

---

# USING SHORT-TERM STATISTICS FOR ANALYSES IN THE REPUBLIC OF MOLDAVIA

**PhD Professor Constantin ANGHELACHE**

*Academy of Economic Studies/ "Artifex" University of Bucharest*

**Ph.D. Student Oleg CARA**

*National Bureau of Statistics of the Republic of Moldova*

## **Abstract**

*This article outlines the authors' preoccupation for the research of short-term statistics particularities. There are defined: the short-term statistics, the main characteristics of infra-annual statistical researches, the implementation of integrated short-term statistical research, aspects regarding the application of the sampling method in the realization of short-term integrated research.*

**Key words:** *research, principles, harmonization, sampling, typical units, atypical units.*

\*\*\*

In the Republic of Moldova, **short-term statistics (STS)** were and still are destined to measure the infra-annual economic evolution (with monthly and/or quarterly frequency) on national plan, but also regional, including a set of indicators, such as: turnover's index, index of production, and price indexes in industry and constructions, index of new orders in industry, effective number and salary incomes of employed persons, working hours etc. STS covers four major economical areas: industrial, constructions, retail trade and other non-financial services oriented towards the market (transport, information and communications etc.)

Until 2011, short-term statistics in the Republic of Moldova were based on infra-annual statistical researches (monthly, quarterly) of the enterprises, traditionally realized and separately developed, inside branch-oriented statistics (statistics of: labor market, industry, constructions and investments in fixed capital, trade and market services etc.).

Among the main characteristics of infra-annual statistical researches regarding STS, we can refer the following:

- **The functional principle of the organization of researches**, which assumes the collection of information from entities without taking into account their main activity branch. I.e., if the enterprise with main activity

---

in processing industry, also runs secondary activity genders, that is trade and services destined to the market, the reporting unit had to submit to the statistical organs, monthly, statistical questionnaires regarding: industrial activity; trade; services delivered to the population; and also a series of questionnaires that do not take into account the main activity branch of an enterprise regarding the statistics of labor, investments, etc.). There were multiple statistical questionnaires the statistical respondent was to present to the statistical organ during the year.

- **The exhaustive principle of collecting short-time statistics**, according to which most short-term statistical criteria were based on the complete study of statistical unit, without taking their size into account. Statistical researches were characterized by a great number of statistical respondents, studied infra-annually. As an example, in the quarterly statistical research no. 5 - „Consumption and expenses of the enterprise”, there were included some 23 thousand enterprises, versus 16 thousands studied every year within the Annual Structural Survey, which in fact is considered the most extended statistical research in enterprises, with a periodicity that is below one year.

- **The existence of double information in various statistical questionnaires.** The same indicators were found in more statistical questionnaires, and there were inconsistencies in their reporting. I.e., the indicators related to the number of personnel were in six questionnaires with monthly and quarterly frequency. Also, the concept used varied in different questionnaires (example: the average number of employees, the number of employees, the recorded average number of employees for the calculation of productivity of labor, the recorded average number of employees occupied within the main activity), fact which caused difficulties to the respondent in reporting of the indicator and negatively influenced the quality of primary data.

- **The existence, in the result of mentioned causes, but also of the different presentation terms of statistical questionnaires, of the high information-related pressure on respondents.**

The criteria and reporting mode by units for some researches were not strictly defined. In a series of statistical researches, respondent units were both the enterprises and the local units (structural sub-divisions in various areas of the country). There were cases of submission to the Statistical Office of the questionnaires: in a **centralized** manner, by the mother-company (which included the economic activity of all sub-division, regardless in which areas/territorial administrative units – TAU they were), or **de-centralized**, by the structural sub-divisions, which had local units in various territorial

---

administrative units, presenting the statistical questionnaire to the territorial statistical office where they activated. There were cases of under-reporting or doubling the collected statistical information.

