

25<sup>th</sup> Meeting of the Wiesbaden Group on Business Registers  
- International Roundtable on Business Survey Frames

Tokyo, 8 – 11 November 2016

*M. Consalvi, G. Garofalo, C. Viviano*  
*Istat*  
*Session No. 2*

*Role of Business Registers*

***The new role of the Business Register within the Integrated System of Registers***

## **1 Introduction**

It is known that the increasing use of administrative data for statistical purposes is relevant for the majority of National Statistical Institutes (NSIs). To maximise the benefit deriving from the huge amount of information available it is necessary to build an integrated system of the administrative sources acquired by NSIs. Considering that administrative data are primarily used for other purposes, their statistical use requires a clear definition of statistical units. In addition, each unit in the system should always be uniquely identified, also over time. This result is best achieved by a unique system of unit identification and integration.

Istat has moved in this direction for years, starting from 2013, by centralising some functions for acquisition, storage, integration and administrative data quality evaluation. In the new system called SIM (Integrated System of Microdata) – built with the aim of supporting the statistical production processes both for social and economic statistics – microdata are integrated and unique identification codes are attributed to: individuals and economic units; places; relationships between individuals and units. The system contains more than 70 administrative sources integrating several hundreds of millions of records and thousands of administrative variables in a unique Data Warehouse.

In later years SIM became the basic infrastructure of the statistical production of Istat, also supplying metadata describing data and processes. Starting from the SIM experience, Istat top management decided to invest in a register-based approach to industrialisation.

A transformation of Istat structure is currently ongoing in order for it to adapt to the rapidly changing external context and be able to continue producing top-quality statistical information. Starting from September 2014, in accordance with the European Statistical System commitment to Vision 2020, Istat has developed a model for defining a new organisational and production structure, implementing its internal standardisation and industrialisation process within the framework of a common Business Architecture.

The current Italian modernization programme is based both on a strong and structured governance and on production lines modelled on registers, where the BR will have an enhanced central role.

The new Integrated System of Registers (SIR<sup>1</sup>) is a core element of this project of modernization.

---

<sup>1</sup> The acronym SIR is derived from its name in Italian “*Sistema Integrato di Registri*”. The same applies for its components described next, RSB “*Registri Statistici di Base*”, RSE “*Registri Statistici Estesi*” and RST “*Registri Statistici Tematici*”.

Primarily it is needed to overcome a stovepipe organization model. Then the new system corresponds to the development of a “multiple” approach in data collection that necessarily have to take into account heterogeneity and variability, even in time, of the sources used for the production of statistical information. Furthermore there is the need for qualitatively increasing the available information, making use of a proper ‘reading’ able to interpret the connections between phenomena, i.e. issues that are at the same time economic and social, or that can be seen from two different perspectives, like work demand and supply.

This results in the need for favouring “process” rather than “thematic” aspects, above all in the context of registers, going towards the deployment of applications and methodologies that are independent from the specificity of information and therefore enable a high generalization.

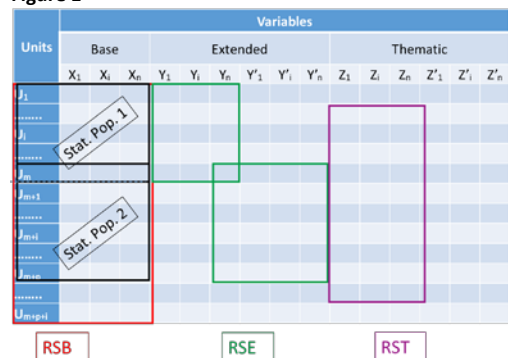
One of the targets is to identify the cross-cutting nature of information: as far as may be possible the integrated system must be thought - in its contents - as the only one supporting more statistical processes.

## 2 An overview of the Integrated System of Registers

The SIR is a single logical and functional environment to support the consistency of statistical production processes in Istat, in particular consistency in “identification” and “estimation” for the whole integrated system of units and variables - of course not for all variables, but at least for some core ones.

There are three main components within the SIR, the Base Statistical Registers (RSB), the Extended Statistical Registers (RSE) and the Thematic Statistical Registers (RST). In the general pattern of the SIR the relations that exist among them can be described by a schematic diagram based on a unit-variable matrix (Figure 1).

