

Joint OECD-Eurostat-Russian Statistical State Committee Workshop on Measurement of the Non-Observed Economy

Hosted by Russian State Statistical Committee

SOCHI, RUSSIAN FEDERATION ,16 - 20 October 2000

**EMPIRICAL PROCEDURE IN MEASURING NON-OBSERVED ECONOMY IN
GDP ESTIMATES**

**Dr. Nikos Mylonas,
Director of National Accounts of the NSSG,**

.

EMPIRICAL PROCEDURE IN MEASURING NON-OBSERVED ECONOMY IN GDP ESTIMATES

Dr. Nikos Mylonas, NSSG

The task to achieve exhaustiveness in the calculation of the national accounts system following the international standards in particular with the SNA 93 or ESA 95 by including non-observed economy (NOE) in the GDP estimates is a difficult area and important issue for all national statistical offices. The NOE comprises production activities that are non market economic activities (such as home production), and legal market activities that are kept hidden for reasons such as tax- evasion while they should be measured by the statistical system. In the attempt to include NOE in the national accounts system a number of the following steps of actions should be undertaken, as for instance:

Action 1: ► refers to the information needed for understanding the strengths and weaknesses of the results of sources and methods used in the compilation of the national accounts system.

In order to suggest possible practical ways of experiences on undertaking checks, comparison, and to make corrections and adjustments in the survey results and in the system of national accounts generally, with the aim to include also NOE we have to understand the conceptual framework of the existing statistical system of the country concerned. This can be done, by collecting information using the questionnaire given below, on the basic characteristics of each survey of the statistical sources. Examining carefully this information, the strengths and weaknesses of the sources and methods used in the national statistical system could be identified.

Questionnaire about the strengths and the weaknesses of the sources and methods used in the statistical system:

Name of survey:

Link to surveys undertaken at the European level (e.g. structural business statistics):

Reporting units (e.g. enterprise/ local KAU/ household):

Periodicity (e.g. annual/quarterly/other- to be specified):

Time of availability of results (e.g. 18 months after the end of the survey period):

Sampling frame: (e.g. name of business register used/ population census):

Survey is compulsory or voluntary?

Main features of survey methodology (e.g. PPS sampling/ panel of respondents/ use of a size threshold for sampling/ postal questionnaire/ telephone interview):

Population size:

Sample size:

Survey response rate:

Method used to impute for missing data:

Variable used for grossing-up to the population (e.g. turnover/ employment):

Sample coverage, as % in terms of variable used for grossing-up (e.g. sample covers 60% of employment recorded on the sampling frame):

Main variables collected:

Further adjustments made to the survey data:

For other data sources which are important for compiling GNP, the following information should be provided:

Name of data source:

Organisation collecting the data, and purposes for which it is collected:

Reporting units:

Periodicity:

Variables collected:

Methods used to allow for missing data:

Adjustments made for conceptual differences from national accounts concepts:

Further adjustments made to the data:

As already stated above, this information is very important for the evaluation of the results of each survey also to propose corrections for measurement of the NOE.

Action 2: ► refers to a minimum of information that should be collected by the statistical sample surveys conducted on the productive activities of the small industry.

Generally, the statistical sources used in national accounts are distinguished according to the approach they are collected to the following categories:

- Data which are collected by special basic census or sampling quarterly or annual surveys from households and from enterprises or from the establishments (in the case they keep separate accounts).
- Data which are based on direct information collected from all public organizations of productive activities and Ministries etc. and
- Data based on educated information and technological coefficients derived from special studies and cooperation with specialists.

In national accounts we are interested to provide independently, reliable and exhaustively collected data from production, the expenditure, the Price-quantity and the income approach statistics.

A survey on production statistics should include the following characteristics:

- Complete coverage of large enterprises (using a very detailed questionnaire, such as the type A questionnaire of the annual industrial survey of the NSSG) plus stratified sampling with the lowest stratum representative of the smallest business (using for instance, a concise questionnaire of type B).
- The use of production survey enables the variables (gross output, intermediate consumption) required for estimating value added to be directly collected according to national accounts definitions. This means for example that intermediate consumption can directly measured rather than estimated.

The questionnaire of type B which is referring to the lowest stratum should collect at least the following variables:

- ⇒ Number of employees
- ⇒ Number of working proprietors (self employed)
- ⇒ Gross output and sales in detail
- ⇒ Compensation of employees
- ⇒ Gross wages and salaries
- ⇒ Intermediate consumption of good and services (in detail to the main input components, if it is possible)
- ⇒ Gross fixed capital formation
- ⇒ Changes in stocks

The collection of statistical data through a special basic census or sampling quarterly or annual survey from households and from enterprises or from the establishments is based on the maintenance of a well informed business register.

