

THE EXHAUSTIVENESS OF PRODUCTION ESTIMATES: NEW CONCEPTS AND METHODOLOGIES

Manlio Calzaroni, Istat Italy
Via Cesare Balbo, 16 – 00184 Rome
calzaron@istat.it

ABSTRACT

In this paper we outline the conceptual framework to be used to verify the exhaustiveness of the production estimates according to the System of National Accounts (SNA, 1993). We describe the types of production units (regular, irregular, informal, not physically identifiable) that we have to investigate and the statistical problems that we have to deal with. As regards the application of this framework, we present the methods of the Italian National Statistical Institute (Istat) used to ensure the exhaustiveness of the GDP estimates.

Key Words: exhaustiveness, input of labour, productive units, non observed economy

1. INTRODUCTION

By "productive units" we refer nowadays to a very large number of different typologies. If our objective is to obtain a complete picture of the volume and characteristics of production activities, i.e. an exhaustive estimate of them, we need to take into account the problems of measuring the activities of all types of units. Furthermore, if our population target is production as defined in the latest version of the System of National Accounts (SNA93), we also need to consider the problems of measuring the production of units that, in order to avoid payment of taxes or social security or because of their economic characteristics, may be invisible to the statistical instruments.

2. DEFINITIONS AND POPULATION TARGET

The SNA93, for the first time, has built a coherent framework for the identification of all the components to include in the National Accounts estimates; therefore, it is possible to capitalise on this considerable work and to adopt at the international level a homogeneous language as a basis for describing and analysing all types of production units and their output.

The foregoing concerns are common to industrialised, transition and developing countries alike, in their efforts to build a more adequate information base for economic analysis.

Because of the experiences developed within the EU to guarantee the "Exhaustiveness of the GDP" we think it is useful to use the description adopted in the European Commission (Decision 94/168/EC, Euratom of 22 February 1994, Article Type 2, Title II, *Definition of Terms*):..... "Within the production boundary, national accounts provide an exhaustive measure of production when they cover production, primary income and expenditures that are directly and **non** directly **observed** in statistical surveys or administrative files".... According to the internationally accepted definition described in the SNA93, the **Non** directly **Observed Economy** (from this decision derive "NOE", drawing this term, as a convenient summary), includes the following "areas": (1) illegal, (2) underground, (3) informal sector.

(1) Illegal activities are classified by the SNA (SNA93: 6.30-6.33) in two categories: (a) the production of goods and services whose sale, distribution or mere possession is forbidden by law (i.e. production and distribution of illegal drugs); (b) production activities which are usually legal but which become illegal when carried out by unauthorised producers (i.e. unlicensed practising of a profession).

Both kinds of production are included within the production boundary of the SNA93, provided that they are genuine processes whose outputs consist of goods and services for which there is an effective market demand.

(2) Underground economy (SNA93: 6.34-6.36) stands for all legal production unknown to public authorities due to various reasons: (a) to avoid the payment of income, value added or other taxes; (b) to avoid the payment of social security contributions; (c) to avoid having to meet legal standards such as: minimum wages, maximum hours, safety standards, etc.; (d) to avoid complying with certain administrative procedures, such as completing statistical questionnaires or other administrative forms.

Therefore, underground activities can be defined either as “**economic underground**”, indicating those activities characterised by the deliberate will not to respect administrative standards, with the purpose of cutting production costs (items (a) (b), (c)), or as “**statistical underground**” indicating those activities that are not surveyed because of inefficiencies in the statistical information system, the characteristics of productive units or the lack of statistical sensitivity on the part that are requested to compile the questionnaires (item (d)).