Figure 1



The RSBs identify certain types of basic objects and key populations that are necessary for official statistics. In particular,

- the Population Register (*RSB of individuals, families and households*);
- the Business Register (*RSB of business units*);
- the Real Estate Register (*RSB of Places - RSB of building and dwellings; geographical database; GIS*);
- the Activity Register (*RSB of Relations among the objects of the base registers*).

The base registers may contain different types of statistical units, connected to each other, and several statistical populations. Each statistical unit can refer to more than one statistical population.

The RSBs contain for each identified object/s:

- a unique statistical identification code, stable over time, which identifies the specific statistical unit;
- all the characteristics that could be useful to identify – both semantically (by means of inclusion rules) and pragmatically (by means of enumeration) – all the statistical populations that are essential for supporting the official statistical production. E.g. the resident population, the present population, the population “standing” on a territory, the active enterprises, the non-profit units, the institutional units and so on;
- all the characteristics that could be useful to identify the object to support the production of lists. E.g. company name, address, first and last name if strictly necessary;
- the minimum set of useful variables to statistically characterize the object to support the extraction of subpopulations or samples. E.g. for individuals: sex, age, citizenship; for business units: Nace code, legal form, size.

The base registers are identified by the set of statistical units that belong to the same object. The statistical units may correspond to those stated in the specific administrative registers or may be the result of transformation or merging/split of units occurring in one or more administrative registers, as conceptually different. Units like “individuals” are an example of the former, while both the concepts of “family” and “enterprise” differ – or may be different – from those encountered in administrative registers.

The base registers contain all the necessary information so that the units can be recognized as belonging to a specific statistical population. In this sense the RSB may contain more than one statistical population. For example, with reference to the Business Register, it contains both the enterprises supporting processes related to Business Statistics and non-profit units (some of which are considered enterprises). In the Population Register, various statistical populations can be defined that relate to residents or to habitually dwellings, or to individuals who “stand” on a territory. In general, the relationship unit-statistical population is of type (0, m), as a unit can be part of more populations or may not relate to any of them separately.

The identification of the statistical populations comes from international regulatory requirements or from specific information needs: the *resident* population is defined and identified by reference to Regulation (EU) No 1260/2013 of the European Parliament and of the Council of 20 November 2013 on European demographic statistics, while the population “*standing*” on a territory is defined and identified on the basis of information needs to support the local government of specific areas.

The Activity Register is a particular type of RSB. The activities allow you to describe the relationships that can exist between different actors (i.e. entities – such as individuals or economic units – which perform actions). Each activity (e.g. study, work, ...) has its own unique identification of its attributes and can involve more individuals and/or more economic units and/or more territorial units.

The RSEs extend the information available in the RSBs for a specific population (or subpopulation) by means of statistical integration of other variables – also derived from statistical surveys – in order to support one or more specific statistical processes.

An example is the *Frame-SBS*<sup>2</sup> that for the specific population of “enterprises” in the RSB Business Register extends the information with individual data relating to the income statement.

The RSEs are informative “extensions” of specific populations identified in the RSB. They are usually defined to support specific regulatory requirements (EC regulations) and, by integrating administrative and surveys data, have the function to reduce burden and costs of a statistical production process or a consistent set of processes. For each unit of a specific statistical population the RSEs identify the values of certain core variables. The statistical result is derived by the sum. The variables in the RSE have one and only one definition, usually determined by the same international regulations.

The RSTs identify the information systems to support more statistical processes through a consistent and shared treatment of some variables, defined as core variables, for specific topics that could be relevant for the official statistics. Unlike the RSEs, they are not bound to specific populations, but rather have the goal of supporting statistics that refer to multiple objects.

The RSTs have characteristics different from the RSEs. If the latter, in fact, must support the rigidities determined by the specificity of populations and variables, the RSTs have fewer constraints. Their goal is to operate a homogeneous and shared treatment of some *core* variables aimed at a specific thematic output and designed to support more processes: they are both *specific*, because dealing with particular issues (education, work, income) and *generic*, because where possible designed as a tool at the service of

---

<sup>2</sup> In 2013 the Istat Department of National Accounts and Economic Statistics has started the implementation of a new information system (called *Frame-SBS*) to generate SBS data, for the production of key economic accounts statistics based on the massive, integrated use of administrative data, complemented by survey data for the estimation of the remaining economic accounts statistics.

multiple statistical processes (production lines) that can also refer to different objects (individuals/economic units).