Action 3: ► refers to the conduction of a survey of a well informed business registers with the aim to use it for the measurement of the NOE.

Data collection of production statistics surveys may be affected from underreporting data, from non responses to the survey, from overestimation on the inputs etc. Missing data or non reliable information or underreporting of data especially, in the case of sample may affect the quality and reliability of the expansion results. With the task to improve the reliability in national accounts a minimum of information apart from the information about the identity of the enterprises (and establishments), that is the activity, the address, the legal status etc., data on employment, revenues and VAT on output and in inputs should be included.

A register survey is considered to include a wide range of different source data on producer units that can be unified into a single coherent business register offering the following main characteristic information:

A. For enterprises

- Enterprise number
- Name
- Address(mailing address)
- Telephone number
- E-mail
- Ownership(individuals, limited partnership, governmental, municipal etc.)
- Municipality code
- Address code
- Activity code according to NACE Rev. 1(Principal code and code for secondary activities)
- VAT and tax number
- Group of enterprise number
- Number of employees
- Turnover
- VAT on output and }
- VAT on inputs } or the VAT rate for the principal activity

B. For establishment

- Name
- Address(where the activity take place)
- Telephone number
- Address code
- Ownership(individuals, limited partnership, governmental, municipal etc.)
- Municipality code
- Establishment number
- Enterprise number
- Number of employees
- Activity code(Principal code and code for secondary activities)
- Turnover

With the aim to improve the quality in the survey results, we need more or less these information, as already mentioned above, as a minimum information, from an well informed business register for the following reasons:

- To use the file of addresses for the enterprises and establishments with the task to conduct in the case of a census covering all production units,
- to design the sampling frame of the production statistics surveys and
- to check the monitoring of an establishment for possible movements from one 4-digit NACE code to another or from one stratum to another.
- to cross-check the data completed by the enterprises production survey statistics with those included in the business.

In conducting a well informed register special attention should be paid to the quality of data on turnover, on the employment number and on the correct classification of a production unit according to its main production to a 4-digit code number classification of the NACE rev.1.

More specifically,

Action 4 : ► refers to the use of the registers-file of enterprises and establishment for checking the exhaustively of data in the production surveys results.

First of all, it is necessary, a cross-checking and comparison of the data, of each the production unit collected by the production statistics survey and by the well informed business register as to the following variables:

- the employment data,
- the activity, and
- the turnover of the production unit.

In some cases of production activities, such as hotel industry the register may include additional variables, so that a stratification could be based according to the number of beds, and/or to the overnights, per category of hotel, separately for the number of hotels operating on an annual on a seasonal basis, by each type of establishment.

A well informed register is updated taking into account the new establishments on the basis of the TAXIS system and/or of the issued licenses by the Ministry of Industry, Social security funds or any other relevant Ministry or organization. The results on employment and on turnover on a production units basis and on a stratum basis are cross-checked with the corresponding data derived from the registers and with fiscal audit data, that is VAT or tax number archive in order to undertake the appropriate corrections in the level of output and inputs.

From business registers data-file of the total “population” of the enterprises and establishments a special stratification may be provided by presenting the number of enterprises (and/or establishments) and the number of employment or the size of turnover per stratum on a 4-digit NACE (Rev.1) classification according to the stratification used in the production statistics survey for each branch.

This kind of stratified information of a register at the reference year, is important, as stated above, for the following reasons:

1. For the comparison of the data concerning revenues, employment and VAT payments of a firm (or an establishment) as given by the register, with the relevant data given by the production statistics survey.
2. For the design of sampling frame of the production statistics surveys and the checking of monitoring of an establishment for possible movements from one 4-digit NACE code to another or from one stratum to another.

A stratification of the number of production units per by industry helps to the design of sample survey and the expansion of its results in a more reliable way.

The NSSG conducts annually industrial surveys on a census basis on all industry activities from III to VII stratum and on a 4-digit code classification of the NACE (rev.1).

The small strata I employing 0 to 4 persons and stratum II employing 5-9 persons are surveyed on a sample basis. The sample of enterprises is derived from the list of enterprises (or establishments) included in the register. As already stated, for these establishments a simplified questionnaire –type B – is used, while for the firms with an employment of more than 10 persons a very detailed questionnaire of the type A is foreseen.