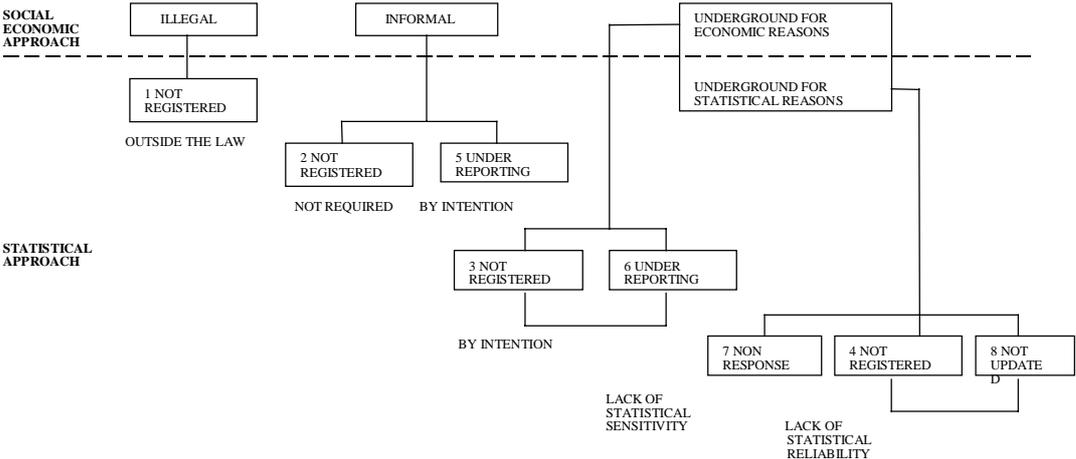
(3) To define the Informal sector, the SNA93 (IV – Annex) refers to institutional production units characterised by: (a) a low level of organisation; (b) little or no division between labour and capital; (c) labour relations based mainly on occasional employment, kinship or personal and social relations, as opposed to formal contracts.

These units belong to the household sector and cannot be associated with other production units. In such units the owner is fully responsible for all financial and non-financial obligations undertaken for the productive activity. There is no need to underline that informal activities are not necessarily carried out in order to evade taxes or other controls related to social security contributions (as mentioned before). On the basis of the laws in force in each country, for example, this sector may be identified by referring to the “size” of the production unit or to the characteristics of the legislation (no obligation whatsoever to register with public authorities).

3. PROBLEMS AFFECTING THE EXHAUSTIVENESS: THE STATISTICAL POINT OF VIEW

From the statistical point of view, measuring NOE is obviously a very difficult endeavour because of the elusive nature of what is being measured and, consequently, the approximations made in the measurement process. Figure 1 shows how the aforementioned phenomena relate to the statistical problems: the origins of such problems, their impact on the statistical system and the different NOE typologies that can be identified

Figure 1 – NON OBSERVED ECONOMY



accordingly. To sum up, from the statistical point of view we have four different problems: Non registered, Underreporting and non updated activities. These issues are then briefly illustrated to highlight the statistical aspects involved in the measurement of the single NOE components.

Non-Registered Activities

The existence of non-registered production activities implies that there are missing institutional units in our data sources (statistical or administrative ones). The most important consequence is the lack of reliable business registers. Let us analyse how this situation arises. The first and simplest reason for non-registered activities is the illegal activity (in fig. 1 this corresponds to NOE type 1 - illegal economy, non registered). Production units in the “informal” sector may be missing when they are not requested to register at all by any kind of legislation (in our typology this corresponds to NOE type 2 - informal sector, non registered). Among the enterprises belonging to the “formal” sector, some units (one enterprise or a part of it) are missing due to the deliberate intention to avoid the existing obligations to register. The main reason for such behaviour is to avoid additional costs of various kinds: value added taxes, social security contributions, costs related to the compliance with health and safety standards, etc. (NOE type 3 - underground economy, non registered). Within the “formal sector”, enterprises may also be missing in our sources due to statistical reasons (independent from their will). This can happen, for example, because of: (1) the great turbulence in the enterprise universe, such as the high rate of turnover of enterprises (considerable where the share of small-sized units is particularly high); (2) the lack of adequate laws about statistics; (3) the lack of efficiency in the statistical system; (4) special laws for specific types of units. The main consequence is the difficulty to maintain reliable business registers and, therefore, the impossibility to guarantee the exhaustiveness of the information derived from business surveys (NOE type 4 - statistical underground, non registered).

Underreporting

This occurs when data provided by a production unit are not consistent or have abnormal characteristic values (i.e. costs/production ratio). In general, this happens because one or more components contributing to make up the production value are underreported or, vice versa, because cost components are over-reported. In all cases, the final result is an underreporting of the enterprise’s profits, highlighted by an underestimate of the value added. Such underreporting can be identified both within the informal sector (when tax returns about such units are available) and within the underground economy (in the first case it is NOE type 5 - underreporting, informal sector; in the second case, it is NOE type 6 - underreporting, underground economy).

Non response

This problem is related to the “statistical sensitivity” of single enterprises, a problem well known to statisticians. Therefore, we will not discuss it here (NOE type 7 - non response, statistical underground).