The reference *population* is related to the particular theme (e.g. all persons who produce income in Italy, resident and non-resident ones, all students enrolled in Italian schools, all units linked by company obligations regardless of their residence, etc.). Therefore, the RST are not limited to a specific statistical population but may include a portion of it, or even segments of more populations (e.g. referring to work: all resident persons employed in Italy and all persons employed of resident companies in Italy) or, finally, they can include units that are not defined in any statistical population (e.g. non-resident individuals or economic units, with no permanent establishment in Italy, that produced an income in the national territory).

The values of a specific *variable* can be determined by multiple metrics because they can refer to multiple definitions (think of the difference between the variable “income from real estate” in National Accounts and the homonym variable in the European Union Statistics on Income and Living Conditions – EU\_SILC).

With their extremely flexible informative structure the RSTs are able to produce information that can introduce new conceptual and definitional frameworks, also allowing to anticipate and possibly induce potential changes in the European regulations.

They also represent a bridge between business statistics and demographic and social statistics. The information on individuals and on economic units are treated in the same informational and methodological environment that goes beyond the traditional households-enterprises dichotomy, not only and not so much from a theoretical point of view (e.g. “supply” vs “demand”), but also from the point of view of the processing of information, with the aim to make the most of all those available, both cross-sectional and longitudinal.

### **3 The SIR Integration Systems**

In order to ensure a coherent integration among objects and characteristics coming from multiple input sources, specific criteria, proper methodologies and suitable processes has been detected. They are summed up within the SIR Integration Systems, which stands out in detail in a *conceptual, physical and statistical* integration.

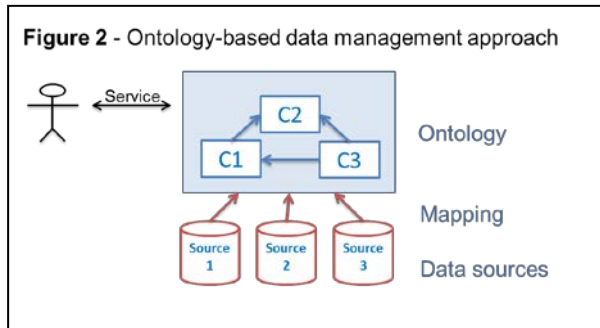
The *conceptual integration system* allows the identification of metadata – definitions and classifications – related to objects and characteristics contained in the registers; it is able to identify the consistency and/or the changes that occur in objects and characteristics of the same register over time; moreover it may possibly identify the level of consistency that exists between objects and characteristics in more registers. The set of problems related to the representation of the information in the registers take on a crucial importance, that is more complex than the typical meta-information systems of a survey. The latter, in fact, on the one hand are defined a priori in a statistical process, on the other hand have a sufficient stability over time. Quite the opposite for what concerns the administrative registers, the documentation is often lacking to some extent for inability to obtain the necessary information and at the same time complex, because it is determined by law and sometimes extremely specific rules.

In this view, an important position is taken by the coordinated management and the representation of sources’ metadata for both objects (units) and attributes, in terms of definitions and classifications even at the single modality level. The conceptual integration system can be effectively designed as a real management and representation system both for metadata and data. In this way, it also performs the functions of quality control and of consistent management of temporal evolution.

A particularly promising approach for the realization of the conceptual integration system is called “Ontology-based data management” (OBDM), to access the data layer of an information system and to represent the information in the registers, as illustrated in Figure 2. It is a new paradigm for accessing, integrating and managing data, whose key idea is to resort to a three-level architecture, constituted by:

- the **ontology** (a formal description of the domain of interest, specified in terms of formal descriptions of concepts, binary relations between concepts, and attributes),
- the **data sources** (the repositories used in the organization by the various processes and the various applications; sources can be external, independent, heterogeneous, laying on different storage media),
- and the **mapping** between the two (the mapping layer explicitly specifies the relationships between the domain concepts-ontologies on the one hand and the data sources on the other).

With this approach, the domain of interest is not only conceptually documented, but also this “effort” can be used to implement and/or manage the physical integration of the data. In other words, the



“translation” of data from “metadata of the sources” to “metadata of the ontology” can be achieved through the mapping.