The respondents circle for statistical researches was, in many cases, established directly by the territorial statistical organisms (OTS) and was lowly managed by the Central Statistical Office. The National Registry of Statistical Units was not fully used in short-term statistical researches. The lists of statistical respondents, created by the territorial statistical organisms were based on the circle of statistical units, which traditionally presented the statistical information. Due to the limited use of the National Registry of Statistical Units there were, also, cases when, for the same enterprise, in various statistical researches, different data regarding the main activity branch, the property form, organizational-judicial form were observed, affecting the quality of dis-aggregated statistics on respective criteria.

Despite the development of national short-term statistics in the process of transition towards the market economy, there were important blind areas regarding the system of indicators and the calculation methodology compared to international standards, especially the European Union's.

There were blind spots in the research of market services, especially the ones provided to the enterprises, the research of industrial enterprises' orders. The problem can be solved both by financial restrictions, human restriction, lack of necessary expertise, and by the "skill" of many statistical users, especially public administration organs, to operate with traditionally produced statistical indicators. On one hand, short-term statistics were not fully relevant for the market economy and did not provide their international comparability and, on the other hand, the statistical organ allocated considerable resources to produce "traditional" data with reduced information value under the new economic conditions. Meanwhile, we appreciate that the national statistical system of a country must organize its activity, to realize statistical researches according to the informational needs of national users, but also international (by virtue of state obligations as member of international organizations, for the information of potential investors etc.), according to international standards, to ensure high quality, including the international comparability of statistical information.

The processes of collecting, processing and storage of statistics resulted from short-term statistics (as in the case of other enterprise researches) for each questionnaire were separated ("furnace" type), not being correlated and, in most cases, based on old software applications, that did not allowed the application of advanced methods and were extremely rigid, any modification of the questionnaire and processing software was generating enormous difficulties.

---

To ensure the international comparability and internal coherence of the short-term statistics, methods and organization mode for these researches, it occurred the necessity to improve the short-term statistics, to create a new modern and efficient short-term statistics system, as the system would allow

- Harmonization of short-term statistics with the European Union specialized standards, especially adaptation to the requirements of EU regulations no. 1165/1998 and 1158/2005, with subsequent modifications;

- The optimization of the program for STS researches, being eliminated the collection of the same indicators;
- Implementation of sample-based methods for the organization of short-term statistical researches/data collection; consolidation of activity and centralization of sampling and statistical inference works;
- Decreasing the information pressure on the economic agents and the optimization of deadlines for presenting short-term statistical reports (questionnaires);
- Improvement of the Registry of Statistical Units and its transformation in the unique source for establishing the ring of statistical units included into the research;
- Assurance of standardization for processes, joint technical infrastructure etc.

In the results of the examination of EU regulations regarding the STS, the relevant experience of other states, especially: France, Romania, Norway, Baltic Countries, in 2010 were examined methodological, technical and logistical solutions which would ensure the viability and efficiency of the **new system of short-term statistics**, both from the NBS and respondents viewpoint.

By taking into account the advantages of holistic approach in the organization of statistical researches, priority launched at the level of European Union, and its practical realizations, in the field of short-term statistics realized by National Bureau of Statistics through the implementation of the „UNICA” survey, the experience of this country proved to be very relevant. **At the design of the new system of STS the example of the new Romanian model was chosen.**

The complexity of the task of reforming short-term statistical researches and creation of „integrated research” has necessitated an absolutely new conceptual approach and, respectively, the realization of some important methodological, organizational and technical activities.

We consider that it should be taken into account the need to: reconsider the system of indicators, definitions, calculation mode, to ensure

---

harmonization with EU standards. The indicators that are no longer actually, with a low information value were to be excluded and replaced by new ones, in compliance with EU requirements for producing short-term statistical indicators in the fields of industry, constructions, commerce and services, personnel and salary gains; clear presentation to respondents of the indicators in questionnaires and indications to complete them, by taking into account the provisions of the accounting standards; revision of the principle of organizing statistical researches/data collection, the functional principle being replaced by the institutional one, that is according to the principal activity branch of enterprises/units; collection of infra-annual statistical data was to be organized only from enterprises, not from the local units (as it was done in the case of some short-term statistical studies); establishment by NBS, in a centralized manner, of the list of units subjected to research, based on the National Registry of Statistical Units; wide-scale implementation of the sample method in the research of small-size units (the main criteria set being the effective number of employees); optimization of the processes regarding data collection, processing and storage, with the elaboration of respective software solutions, elaboration of applications to ensure information processing and storage based on modern IT solution, which would allow integration in the future improved information system.