Lack of updated information

Usually this is due to the fact that the registers (or more generally speaking the population target) used by National Statistical Institutes (NSIs) are not reliable. In general, we refer to problems related to out-of-date information on production units inside statistical or administrative “registers”. The lack of updating can be due to various kinds of changes in the production structure of enterprises, such as: 1) enterprises that do not exist any longer, but are still considered as “active” ones; 2) changes in the structure (mergers, demergers, etc.); 3) changes in size/dimension (in terms of employees or of turnover, etc.); 4) changes in economic activity; 5) changes in the geographical distribution of production units. Obviously, the above mentioned items affect the exhaustiveness and the quality of National Accounts (NA), as they affect both the level of estimates and their analysis by geographical area and by economic activity (NOE Type 8 - statistical underground, non updated).

4. THE ITALIAN APPROACH

The Italian economy is characterised by a strong presence of small productive units, often unrecorded, and a high rate of irregular employment in the labour market. In order to ensure coverage of these two problem areas in GDP estimates, in the 1980s the Istat (Istat, 1993) have developed an original method, the “Input of Labour Approach”. The procedure recommended by the SNA93 as most appropriate to estimate the input of labour in terms of full time equivalence (FTE) is very close to the method used by the Istat. In the Italian methodology FTEs are used as a tool to estimate the total labour input including NOE activities, and to obtain the total estimates of output and value added (by multiplying the number of FTEs by the per capita values of output and value added). In light of this, it can be said that not only is the NOE (excluding illegal) already included in current estimates it constitutes an integral part of all the economic aggregates influenced by it.

As can certainly be seen from the following explanation, brief though it may be, this type of methodology contains numerous advantages such as organic unity and the systematic nature with which the problem of the Non Observed Economy is treated, as well as the replicability resulting from its standardisation.

The techniques used for estimating the production and value added are diversified by branch of economic activity, on the basis of the best results obtainable in exhaustiveness terms:

- A. estimates “quantity × price”, this technique is used for estimating the activities of the agricultural and energy sectors and part of construction;
- B. estimates through expenditure (part of constructions, rents and private services for education and research, health, entertainment and leisure);
- C. estimates through direct gathering of costs and earnings from balance sheets (credit, insurance and some branches mostly belonging to public enterprises);
- D. estimates through distributed incomes (non-market services);
- E. estimates through expansion of per capita values for FTE, after having estimated the overall labour underlying the product and after having corrected the per capita values for possible underreporting (technique “input of labour × average per capita values”, used for estimating all other branches).

These criteria show the fundamental role played by employment in methods which estimate the product from the point of view of formation . Approximately 70% of the value added is estimated with the E technique.

By combining the different estimation techniques, of which the E technique is the most original one, Italian accountants believe that they can cover exhaustively the “non observed economy” linked to legal activities (i.e. non-criminal ones, that EU decided not to estimate for the time being).

To sum up, Italian accountants consider (i) the statistical underground which is due, first of all, to the weight of small enterprises (in fig. 1 statistical underground), (ii) the utilisation of irregular labour within the productive process (economic underground) and (iii) the under-declaration of the production obtained by means of regular labour (economic underground) the major aspects that characterise the Italian NOE reality.

The procedure for estimating the aggregates of NA (such as production, value added, compensation of employees) analysed by branch of economic activity can be summarised with the following formula:

$$Y = \sum_{i=1}^m \sum_{j=1}^8 x_{ij} \cdot U_{ij} + \sum_{i=m+1}^{101} Y_i \quad [1]$$

where:

Y = overall estimate of the aggregate (for example value added)

i = indicator of the branch of economic activity (101 branches are chosen in accordance to the Italian productive system)

j = indicator of the size of the enterprise (1-5 employed, 6-9, 10-19, 20-49, 50-99, 100-249, 250 and over)

x = average per capita value of the aggregate (for example: production or value added per employed)

U = Full Time Equivalence (FTE)

$\sum_{i=m+1}^{101} Y_i$ = part of the aggregate not estimated through the units of labour technique

Input of labour estimates are obtained with the same methodology for all industries. In those branches where the technique “E” is not used, input of labour estimates are applied for coherence controls.

With reference to the consequences which the NOE has on the exhaustiveness of the statistical sources we have to solve three main problems: the non registered activity, non updated information and underreporting. In accordance with the Italian approach to solve them we need an exhaustive estimate of the input of labour (“U” in the formula) for the non registered and non updated components, and an adjustment of the per-capita (“x” in formula) for the underreporting components.