Checking the differences between the primary frame of the survey and the final results by stratum.

In the next stage the sampling results mentioned above are extrapolated to the total population by each stratum. For the final approval of the results of the surveys some corrections are made by the Division of Methodology which checks by stratum the differences between the primary frame of the survey and the final results and also makes all the necessary adjustments according to the international practice. Mainly this is the case for the establishments non-responding to the survey. Consequently, the framework of the survey and the extrapolation of the final results, depend on the establishment of a very well-informed register.

Action 5: ► refers to the adjustments made by the National Accounts Division as to include the second (or third job employment) and to compiled annually employment matrix for each industry .

The difference in employment data between the Labour Force Survey employment matrix data of each industry and the corresponding employment data given by the production statistics surveys by industry provide us with the so called “extra employment”. For this extra employment an extra output and also inputs are calculated in addition to those given by the production statistics in the small stratum I and II of the manufacturing industries. These estimations of output and also inputs are based on the productivity per head or the technical coefficients of the surveyed industry concerned.

Diagram 1, presented below, shows a working table as example for the calculations procedure followed in the production accounts by industry utilizing all kind of available data sources.

Action 6: ► refers to the adjustments made by the National Accounts Division as to adjust output of some industries as to the amount of non-deductible VAT which was found as the difference between the theoretical VAT and empirical(or received) VAT. This calculations were only made for the case of the base year.

In some countries the VAT data which are declared by the establishments to the register and to the questionnaires of the survey are cross-checked. This cross-checking is carried out on the basis of the nominal VAT rates per each product estimating the so called “theoretical” VAT. This is the VAT which should have been paid to the state if there wasn’t any tax evasion.

Comparing this theoretical imputed result with the relevant value of the VAT which is declared from the enterprises to the production statistics and the VAT declarations, an extra output (and inputs) is estimated reflecting the amount of tax

evasion. Nevertheless there is the problem of the fraud VAT with agreement or not of the purchaser. In the first case no corrections to the output should be made.

Diagram 1: Working table for the calculation of the extra output in the production accounts by industry.

Industry STATISTICAL SOURCES ⇒ ⇓Data	Exh aus tive So urc es	Exp and ed So urc es	Com plem entar y Sou rces	Statistic al Declarat ion	Output Other Ind.	Rem aind er	TOT AL
Secondary output of other industries + main production of this industry	////	////	////////	////////	0	////////	0
MAIN production	0	0	0	0	////////	0	0
Secondary Production	0	0	0	0	////////	0	0
Total intermediate inputs	0	0	0	0	////////	0	0
Compensation of employees	0	0	0	0		0	
Wages and salaries	0	0	0	0	////////	0	
Actual social contributions	0	0	0	0	////////	0	
Imputed social contributions	0	0	0	0	////////	0	
Other taxes net linked to production	0	0	0	0	////////	0	
Gross operating surplus	0	0	0	0	////////	0	
Employed persons	0	0	0	0	////////	0	
Employees	0	0	0	0	////////	0	
Self-employed	0	0	0	0	////////	0	
Unpaid family workers	0	0	0	0	////////	0	
Total employed persons	0	0	0	0	////////	0	
Number of establishments	0	0	0	0	////////	0	
VAT invoiced on main production	0	0	0	0	////////	0	

Action 7: ► refers to the presentation of methods used for the integration of the production and the expenditure results on the macro or the product level and the need of calculating GDP by independent sources.

There are three basic approaches of calculating GDP , and although they use different variables of the economic system, they have to give the same result for GDP if each component (or item) is estimated correctly by an independent way .

These three approaches are:

- **The output approach or production statistics**, which is based on the calculation of production and intermediate consumption(IC) of the various industries of the economy using the results from the production statistics data.

These production statistics give us gross value added and thus the contribution to GDP . by the following equation:

Production (at basic prices) - IC(at purchasers' prices) = Gross value added(GVA) (at basic prices) + taxes on products – subsidies on products = GDP.

- **The expenditure approach** which is based on estimates of the components of final demand:

GDP = FC(Final consumption at purchasers' prices) + GFCF(Gross fixed capital formation at purchasers' prices) ±J(Changes in stocks at purchasers' prices) + E(Exports fob) – M (Imports cif).

- **The income approach** is based on the separate estimates of the components of GVA by industry.

GDP = Copensation of employees + Operating surplus/mixed income + taxes on production and imports – Subsidies.