In the result of preparation works realized, in January 2011 was launched the **Integrated Short-Term Statistical Research** (integrated research, or CITS), on a monthly basis.

Within the implementation of CITS, the integration from the methodological and organizational viewpoint was made for eleven statistical questionnaires, from industry, constructions, trade and services, salaries and labor market, which were realized separately inside several directions for statistical production of the NBS.

The integrated statistical research comprises four monthly specific statistical questionnaires, depending on the main activity field of the enterprises (regarding short-term indicators in: industry; constructions; trade, non-financial services for the market, agriculture; salary gains in the units of other economical sectors) and two quarterly surveys (regarding consumption, expenses and investments of the enterprise).

In order to compile the macroeconomic indicators according to National Account System within the infra-annual statistics, the quarterly research regarding incomes, consumptions and expenses of the enterprise was kept.

The questionnaires include common chapters for all reporting units, regardless their main activity branch: data regarding the number of occupied

---

persons (monthly); data regarding retribution of labor (monthly); data related to the turnover (monthly), with the exception of the units, for which the indicator is not relevant (budget units, financial institutions etc); data related to investments in fixed capital (quarterly).

Monthly questionnaires include also specific modules, depending on the main activity branch of the enterprise, regarding: industrial activities; activities in constructions.

During the stage of questionnaire design, specialists in National Accounting Standards were invited to join. The questionnaires, and also the methodological indications regarding its completion, were elaborated, being observed, from the peculiarities of the national accounting evidence and laws, to facilitate the understanding and completion of questionnaires by respondents.

The test of the statistical toolbox for the integrated research, realized in enterprises that run various types of activities (processing industry, constructions, engross and retail trade, market services) and various sizes (big, medium and small enterprises), allowed the identification of problems related to the completion of questionnaires and adjustment.

The result of the implementation of designated research, the **Short-term statistical research**, since 2011, distinguishes itself through:

Organization of information collection from statistical units is made on the basis of the institutional principle – depending on their main activity gender. The reporting entity presents only one monthly statistical questionnaire and a quarterly one.

The statistical research was realized on a survey basis and includes some ten thousand reporting units (two or three times less than the most extended infra-annual research no. 5 C existing until 2010).

Statistical units were included in the **exhaustively** integrated research (enterprises with at least twenty employees on average, budget institutions being exhaustively researched, regardless their number of employees) or **sample-based** (enterprises with an average number of employees between 4 and 19);

Enterprises with less than 4 employees were not included in statistical researches during the year.

The responding unit was clearly defines and is the enterprise, data are reported on the whole unit, including structural sub-divisions (subsidiaries, production units, own stores etc.). Thus, there are emphasized situation of double information collection or data sub-reporting in the case of incomplete collection from local units.

New terms were established, extended, that allow the presentation of qualitatively improved data by entities. The integrated short-term research is

---

presented by respondent units at territorial statistical organisms (OTS) on a monthly basis – until 15<sup>th</sup> and quarterly – until 25<sup>th</sup> after the reference period.

The circle of respondents for each monthly or quarterly questionnaire was established by the central office of NBS, based on the National Registry of Statistical Units. There was strictly monitored the situation regarding the statistics of reporting, regarding the statistical respondents.

The implementation of the research allowed the harmonization of definitions and adaptation of short-term statistics to the requirements of the European Union.

There was launched the elaboration of new methodologies for statistical indicators, on a monthly basis, regarding: **turnover** with the assurance of the joint approach, harmonized, data being collected from all units for which it was relevant, including the agricultural enterprises and non-financial market services companies; the **nominal turnover index** for enterprises with main activity in retail trade, engross trade, market services; **real turnover index** for enterprises in retail trade; **value of new orders (contracts)** for industrial units; **salary gains and number of persons** occupied on homogenous activities; **labor expenses** etc.

