table 1 - Households by social group" we find an estimate of 188,6 thousand households of the unemployed, and want to know the confidence of this estimate. So we look up in Table I in the column "Households - estimate - thousands" the row closest to the number 189, namely 180. In this row we find the particular confidence interval, which in this case amounts to ± 16,6 thousand, for relative frequency the confidence interval is 4,45 ± 0,41 %. One could still refine the estimate using simple linear interpolation.

Table II Estimates of 95 % confidence intervals of subpopulation totals for households

The table serves to determine an approximate 95 % confidence interval of subpopulation totals of frequencies from the set of households at the level of the Czech Republic as a whole. So provided we want to find out the confidence of estimate of frequency of lone-parent families in unemployed households, which was 25,5 % of 188,6 thousand, we will look up in table II the closest row to the number 189, namely 180 again and the column closest to the number 25,5, namely 25. The confidence interval for the relative frequency amounts to 25,5 ± 4,08 %. Again. one can use linear interpolation to further refine the interval.

3.3 Confidence intervals in general

If the variable is not distributed binomially, one cannot apply the previously mentioned approximation, but has to compute the standard error directly from the individual data. As we estimate averages or totals, we can apply the central limit theorem and determine an α% confidence interval for the estimate *h* of the characteristic *H* using this formula:

$$h \mp u_{1-\alpha/2} . s_h, \qquad (2a)$$

where h is the estimate of characteristic H,
s_h is the standard error of the estimate h
and $u_{1-\alpha/2}$ is the quantile of normal distribution.

Confidence intervals for average income per capita

We collect data about incomes for the whole household. Therefore the average income per capita is computed as ratio of 2 random variables *y* - total of incomes and *x* - total of persons. Provided simple random sampling without replacement applies and we weight the sample data with weights *w,* one can determine the confidence interval using this formula:

$$\frac{y_w}{x_w} \pm \frac{u_{1-\alpha/2}}{x_w} \sqrt{\left(1-\frac{n}{N}\right)\frac{n}{n-1}\frac{n}{\sum_{i=1}^{n} w_i}\sum_{i=1}^{n}\left[w_i\left(y_i-\frac{y_w}{x_w}x_i\right)^2\right]} \qquad (2b)$$

where $u_{1-\alpha/2}$ is the quantile of normal distribution (in our case 1,96),

n the sample size,

and $x_w (y_w)$ are weighted sample totals $x_w = \frac{n}{\sum w_i}\sum_{i=1}^{n} w_i x_i$ resp. $y_w = \frac{n}{\sum w_i}\sum_{i=1}^{n} w_i y_i$

Although computed confidence intervals in tables III, IV and V were based on this formula assuming simple random sample, the influence of ***design effect*** was additionally taken into account. Simplified, it is the influence of complicated sampling scheme on the variability of estimated characteristic compared to the same result assuming simple random sampling. In reality, as previously described, the sample was stratified at the level of NUTS3 and 4 size-groups of municipalities and was carried out in two stages (see chapter 1.1).

Generally the design effect is quantified in compliance with this formula:

$$\text{deff (h)} = s_h^2 / s_h^2\{\text{srs}\}, \qquad (3)$$

where s_h^2 is the variance of variable *h* at the real sampling design

and $s_h^2\{srs\}$ is the variance of variable *h* at simple random sample.

It is known from the theory that stratification decreases variance, whereas multistage sampling causes estimates with equal observations to be less efficient. Due to higher total number of dwelling units selected, also (both relatively and absolutely) more CEUs (census enumeration districts) were included. The influence of the abovementioned deff therefore decreased in accordance with expectation, and its values for not rarely occurring income categories for the whole Czech Republic varied between 1,0 and 1,3.
E.g. in the case of net monetary income per capita in the Czech Republic the design effect equalled 1,10.

A modification of formulas (2a) and (2b) was used to compute values in tables III, IV and V. Total variability was in each case decomposed to its components corresponding to each sampling stage.

x x x

In this publication, it was possible to bring out only some results of the third survey on incomes and living conditions in the CR dubbed "Living conditions, 2007". The data collected in the survey make it possible to publish various other breakdowns that are not included in this output. For further information contact Information Services - tel. +420 274 052 304 or e-mail address infoservis@czso.cz