Accordingly the best results for GDP estimates are obtained if, for each of the items of the three approaches estimates, from independent sources are available. However, in practice, several items are estimated from the same sources, making the estimates less independent. For example, the estimates of household consumption are often, for a large part, based on production data. This is reducing the independence of the production and the expenditure approach. Also, estimates of the operating surplus of the industries are often based on data on production, intermediate inputs and compensation of employees. In that case , the production and income methods are not independent either.

In answering the question what is the best estimate of GDP and other national accounts variables, we have explained that this depends on the quality of the sources underlying the variables in each of the three approaches.

Action 8: ► refers to the choice of integration method of the three approaches of estimating GDP with the aim to measure in the final synthesis procedure of the system remaining still additional NOE activities.

Another issue is related to the choice of integration method of the three approaches of estimating GDP, since there are two methods of integration of the three approaches mentioned above:

The one is:

- **The integration method at the macro level** when the results of the three approaches, comparison of the sources of the different approaches. As a result, this method does mentioned above, are only compared at the latest stage. In this method there is a limited not guarantee a fully consistent system, and causes errors that are not determined. This method, which is used in several countries, integrates the results of the three methods at the macro level for calculating national accounts data on GDP.

The other is:

- **The integration method at the product level** when the results of the three approaches, mentioned above, are compared at a detailed product level, as it is done in a supply/use framework. The key characteristic of this system is a common product classification for all purposes (i.e. production, intermediate consumption, household consumption, capital formation, foreign trade, etc). The product classification enables an analysis at the detail level where the supply and use of each product group should be equal.

According to this method of integration there is only one GDP, resulting from the system that is made consistent on the most detailed level. From the point of view of the reliability and exhaustiveness of national accounts' estimates, the method of integration at the product level is clearly favorable.

Using a synthesis system such as the supply/use framework makes it possible:

- to compare estimates for the same variables resulting from different sources on a detail basis,
- to identify gaps in the available sources and
- to make consistent estimates by residual for variables for which no data are available.

What's even more important:

Such checks and comparisons can be made in a very systematic way, product by product and industry by industry. Thus the effects of changes made to some variable on other variables are immediately clear.

Action 9: ► refers to final conclusion of measuring NOE activities by using the supply / use framework.

In some countries the estimates from the income, output and expenditure method are compared more or less only at an aggregate level. In these countries national accounts does not include an integrated supply/use framework. In case the difference between the production and the expenditure method is regarded as too large, an investigation is mounted into the causes of the difference. This leads to adjustments to either the production or the expenditure components until this statistical difference has been reduced to an acceptable level. In the expenditure side as statistical discrepancy is then included this remaining difference. The income approach plays a role mainly only for the non market activities.

In some other such as in our country a working procedure using the supply/use framework (of about 380 homogeneous groups of products by 123 industries) brings the data from various sources together and balances them on a commodity level. This involves comparison of the data and analysis of errors at a detailed level, simultaneously in current and constant prices.

The establishment of the supply and use framework is generally recognized as being the best integration method for the national accounts system, since it improves the reliability and the exhaustiveness of national accounts data. This cannot be done from the one day to the other; it presupposes a special educational program and the incorporation of several transversal tasks of processing the data sources, methods and definitions in the establishing of the supply/use framework for calculating quarterly and annual national accounts data.

In order improve the statistical information in the national accounts system the following special pilot surveys should be conducted:

The attempt to improve the information on statistical sources through special surveys only positively could be assessed. The priority which is given to the work program on the component of these special surveys will assist to cover part of the underground economy activities. This information is needed in order to compare different data from the statistical sources such as:

- **to compare employment data from surveys and the Labor Force Survey(LFS) results,**
- **to include to the results of the LFS an estimate by activity as the second or third job of employment,**
- **to calculate the employment matrix by industry including also the extra employment for which an extra output and intermediate consumption should be calculated,**
- **to check the survey results as to the detail categories of the BOP, of business tax data, of the employment data in services activities or of the structural business statistics and data from special studies,**
- **to exploit information about hidden activities and to deal with the problem of tax evasion or the legal informal economy and other kinds of legal or illegal 'black' or underground services industries of hotels and tips in restaurants etc. , transportation and construction industry,**
- **to examine the possibility to improve progressively the content and quality of national accounts.**

It's well known how difficult is to evaluate some information coming from HBS concerning expenditures such as smoking , drinks, entertainment, payment in kind, tips etc. Thus the existence of independent complementary statistical sources is necessary to improve the quality of the statistical information coming from some times unique statistical source.