As a result of CITS implementation, there was a considerable decrease of informational pressure on small units, especially micro-enterprises, and also the expenses related to reporting by companies. The efficiency of statistical organs activity increased, the main focus being put on quality assurance for primary data.

The launch of the integrated research needed a modern joint technical infrastructure. In collaboration with NBS specialists, a reliable information system was created for introduction, processing and storage of data in a single database.

Other details regarding the application of the **sampling method in the integrated short-term statistics** is presented below.

The applied sampling plan is the optimal **stratified sample**, and the selection method used is the **non-recurring random simple selection**, each unit in the layer has the same probability to be included in the sample.

The reference population, which made the object of the research, is the collectivity of active enterprises (units with turnover above zero or at least one employee), regardless the activity gender, according to CAEM Rev.1 (excepting the sections P „Services for private households by employed personnel and private households activity regarding the production of goods and services for own consumption” and Q „Activities of extra-territorial organizations”), which act in the Republic of Moldova, excluding Transnistria. The contents of the sample basis was according to the requirements of the sample, without situations of over or under coverage for some activities.

---

The selection unit was the enterprise.

The layer variables were: economic activity – at level of division for CAEM Rev.1 and the class regarding the size depending on the number of employees.

The layers were formed on the intersection of layer variables. In each layer, the selection method used was the non-recurring random simple selection, each unit in the layer having the same probability to be included in the sample.

The sample for 2011 was around 10 th. units. The inquiry sample ensured the representative data at the level of national economy (division level CAEM Rev.1 and size classes 0 - 9, 10 – 19 and 20 and more employees).

The determination of the sample was made under the condition that the maximum admissible error of the estimation – turnover to be at most 5%, the results guarantee them with a probability of 95%, after the relationship:

$$n = \frac{z^2 \cdot \bar{\sigma}^2}{\Delta_x^2 + \frac{z^2 \cdot \bar{\sigma}^2}{N}}, \quad (1)$$

where:

$z$  – value of the trust quotient determined depending on the probability at which results are guaranteed, in the tables of Laplace function;

$$\bar{\sigma}^2 - \text{intra-groups variation} \left( \bar{\sigma}^2 = \frac{\sum \sigma_i^2 \cdot n_i}{\sum n_i} \right); \quad (2)$$

$\Delta_x$  - limited error of the selection average

$$\Delta_x = z \cdot \sigma_{\bar{x}} = z \cdot \sqrt{\frac{\bar{\sigma}^2}{n} \left( 1 - \frac{n}{N} \right)}; \quad (3)$$

$N$  – population (total collectivity).

The allocation of the measure of the layer sample was made according to the Neyman method, which takes into consideration the value of standard deviation of the turnover within the layer, with the help of the relationship:



---


$$n_i = \frac{N_i \sigma_i}{\sum_{i=1}^k N_i \sigma_i} n \quad (4)$$

where:

$n_i$  - size of the sample in layer  $i$ ;  
 $N_i$  - general collective population in layer  $i$ ;  
 $\sigma_i$  - average square deviation of the turnover in layer  $i$ ;  
 $n$  - the volume of the sample calculated in the previous step.

Resulting from the need to ensure the compatibility of time series, it has been decided that the exhaustive research limit to be fixed at 20 persons. It was decided that all budget units to be researched exhaustively, regardless their size. Also, were subjected to exhaustive research: all commercial banks (group 651, CAEM Rev.1); all insurance units (66 division, CAEM Rev.1); all units created in 2010, whose statutory fund is at least 0,5 mil. lei.

If, after formation of layers, there were situations in which the effective in a certain cell was below 10 units, units within layers were included in an exhaustive manner (for 2012 the minimal size of the sample was 5 units).

Together with the units that satisfied the exhaustive criteria mentioned, there was a category of units exhaustively researched – atypical units. As atypical units, in literature, we understand statistical units part of a certain collectivity, but whose characteristics(one or more) are essentially different than the ones of most units. This fact led to the presence of a high not justifiable variability in the population, which forced the extraction of an ineffective sample due to this fact, respective units were treated in a distinguished manner from the others, in order to avoid the achievement of some erroneous results. In our case, the identification of atypical units was made on turnover at layer level.