The importance of the ISTAT experience in this field is also demonstrated by Eurostat’s decision to adopt this approach for verifying the exhaustiveness of GDP estimates in the EU. The reasoning behind the methodology developed by ISTAT is fairly simple. If the same sources which are needed to estimate value added in the NA (enterprises’ side) can also be used to yield an estimate of employment, then that employment estimate can also be assessed for completeness against the estimates of employment available from demographic data sources

(households' side). As summarised by Hayes and Lozano (1998), the exercise can be described as a process in four steps: Defining 'employment underlying GDP', so that the link with the variables of production and value added are as straight-forward and as close as possible; Standardising the definition of employment, to compare different sources a meaningful way; Assessing the employment comparisons, for the whole economy or at branch level (the implicit assumption made concerning households is that they have less interest than the enterprises to hide their real working condition, regular or "non regular" though it may be, so that the discrepancies that may emerge can be economically meaningful); Assessing the impact of the resulting best employment estimates on the estimate of value added and hence GDP.

4.1. Estimating non registered and non updated components

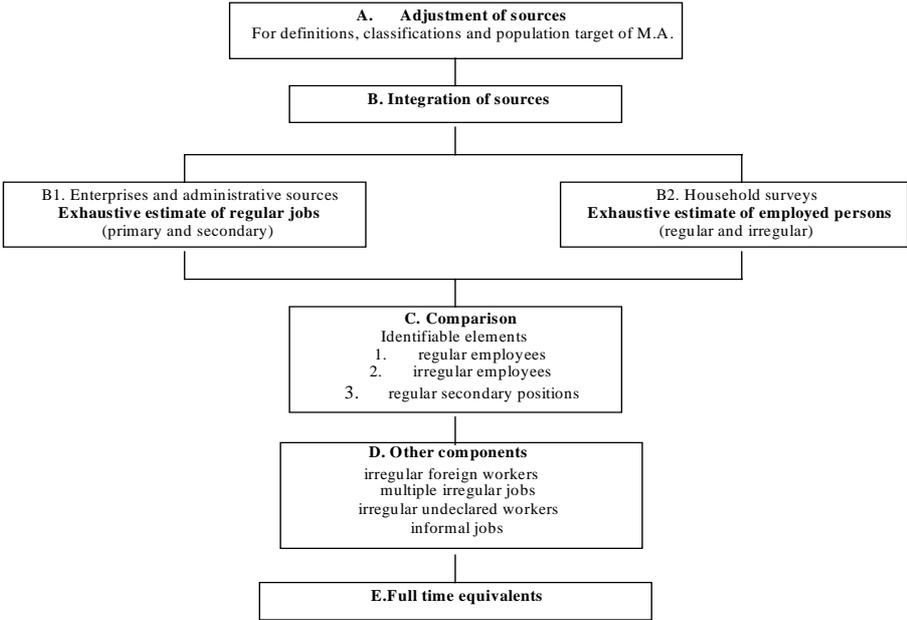
Before moving on to a description of the main methodological steps that have been implemented for obtaining an exhaustive estimate of labour input it is necessary to highlight the different informational capacity that the two survey units (enterprises and households) have, at least in the Italian context. Enterprise data (demand side) provide a measure of regular jobs (both primary and secondary), i.e., they provide information on employment for which legal provisions and obligations are full filled. On the other hand, data collected via households (supply side) measure the number of employed persons, both those who are regularly and irregularly employed. The existence of such a situation is not a hypothesis but rather a reality that has been repeatedly verified by researches carried out by Istat.

The labour input estimation technique foresees the following main steps, in accordance with the fig. 2:

- A. Adjustment of sources.** Temporal and territorial harmonisation; conceptual harmonisation to national accounts definitions (the concept of "domestic employment" as a productive factor of GDP), etc.;
- B. Integration of sources.** On the supply side (B1) and on the demand side (B2) of labour so as to have the most exhaustive measure of employment from the demand side and the supply side derived independently;
- C. Comparison** of sources and quantification of the various segments of employment in terms of "jobs" attributing economic significance to the discrepancies;
- D. Other components;**
- E. Full time equivalence.**

With reference to the five steps mentioned above, only the more complex operations are recalled here, while further in-depth analysis can be found by consulting the relevant text (Calzaroni, 1999).