This approach has enormous advantages in terms of industrialization of the process, of generation of the embedded data and the quality controls, of flexibility in adding new sources or modifying the existing ones, etc. Actually a difficult problem to face, in particular for complex sources such as tax register or social security, is the treatment of changes in time of definitions, classifications and scope. The objective of

the conceptual integration system is to follow the evolution over time of the variables and their modalities, intercepting changes and making their metadata suitable for processing as much as possible. Another essential aspect of this horizontal cross-cutting activity is the cross analysis of content to multiple sources: not only to follow a register in time, but also to assess if the significance of the variables is the same in multiple inputs having the same names, codes (tax codes, for example), descriptions, definitions, scope. This activity seems crucial for particular topics that has a cross-cutting nature to the processes that branch off from the base registers such as “employment” and “income”.

A coordinated management that could take into account several different purposes is a challenge: this system is particularly complex and requires to track down a very huge number of dedicated resources, able to develop and manage a unique metadata management system for the metadata related to sources, integration process, analysis and dissemination. In any way metadata will be defined, they will have to manage the meta-information for all the production processes that will use them.

The *physical integration system* delineates the operating, deterministic and probabilistic procedures for the unit identification. Starting from the data collection (microdata collected through statistical surveys or administrative sources), it takes the form of a linkage operation having the goal of identifying the presence of the same elementary unit (a natural person, a legal person, a place, a relation) in all the various sources. It is able to identify the same object in different sources with a unique and stable (over time) ID number; moreover it can delineate, for each object, the logical and physical relationships, in time and space, among the information available from the various sources; additionally it is able to identify and define the logical and physical relationships which occur among objects of the same type (e.g. relations between individuals) or between objects that belong to different types (e.g. relations between individuals and economic units). Thus, the unique and stable identification keys and the temporal validity references (starting year, end year) could be assigned to the identified objects. In this way it is possible to define the combination of the sources in which a given element is present in a set time interval and thus to create a logical/physical connection between the totality of the information concerning it in all input sources.

The *statistical integration system* identifies all the criteria, methodologies and processes to derive units, populations and statistical variables from the objects and the characteristics of the input sources. It allows to define the consistency between information coming from integrated administrative archives and information deriving from statistical surveys, in order to correct and/or evaluate the statistical information

coming directly from administrative sources and with the aim to produce individual indicators extended to all the persons inside the administrative archives that are not included in the survey sample.

Actually in Istat this type of activity has already been developed in the old stovepipe production system, both in the set-up of the base registers and in the statistical production (e.g. in the processes and methodologies for creating the so-called *Frame-SBS* or in the imputation and integration systems used in EuSilc).

The renewed and specific emphasis on an integrated management of the input sources derived from two interconnected circumstances, outgrowth of ESS policy objectives. On the one hand the increase in quality and detail of the estimates and the simultaneous reduction of the statistical cost and burden; on the other hand, the dropping of the “stovepipe” model in favour of cross-cutting and integrated approaches. The statistical integration processes are therefore the strategic key to the realization of the statistical production models able to meet these objectives. In this innovative context, the opportunities to achieve an efficient use of the available information are placed not only in downstream phases of the statistical production processes, but also and mainly in the upstream phases, which directly involved the planning, the design and the logic itself of that process. On these issues a big effort has been made by Eurostat in recent years through the promotion of a large number of ESSnets and other initiatives to encouraging actions on the suitable methods to support these policies.

Nevertheless the strategic importance of the statistical integration processes is particularly evident for the integrated use of information coming from all administrative sources with those collected through sample surveys. Especially in domains characterized by the combination of both types of sources, the statistical integration process is able to transform the result of the reconciliation of the sources’ errors (of different nature and intensity) in a statistical product. More generally, it is to reconcile the sampling and non-sampling errors of surveys with those of the administrative sources, such as measurement errors and misalignments with the definitions underlying the statistical regulations. There is the need to reconcile the “objective” elements of administrative sources and the “subjective” ones resulting from surveys. The statistical integration also proposes a switch from integrated samples, long used in Istat, to the integration of the information available on the entire population starting from the integrated sample. This step must necessarily be based on the probabilistic approaches which need to exploit the information available on a longitudinal basis in order to fully grasp the opportunities offered by administrative sources.

