INTRODUCTION

PURPOSE OF THE SURVEY

The Statistical Office conducts the Household Income and Expenditure Survey (HIES) regularly since July 1995.

The purposes of the survey are:
- *—— to get reliable information about the economic situation of households,*
- *—— to calculate indicators reflecting socio-economic development of the society (standard of living, cost of living, etc.),*
- *—— to project socio-political measures,*
- *—— to estimate the effectiveness of socio-political measures.*

The methodology of the survey, which was worked out by the scientists of the Family Laboratory and the Institute of Mathematical Statistics of Tartu University, is close to the methodology of surveys conducted in other European countries.

In addition to income and expenditure, the survey provides information about the housing conditions of households, about the availability of durable goods (technical items) in the household and additional sources of income (e.g. services used under favourable conditions, possibilities to use agricultural land and food products grown on that land). It was also surveyed how the households estimated their economic possibilities and how large income the households would like to receive in order to manage comfortably without excessive luxury.

The present yearbook summarizes the data collected by the survey during 1997. The user will find from the book information on the households average monthly income and expenditure in the country as a whole, in urban and rural areas, by regions and by different types of households. There is also information about the quantities and average prices of foodstuffs bought and on the consumption of self-produced or free foodstuffs. The summaries of answers given to the questionnaire "Economic Situation of the Household" have been published (in percentages).

Due to rounding, the column sums are not always equal with the total in the households income and expenditure tables. The difference can be up to some last decimal places.

SAMPLING DESIGN AND ESTIMATION OF PARAMETERS

The size of the sample in one quarter of 1997 was 4,800 households. For obtaining quarterly regional information, we use an additional sample from Hiiumaa (90 households in each quarter).

All households who live officially and permanently in Estonia are considered as the population to be studied. Persons who live in institutional households are excluded.

The frame is the population database of the Ltd Andmevara. As there exists undercoverage in the frame among the younger age groups we presume that persons who are younger than 16 years do not form households of their own. Systematic random sampling is used among the people who are 16 or older at the beginning of the surveyed year. As the people in population register appear in random order, the systematic sampling could be viewed as a simple random sampling. Each person selected into the sample brings his household to the survey. The person who determines the household included in the sample is called an address person. If the address person is dead, has left Estonia or is living in an institutional household, we consider him as frame overcoverage and exclude him from the survey (together with his/her household).

If the address person does not live at the address given in the population database, the household which actually lives at that address will be surveyed as the replacement household.

The inclusion probability of each person is $\frac{n}{K}$, where K is the number of persons in population who are 16 or older and n is the sample size. As the studied household is determined by the person, the inclusion probability of the household depends on the number of persons in the household who are 16 or older. Household where there are p 16 years old or older members have p times higher probability to be in the sample. The latter is assumed while giving the formulae of estimates.

Since there are errors in the frame, the number of 16-year-old or older persons who live officially in Estonia have been used when finding the estimates. On 1 January of 1997 that number was 1,151,970 persons (data

source is the Statistical Office of Estonia). Estimation functions and methods how to eliminate bias caused by nonresponse and frame errors have also been provided.

Notations

y_i — *total of surveyed variable in the i-th household,*

n — *number of households in the sample (the frame error excluded),*

m — *number of respondent households,*

N — *total number of households in the population,*

L — *number of people in the population,*

L_d — *number of people in the d-th population domain,*

p_i — *number of 16-year-old or older members in the i-th household,*

β_i — *number of members in the i-th household,*

K_g — *number of 16-year-old or older people in the g-th strata,*

m_g — *number of respondent households in the g-th strata,*

β_i^l — *number of children in the i-th household.*

Response rate $\dfrac{m}{n} \cdot 100\,\%$.

We find estimates to the population parameters using Horvitz-Thompson estimator. Bias from nonresponse and frame error is compensated by weighting with the help of four post strata: Tallinn, Hiiu county, other Estonia, which is divided into two: urban and rural.

Used estimation functions are:

1. Number of households
$$\hat{N} = \sum_{g=1}^{4} \frac{K_g}{m_g} \sum_{i=1}^{m_g} \frac{1}{p_i}.$$

2. Number of people in the population
$$\hat{L} = \sum_{g=1}^{4} \frac{K_g}{m_g} \sum_{i=1}^{m_g} \frac{\beta_i}{p_i}.$$

3. Average size of the household
$$\frac{\hat{L}}{\hat{N}}.$$

4. Number of households in domain
$$\hat{L}_d = \sum_{g=1}^{4} \frac{K_g}{m_g} \sum_{i=1}^{m_{dg}} \frac{1}{p_i}.$$

5. Number of people in domain
$$\hat{N}_d = \sum_{g=1}^{4} \frac{K_g}{m_g} \sum_{i=1}^{m_{dg}} \frac{\beta_i}{p_i}.$$

6. Average size of the household in domain
$$\overline{\beta}_d = \frac{\hat{L}_d}{\hat{N}_d}.$$

7. Average number of children per household in domain
$$\overline{\beta}_d^l = \frac{\hat{L}_d^l}{\hat{N}_d} \ , \ where \ \hat{L}_d^l = \sum_{g=1}^{4} \frac{K_g}{m_g} \sum_{i=1}^{m_{dg}} \frac{\beta_i^l}{p_i}.$$