Fig. 2 Methodological scheme to estimate the input of labour



Integration of sources on enterprises (B.1 in fig.2)

The objective of this phase is to estimate and classify workers identified through informational sources that utilise the enterprise as the survey unit; thereby determining an exhaustive estimate of the “regular” component of employment in terms of both its primary and secondary job. This objective is reached by using a register of enterprises that was constructed on the basis of Regulations defined at the EU level and applied to all member countries (Eurostat, 1999).

The register is based on both administrative sources and data gathered through statistical surveys. The main administrative sources are: fiscal data, social security data for pensions, social security data for insurance and against accidents at work, data on enterprises that are registered with Chambers of Commerce, data on the utilisation of electrical energy and telephone services. The main statistical data sources are: census of agriculture, census of industry and services, surveys on the budgets of enterprises, surveys for the up-dating of the business register, monthly surveys on labour indicators in large enterprises in industry and services.

Enterprises systematically produce administrative acts during their lives: they pay taxes, stipulate telephone and electrical energy services contracts, insure the employees against accidents at work. All the aforementioned administrative acts are potential sources of useful information to describe economic activity from a statistical point of view.

Every administrative body has its own function to collect data and manage the corresponding records, under specific legislation and rules which govern relations between various individuals and between them and the public administration. The administrative body defines, classifies, collects and records information on economic agents and their characteristics that, in the strict sense of word, do not have statistical validity. In other words using administrative data causes a problem for statisticians: the inconsistency of data.

The use of administrative data for statistical purposes imposes the necessity to solve the usual problems of a statistical survey - accuracy, completeness, timeliness – with a new conceptual and methodological approach.

Within a survey, consistency is a problem evaluated ex-ante as well as it is strongly linked to the process of microdata collection and macrodata production. When we want to use data stored in non-statistical (administrative) databases, for which statisticians do not have any control of the production process, the problem of consistency is set in a different context and it is resolvable only ex-post.

The main problem that arises using administrative sources for statistical purposes is to identify the correspondences between the statistical concepts and the administrative rules through which those sources observe the population of reference. It is therefore necessary to handle the administrative sources in order to align them with the statistical concepts and definitions. This is possible if, on the one hand, we have an in depth knowledge of the sources to be used and, on the other, suitable statistical methodologies are available.

It is possible to synthesise the logical process for the use of the information derived from administrative sources according to the following three conceptual steps:

1. definition of the reference conceptual frame: statistical definitions and classifications;
2. knowledge of the observed universe (administrative files) referring to coverage, definition of the units and characters, classification used, time and modalities of updating;
3. identification of the rules to convert administrative data into statistical data.

The use of an exclusively administrative source, for example the fiscal one used in some European countries as basis for statistical registers, could cause serious problems. Referring to a defined statistical universe, the typologies of errors generated in the use of only one administrative source for statistical purposes (described in fig. 1 as NOE due to non registered and non updated problems), can be summarised as follows:

- | | |
|--|--|
| <i>E1 – error of under-recording</i> | a) missing records of legal subjects due to delays, etc...;
b) unrecording of legal subjects not obliged to the registration. |
| <i>E2 – error of over-recording</i> | a) registration of not active legal subjects due to duplications, delays or cessation recording;
b) registration of legal subjects without any enterprise features. |
| <i>E3 – error assignment of characters</i> | a) incorrect recording due to delays in variations acquired or to errors in declarations, in recording, in checking;
b) incorrect recording due to different definitions and classifications. |
| <i>E4 – missing assignment of char.</i> | a) partial or total lack of attribution of a character. |

For the above mentioned reason together with the conceptual steps previously illustrated, it is necessary to develop a further function of "identification of rules for the integration of data coming from more administrative sources". After the implementation of such appropriate statistical methodologies, the integration process is a tool useful to ensure the exhaustiveness of units and of the characters of the units, obtaining in such way a reduction of type E1 and E4 errors. Such process must be less useful for the reduction of over recording errors and of wrong character attribution. In fact, using more sources can cause an increase of type E2 error; while if each source is really and considerably better than the other, further information for imputation of statistical characters would cause troubles. Besides, the presence of unchecked matching procedures among sources could cause record duplications and therefore an overestimation of units and statistical aggregates.