The notion of atypical unit is highly relative, so literature indicates more methods to identify it. In our approach, we followed the two most used methods, based on the three sigma and the inter-quartile interval.

By using the method of the three sigma (standard deviation), were considered as atypical units the ones whose turnover was higher than the average turnover, on stratification cells, with +/-3 standard deviations.

In the present case, the CAEM classification was taken, division level (2 or 3 signs ) in intersection with 2 groups of employees 0-9 and 10-19 and the standardization was made on the formula:

---


$$z_{is} = \frac{x_{is} - \bar{x}_s}{\sigma_{xs}}, \quad (5)$$

where:

$z_{is}$  – variable  $z$  for unit  $i$  in layer  $s$ ;

$x_{is}$  – turnover of unit  $i$  in layer  $s$ ;

$\bar{x}_s$  – average turnover in layer  $s$ ;

$\sigma_{xs}$  – standard deviation of turnover in layer  $s$ ;

Units whose a score  $z_{is}$  is above  $\pm 3$  were considered atypical, being researched exhaustively.

The method of inter-quartile interval for the identification of atypical (outlier) units assumed the comparison of each unit, on the value of the turnover, with two values resulted from the following relationship:

$$\text{Outlier} < q_1 - 1.5(q_3 - q_1)$$

$$\text{Outlier} > q_3 + 1.5(q_3 - q_1) \quad (6)$$

where:

$q_1$  – quartile I, value until which 25% from the effective of the researched collectivity is placed;

$q_3$  – quartile III, value until which 75% from the effective of the researched collectivity is placed;

The location of the quartiles was determined with the relationship:

$$L_p = (N + 1) \frac{P}{100}, \quad (7)$$

-  $N$  size of the collectivity in a certain layer, and  $P$  takes the value 25 for  $q_1$  and 75 for  $q_3$ .

All units that have satisfied the conditions were treated as atypical units, being exhaustively included in the research.

To increase the degree of homogeneity for the collectivity subjected to sampling research, two variants were examined for the determination of atypical units, on: layers mad of only CAEM divisions; layers formed at the intersection of CAEM divisions with the groups of employees 0-9 and 10-19.

In case of both variants, two methods were tested:

- Method of determination of atypical units by the rule  $\bar{x} \pm 3 \cdot \sigma$ ;
- Method of determination of atypical units with the help of the inter-quartile interval.

---

Following the analysis, the following data resulted, on the four variants:

**Distribution of statistical units in economic field on types, depending on the variant approached and applied method**

	Variants			
	1a	2a	1b	2b
Total (active units)	26898	26898	26898	26898
Exhaustive ring	4086	4086	4086	4086
Atypical units	222	252	1547	1458
Collectivities by sampling	14997	14967	13672	13761
Informative: total (active and inactive units)	46203	46203	46203	46203

It results that in the case of using the method of the three sigma, the number of atypical units is small for both variants (222 and 252), and the method of the inter-quartile interval emphasizes a great number of atypical units (1547 and 1458). The collectivity subjected to research had a low degree of homogeneity and, subsequently, it was mandatory to use some robust indicators of variance, the inter-quartile deviance being one of the indicators.

To increase the efficiency of the sample and avoid the inclusion in the sample of very small legal units (from the perspective of the number of employees), that, among with the fact that were hard to contact, had also a very low weight in the total population, units with less than four employees were not included in the sample. According to European good practices, statistical units kept in the sampling base must ensure a representative character of the turnover of at least 90% at the level of each CAEM division (for activities for which the turnover was available).

Before the real extraction of the sample, several simulation were performed, with the purpose to compare the variants received and to chose the optimum variant.