8. The estimation of study variable average per household member

$$\hat{\overline{Y}} = \frac{\hat{Y}}{\hat{L}} = \frac{\displaystyle\sum_{g=1}^{4} \frac{K_g}{m_g} \sum_{i=1}^{m_g} \frac{y_i}{p_i}}{\displaystyle\sum_{g=1}^{4} \frac{K_g}{m_g} \sum_{i=1}^{m_g} \frac{\beta_i}{p_i}}.$$

9. The estimation of study variable average per household member for the domain of the population

$$\hat{\bar{Y}}_d = \frac{\hat{Y}_d}{\hat{L}_d} = \frac{\sum\limits_{g=1}^{4} \frac{K_g}{m_g} \sum\limits_{i=1}^{m_{dg}} \frac{y_i}{p_i}}{\sum\limits_{g=1}^{4} \frac{K_g}{m_g} \sum\limits_{i=1}^{m_{dg}} \frac{\beta_i}{p_i}}.$$

10. Distribution of households by income deciles.

For every household, the income per one member of the household ($\frac{y_i}{\beta_i}$, where y_i is the total income of the i-th household) is calculated. The households together with their weights ($w_i = \frac{K_g}{m_g p_i}$) are set in ascending order by the income per one household member, and according to this order new serial numbers $k = 1,\dots,n$ are given to households. The households for which $\frac{(f-1)\hat{N}}{10} < \sum\limits_{j=1}^{k} w_j \leq \frac{f \cdot \hat{N}}{10}$ belong to the income decile $f \in \{1,\dots,10\}$.

In the same way as income deciles, the expenditure deciles were formed.

Since the household fills in the diary book of food expenditure for half a month, thus, in order to estimate the food expenditure we use coefficient which depends on the number of days in the month (how many days are in the first part of the month and how many in the second part).

The result of the year is the mean of quarterly results (exc. deciles where it is considered that some households are in the sample more than once).

RELIABILITY OF ESTIMATES

The results of a sample survey are subject to several types of errors.

Estimates always contain a sampling error, which is caused by the fact that the survey is based on a sample instead of a complete census of the population. The degree of the sampling variance depends on the size of the sample, the length of the period for recording data, and the way in which the sample has been selected. Comparing the estimates of the two parameters of the population, it cannot be asserted that the value of one parameter is bigger or smaller than that of the other one, without having checked the statistical importance of the difference. It is possible to estimate sampling variance using data from the monthly bulletin "Estonian Statistics" No 12, 1997, because we know that when the sample size increases k times the sampling error decreases approximately $\sqrt{k}$ times.

The sampling error estimates also contain bias caused by nonresponse, frame error, measurement error and processing error. In the Estonian Household Budget Survey the biggest problem is the bias caused by nonresponse (average response rate in 1997 was 55.3). Those who did not respond generally differ from those who did. Thus, the generalisations about the population contain a certain error.

THE SYSTEM OF DATA COLLECTION

The data are collected from households by face-to face interviews and by the diary questionnaire method.

*1. **Household Picture (Household Picture — Changes)** —filled in by the interviewer at the first meeting with the household. At the second and third visit to the household the interviewer fills in the "Household Picture — Changes", where changes in the household which have taken place during three months are recorded.*

Household Picture contains general data about the size and composition of the household (number of members, age, gender, relationship to the head of the household, nationality, educational level, ability to work, relation to work, area of occupation, form of ownership of enterprise, social group and group of occupation of household members).

Part A of the Household Picture contains data about the economic situation of the household: housing conditions, estimation of the sufficiency of the household's income, possibilities of using free services, use of land, purchasing power of the household, ownership of technical items (durable goods) and desirable income needed in order to manage comfortably without luxury.

*2. **The diary book of food expenditure** contains data about the food expenditure of the household during half a month (half of the sample fills in the diary in the first half of the month — from 1st to 15th day — and half of the sample does it in the second half of the month — from 16th to 30th (31st) day of the month). The cases of eating out and the consumption of self-produced or free food products will be registered as well.*

*3. **The diary book of income, taxes and expenditure** contains data about monetary and non-monetary income of the surveyed month, taxes paid by the household and goods and services bought by the household. The separate parts in the diary are for recording the expenditure on construction and renovation activities and expenditure related to the household's economic and production activities and lastly expenditure on the investment of money and other transactions (depositing, borrowing, lending of money, etc).*

ARRANGEMENT OF THE SURVEY

In order to carry out the Household Income and Expenditure Survey, the Statistical Office uses its own interviewers. The length of one survey period is one month (12 months during the year). The key person in the organisation of the survey is the interviewer. About 160 specially trained household interviewers are permanently involved in the survey. The work of interviewers is coordinated by 18 supervisors from the Interviewers Set Department of the Statistical Office.