Referring to the formal aspect of the integration process, let x_i represent the real value of the i -th unit related to the attribute X and $x_{i1} \dots x_{ij}, \dots, x_{im}$ are the values recorded in M available sources. The relation between the available and the real values can be described as follows: $x_{ij} = g(x_i, e_j, \varepsilon_{ij})$

where e_j is the error due to the bias (structural errors of type "b" previously described) of j -th source, and ε_{ij} describes the random error (errors type "a").

When the knowledge of the j -th source is completed, it is possible to locate rules which standardise (or harmonise or normalise) the units and the variables of input source in statistical units and variables. So the standardisation function is defined as the following application:

$f_s: X_j \Rightarrow X$ which changes the values $x_j \in X_j$ in values $x_i \in X_i$

In other words, a standardisation rule converts administrative concepts and classifications into statistical ones. This rule, generally deterministic, can be divided into three types:

- *coding rules*: which convert coding (e.g. economic activity, legal form, and location) into statistical classifications (Nace, Nuts, etc.);
- *link rules*: by which the different records corresponding to legal or administrative units in one source can be combined to define one statistical unit (enterprise or local unit) ;
- *conversion rules*: to obtain statistical variables from administrative characters

After the standardisation process the erratic component of the model is reduced to the random error ε_{ij} ; therefore the sources are independent and unbiased random variables with same or at least constant quality in order to adopt procedures appropriate to some experimental frameworks such as the theory of repetition of an experiment. To reduce the systematic errors, produced by administrative functions (an example is the trend of enterprises in trade sectors to be classified, for fiscal facilities, as manufacturing enterprises), because the structure of input data it is not useful to use statistics based on linear functions of available x_{ij} (e.g. the mean) in order to avoid distorted estimations of x_i . The methodologies adopted in the imputation of characters must be based on the concept "choice among alternative values" and not on "combining the available values", when there are more values of an attribute for the same unit (Abbate Garofalo, 1998). This means to build a statistical business register.

Integration of sources on household (B.2. in fig. 2)

The approach is based, as for the previous point, on a micro level comparison between the two primary sources that utilise the household as the survey unit, the Population Census (PC) and the Labour Force Survey (LF). The objective is to verify the degree of compatibility between the various declarations provided by the same person and eventually correcting the discrepancies found, with reference to those variables linked to the "employment" phenomenon.

In reality, one can already find many important references in the literature (Cochran, 1977), that point out how important integration can be to improve the accuracy of the estimates of the variables: a census type survey (PC), characterised by non-casual errors but nonetheless capable of capturing statistically irrelevant phenomena that are difficult to capture and quantify in a sample survey; and a sample survey (LF), aimed at a specific phenomenon, in our case "employment", marked by a level of confidence in terms of its ability to both capture and check the data obtained, making it superior to a census survey, at least from these perspectives.

The adopted methodology foresees multiple phases. Here, however, we will limit the discussion to describing those that are most relevant to guaranteeing the exhaustiveness of the estimates.

The merging of the two surveys at the level of the single unit of analysis (that pertaining on the individual person) occurs via a "key" that includes the following variables: place of residence, civil status, sex, academic degree

held, day, month and year of birth. This procedure allows to obtain a dual declaration: one from the PC and another one from the LF, or, in other words, an “overlapping interview”, even if somewhat different from that generally referred to in the literature given the structural differences that exist between a census and a sample survey. This “overlapping interview” concept is used as an instrument for improving the “quality” of the information gathered, both in terms of the significance that is normally attributed from a strictly statistical point of view as well as with respect to the main NA objective, which is the measurement of those components of employment that are not immediately “visible” and the effective discordance present in the information provided by the interviewee in the two surveys.

At this point, LF interviewees make available an overlapping series of information pertaining to their relationship with the labour market. What should be checked, therefore, is the compatibility of this information in terms of the employment phenomenon that can be obtained from the two surveys and, at the same time, identify those non-compatible elements between them that are useful for evaluating components of regular and non-regular employment (i.e., elements that a single survey does not capture).

To sum up, we define: “strong employment” if the respondents declare themselves working in both surveys; “weak employment” if they declare themselves employed in one survey and unemployed in the other one.

The interpretative hypothesis pertaining to the divergent declarations are the following:

- presence of measurement errors in one of the productive processes of the data, from the time of the interview to the validation and tabulation phase (non-sample errors);
- persons that given their marginal position, with respect to the labour market, are able to provide discordant responses. For example, a different “perception” with regard to their actual situation vis-a-vis the labour market in the two separate moments (the subjectivity of the interviewee).