**Volume of the sample depending on approach variant and method applied**

Collectivity by sample	Variants			
	1a	2a	1b	2b
Sample	14997	14967	13672	13761
Sample fraction	4587	7932	2759	2992
Sample + atypical	0.31	0.53	0.20	0.22
Sample + atypical + exhaustive	4809	8184	4306	4450
Sub-population by sample	8895	12270	8392	8536

The most elevated efficiency, both from the viewpoint of sampling fraction, and of the total number of economic units subjected to research, was achieved for variant 1b, where only CAEM divisions were used as layers to determine atypical enterprises, by applying the inter-quartile interval at the identification of atypical units. In this variant, the highest number of atypical units was also achieved, the appropriate mandatory conclusion being the fact that the elimination of atypical units leads to the homogenization of the collectivity. In this case, the number of units in the economic branch that were to be researched every month during the integrated research for 2011 was 8392.

The basic expansion quotients, or the selection weight ( $w_{ih}^b$ ), a Hortvitz-Thompson weight, is calculated as the reverse of the selection probability, on the relation:

$$w_{ih}^b = \frac{1}{p_{ih}} = \frac{1}{\frac{n_h}{N_h}} = \frac{N_h}{n_h} \quad (8)$$

where:

$p_{ih}$  = selection probability of unit i from selection cell h;

$N_h$  = number of units from selection cell h;

$n_h$  = number of units selected in the sample from selection cell h.

The adjustment of the expansion quotients to the non-response ratio was made at the level of each selection cell, in the relationship:

$$w_h^{aj} = w_h^b \cdot \frac{1}{p_{rh}} = w_h^b \cdot \frac{1}{\frac{m_h}{n_h}} = w_h^b \cdot \frac{n_h}{m_h} = w_h^b \cdot \frac{n_h - c_h^4 - c_h^6 - c_h^7}{n_h - c_h^2 - c_h^3 - c_h^4 - c_h^6 - c_h^7} \quad (9)$$

where:

$p_{rh}$  = response probability in the selection cell h;

---

$n_h$  = number of units selected in the sample from the selection cell h;  
 $m_h$  = number of units selected in the sample from the selection cell h  
that have responded to the survey;  
 $c_h^l$  - number of units selected in the sample from the selection cell h  
that attached the response status l and  $l = (2,3,4,6,7)$  .

After the adjustment of expansion quotients to non-answers, the atypical units were identified in the stage of statistical inference. This procedure is recommended to be run in this stage too, in order to prevent brusque oscillations of indicators within time series, due to rapid development of some units in the reference period against the situation in the sample base. Atypical values were determined on the same rule as in the sampling stage, that is by using the inter-quartile deviation method.

Depending on values reported by the enterprise, the results of analysis for variation quotients and atypical value markers, the final expansion quotient of the unit was replaced by the value "1".

### Conclusion

**Successful launch of a CITS allowed the achievement of proposed objective, to harmonize short-term statistics**, relevant, to the standards of the European Union, with the increase of efficiency of statistical activity both from the perspective of NBS, and respondents. Harmonized statistical data are elaborated and presented on a regular basis, in compliance with the provisions of PLS

Another important realization of CITS was the efficient improvement of statistic activity. Within the results of implementation for integrated research, the information burden on economic agents reduced averagely by half, especially for representatives of small business.

### Selective bibliography

- Anghelache C. *Tratat de statistică teoretică și economică*. București: Editura Economică, 2008
- Cara O., Mocanu Iu., Galer L. *Abordarea integratoare a statisticii pe termen scurt: experiența Republicii Moldova*. *Revista Română de Statistică*, no. 6, 2012
- Cook L. *Strategic Planning in Statistical Offices – some Insights from the UK Experience*. *Statistics in transition*, 2003, vol. 6, No 3
- Cox D.R., Donnelly C.A. *Principles of Applied Statistics*. Cambridge University Press. Cambridge, 2011
- Isaic-Maniu A., Mitruț C., Voineagu V. *Statistica pentru managementul afacerilor*, București, 1999.
- \*\*\* *Hotărârea Guvernului Republicii Moldova nr. 305 din 14 aprilie 1999 cu privire la Concepția de reformare a sistemului statistic național în anii 1999-2002*, *Monitorul Oficial al Republicii Moldova nr.42-44 din 29.04.1999*.