The interviewer is a person, who
− *contacts the household and convinces it to take part in the survey (by letter),*
− *makes a face-to-face introductory interview (Household Picture) with the household,*
− *advises the household how to fill in the diary books and makes arrangements for further contacts with the household,*
− *collects the filled diary books after the survey period and carries out the first checking of diaries with the household.*

It is the responsibility of the interviewer that the correctly filled household pictures and other diaries would be handed over on time to the county supervisor. Each interviewer will also fill in the report form on the work done, where he/she records all visits to households, the reasons of non-response and grades the correctness of filling in of the diary books by households.

DATA INPUT, PROCESSING AND PUBLISHING

The filled diary books and household pictures are checked in Tallinn by the Section of Living Conditions Survey and Analysis.

The food expenditure diaries and income, taxes and expenditure diaries will be coded in the section according to a special coding system. The process of data entering is centralized. The data input and processing programme is FoxPro. The data editing includes the first logical control. Before processing of tables the database is additionally checked in the Section of Living Conditions Survey and Analysis. The data are published each quarter after the last survey month in the monthly bulletin "Estonian Statistics", annual data are published in the "Statistical Yearbook of Estonia" and in the annual publication "Household's Income and Expenditure".

THE HIES CONDUCTED BY THE STATISTICAL OFFICE OF ESTONIA IN THE CONTEXT OF EUROPEAN HIESs

By the sampling method, the HIES carried out in Estonia by the Statistical Office is well suited to be compared with the HIESs carried out in other European countries in 1994, being:
− *near the average by the sample size,*
− *on the better side by the questioning intensity,*
− *the best by the population/studied households ratio,*

- *corresponding to the standards by the sampling design (probabilistic sampling),*
- *below the average, but not among the worst countries, by the response rate.*

All the elements used in the sample design in Estonia have been used in other European countries as well. The sample design used in Estonia has the closest resemblance to the sample design used in Sweden.

DEFINITIONS OF TERMS

HOUSEHOLD'S INCOME AND EXPENDITURE

Household *is a group of people living at the same address and sharing joint monetary resources and whose members consider themselves to be members of one household.*

Head of the household *is the member of the household with the largest income (longterm contribution to the household).*

Urban *area includes towns, townships and county towns.*

Rural *area includes villages, separate farms, etc.*

Language spoken at home *is the most currently spoken language in the household (between the household members).*

Revenue net *includes disposable income and saving items — revenues.*

Disposable income (net) *includes monetary and non-monetary net income which is received as earnings from employment, income from self-employment (agricultural and non-farm self-employment), property income, pensions and different social benefits, grants, scholarships and other income (selling of goods, settlements of accounts (taxes), refunded insurance premiums, lottery prizes).*

Income from labour *includes received earnings from employment (wages and salaries, advance payments and premiums), holiday compensations and compensations for business trip without income tax.*

Income from self-employment *includes income from agricultural and forestry activity (monetary and non-monetary) and income from non-farm self-employment. Calculations of income from self-employment include only net income — the current expenditure is deducted. The estimate is negative, if the expenditures of the current month are bigger than income.*

> **Income from agricultural and forestry activity** *includes income from the sale of horticultural and agricultural products, as well as from the sale of livestock, poultry, hunting and fishery among this the cost of self-produced food is taken into consideration. Expenditure on agricultural production is deducted from the sum.*

> **Income from non-farm self-employment** *includes income from registered self-employment, royalties, and income from the sale of self-produced food, goods and services. Expenditure on the corresponding activity is deducted from the sum.*

Property income *includes income from the rent of real estate, income from interests, income from dividends and income from intellectual property (copyright etc.).*

Transfers *include pensions, unemployment benefits, child benefits, sick payments, alimonies, social support, grants and gifts of money.*

Other income *includes income from the sale of personal goods, personal income tax returned, settlement of other taxes, refunded insurance premiums and lottery prizes.*

Non-monetary income *includes income paid for labour or goods and services received as a gift and calculated into monetary value.*

Saving items-revenues include:

> **Financial items-revenues,** *such as the use of personal deposits and income from the sale of securities, loans taken and repayments of loans by a private person.*

Non-financial items-revenues, *such as income from the sale of non-productive real estate, income from the sale of agricultural stock and income from the sale of land.*

Outgoing net *includes all monetary and non-monetary expenditure and saving items-outgoing.*

Expenditure *includes monetary and non-monetary consumption expenditure and other expenditure (alimonies, maintenance costs, life insurance, property insurance, pension insurance and other kind of insurance, fines, gifts, etc.).*

Consumption expenditure *includes monetary and non-monetary expenditure per household member which are connected with consumption (food, clothing, dwelling, health, education, etc.).*

Other expenditure *includes various monetary expenditures such as alimonies, maintenance costs, life insurance, property insurance, pension insurance and other kinds of insurance, fines, gifts, etc.*

Saving items-outgoing *include:*

Financial items-outgoing, *such as deposits paid in and the purchase of securities, repayment of credits and loans and giving of loans and credits to private persons;*

Non-financial items-outgoing, *such as the purchase and extensive repairs of non-productive real estate, the purchase and extensive repairs of agricultural stock and the purchase of land.*