On the basis of the LF information - which is known to be of higher quality – the employment condition declared therein is the one that is accepted. In fact, we maintain that the employment phenomenon targeted by the LF is not subject to bias attributable to the presence of interviewees that declare themselves as being employed even though they are not, while the opposite situation (employed persons who declare themselves as being unemployed) is acceptable. For example carrying out a work activity that is “non stable” (i.e., one that the interviewee does not view as being the same as “employed” because it is: not ongoing over time, carried out for only a limited number of hours), exists outside of the existing normative framework, not connected with the level and/or typology of the professional skills possessed. Any one of these conditions is sufficient for not declaring oneself as employed. Therefore the mismatch between those that have declared themselves as not being employed in the LF but employed in the PC can be considered as a manifestation of their non-regular status which can be identified in through the integrated utilisation of the two sources.

The quality of census data and the underlying motivations of the respondents do not allow us to attribute in a deterministic manner the employment condition of the individual within the group of persons that declared themselves as being employed in the PC but as not being employed in the LF.

The method utilised for resolving incongruous information is based upon research on a “donor” taken from the matched records containing concordant information (employed or not employed in both surveys). The donor record is selected on the basis of a distance function between the information provided directly by the donor itself through a software developed by Istat (Calzaroni, 1999), which identifies a donor that represents the minimum distance with respect to the record to reconstruct. This distance is calculated with respect to the variables that are, by construction, available for the microdata existing in the original data sources.

Analysis on the records that contain a declaration that differs in terms of occupation has shown that, according to the LF, only 21% were part of the active population (8.8% unemployed and 12.5% looking for their first job) while the remaining 79% was comprised of the inactive population, housewives, retirees and students (the vast are housewives – 39%). As for economic activity, the amount of records to correct is the highest in agriculture (23.2%), in retail trade and public activities (10.4%) and in other services (9.9%); according to professional position, the number of records is highest among the unpaid family workers (15,7%), then among other self-employed persons (10,2%) and, finally, among salaried employees (9,1%). This confirms the hypothesis adopted. In fact, it can be seen that, in the Italian reality, “weak” or “marginal” employment is most predominant in sectors such as agriculture and in some branches of the service sector (as opposed to what occurs in manufactories) and among unpaid family workers, which often have an “informal” relationship with the enterprises not always perceived as a real occupation.

Comparison (C in fig. 2)

The work phases described thus far have permitted us to build a coherent and exhaustive information from both the point of views of the households and the enterprises. The estimation of the components that are most relevant to labour input (approximately 90% of the total) is carried out on the basis of such an information set and, in particular, they are identified from the comparison between households and enterprises data.

The assumption at the basis of the estimates of the different employment segments is that in a condition of unchanged field of observation of phenomena, of equal time and space reference of sources, of absence of use of irregular and multiple jobs, jobs detected by surveying the enterprises should equal the number of persons who declared themselves employed in households surveys. By comparing sources from the side of the labour demand with those from the side of labour supply, at a detailed level of territorial analysis and economic activity, separately for employees, the self employed and the unpaid family workers, the different segments of employment are obtained, depending on whether the number of persons who declared themselves employed is greater or smaller than the jobs declared by the enterprises. The following definitions are used:

- regulars: employed people who equal the number of jobs;
- full-time irregulars: employed people exceeding the number of jobs;
- regulars multiple jobs: jobs exceeding the number of employed people.

Other components of labour input (D in fig. 2)

There remain additional occupational segments to estimate on the basis of existing national accounts frameworks which are estimated outside of this procedure: due to the fact that they are directly captured through specific statistical surveys or because they remain outside of the field of observation and therefore are estimated in an indirect manner. For example: considering irregular multiple jobs, indirect sources are used, collecting data from sources on expenditure side; non resident foreigners are estimated on the basis of information provided by the Ministry of Internal Affairs; “informal” jobs, estimated by using “ad hoc” surveys.

Full Time Equivalents (E in fig. 2)

The employment figure obtained by the process as described thus far is quantified in terms of jobs. Now it is necessary to quantify the volume of work (“U” in formula [1]). A correct measurement of its size must take into account the actual quantity of work provided by all jobs in the entire productive process. The first best solution is to be able to quantify how many hours has been performed; however this is not always possible, therefore we convert jobs in FTEs, as suggested by the SNA93. This is achieved by:

⇒ for those components that provide their work on a less than full time basis (secondary jobs, part time), this pertains to regular, underground and informal employment;

⇒ for those components that provide employment on a full time basis, to take into consideration the effective quantity of work provided (for example, hours unworked due to lack of orders).

The conversion into FTEs is carried out through the estimation of “k” coefficients, defined as the relationship between the quantity of work by someone who occupies a position on other than a full time basis and that provided by someone who is employed full time, obtained on the basis of the available information for each of the components of estimated jobs (hours worked, days billed, per capita turnover by full time employees, etc.).

4.2 The component due to Underreporting

Average per capita values (“x” in formula [1]) to be allocated to the input of labour are estimated through surveys on enterprises budgets. These are conducted yearly and they cover the whole universe of enterprises.

Per capita values are corrected with special control and normalisation procedures that are particularly important as they aim at correcting underreporting of turnover. The hypothesis underlying such corrections is that the income of the self-employed worker of an enterprise should be equal at least to the average wage of the regular employees. The income of self-employed workers is obtained subtracting from the value added of the enterprise the compensations of employees, the capital consumption and other components in accordance to the SNA93 scheme (passive interests, banking expenses, rents, etc.). Therefore when the level of income of self-employed workers is less then the level of compensation of employees, the former is re-evaluated. As a consequence the production and the value added are also adjusted by the same amount.

5. CONCLUSION

On the basis of what has been described, the methodology that has been implemented permits a clear and coherent analysis while applying international definitions of the Non Observed Economy phenomenon.

One of the main aspects to be highlighted is that the presence of data derived from the integration of many administrative and statistical sources, completed with the realisation of the Italian business register, has formalised the adoption of the “productive unit” concept as opposed to the traditional “physical entity” (establishment) approach that was statistically surveyable (using the classic instruments of statistical science). By associating the data with the legal characteristics of the unit which is homogenous with the concepts contained in the new SNA93, albeit not necessarily physically identifiable, i.e., only identifiable with new instruments related to administrative sources. For example, we may cite the difficulty in identifying all of the independent or free lance workers – an increasing proportion of the workforce – and therefore necessitate the use of sources that are able to provide measures relative to non physically identifiable units.

One of the main consequences of this approach is that an exhaustive estimate of regular employment allows us to minimise the statistically non observed component of the economy and thereby more accurately identify the economic underground. Otherwise we might have an erroneous interpretations of the economic conditions that are being described.

The integrated use of numerous sources, aside from improving the exhaustiveness of the estimates, allows a more articulated analysis of employment itself given that the sources often identify particular work relationships and, as such, specific behavioural aspects of the production system. The typology of these relationships has become highly diversified over the course of the 1990's and, as a consequence, so has the demand for information on these typologies (part time, contracted, consultancy, false free-lance, etc.)

Finally, the integration of sources at the microdata level approach, utilised in all cases where it was possible to do so, has allowed us to construct an information system in which employment information serves as the bridge between economic data (provided by the NA aggregates that on the bases of the methodology, are consistent with employment as estimated in NA) and social data derived from the LF and PC surveys.

In general, the possibility of analysing employment in a framework bounded by the NA scheme represents a first step towards an approach that integrates monetary variables included in the NA scheme with underlying nonmonetary variables, that represent one of the most important innovations proposed by the accounting system described in the new versions of the SNA.

6. BIBLIOGRAPHIC REFERENCES

- ABBATE C., GAROFALO G. (1998) "Recent innovation in business register and sample survey on enterprises", IASS/IAOS Conferences, Aguascalientes, Mexico.
- CALZARONI M. (1999), "L'occupazione come strumento per la stima esaustiva del PIL e la misura del sommerso", Istat (forthcoming).
- COCHRAN (1977), "Sampling techniques", Wiley.
- EUROSTAT (1997), proceeding of international workshop: "Use of administrative sources for statistical purposes", Luxembourg.
- EUROSTAT (1999), "Use of administrative sources for business statistics purposes. Handbook of good practices", Luxembourg.
- HAYES K., LOZANO E. (1998), "Validating the exhaustiveness of the GNP estimates of the European Union member countries", IASS/IAOS Conferences, Aguascalientes, Mexico.
- ISTAT (1993), "The underground economy in Italian economic accounts", Annali di Statistica Serie X, vol.2, Roma.
- SNA (1993), System of National Accounts, Commission of the European Communities, IMF, OECD, UN, WB.