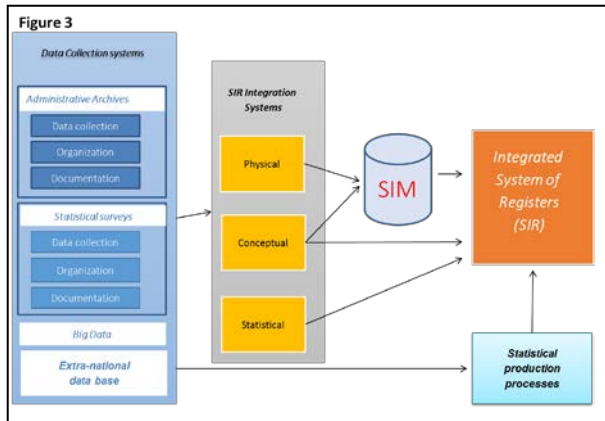
In addition the scope of the sources allows the analysis of phenomena related to more populations and not necessarily related to the most frequently used concepts in official statistics (such as the concept of resident population or present population). This is because the administrative sources are able to record events associated with statistical units that relate to a fairly broad range of ways in public institutions. This allows you to single out non-standard subsets, such as the resident workers who also work in non-resident units and workers (resident or not) working in resident units.

One consequence is the possibility that using the administrative sources some not standard definitions or not standard populations could result, useful for informational purposes and to support the policy. A typical example is the concept of population “standing” on a territory, measured through the registers of workplaces and study places.

#### **4 The functional relations among the present and future registers within the SIR**

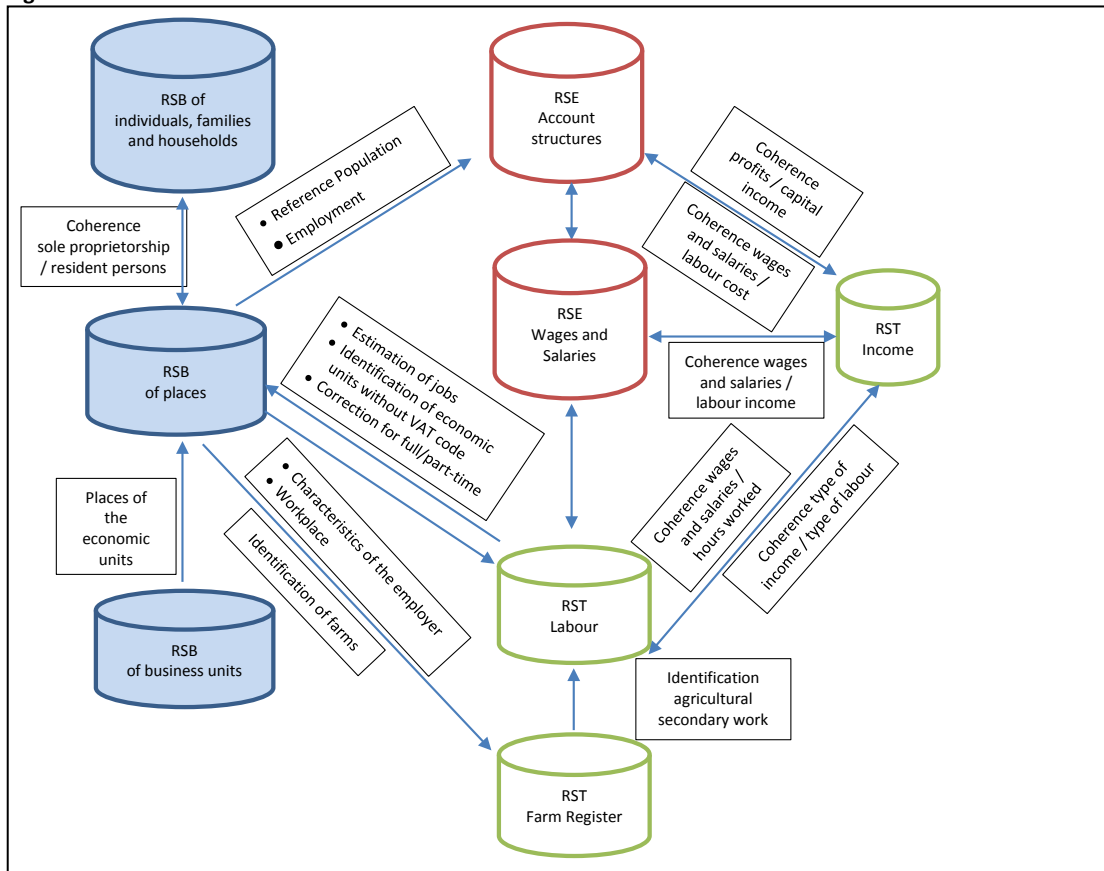
In order to represent all the functional relations, the ones within the SIR and those with the outside components, two separated diagrams will be presented, to show the circular information flow within the system.

Figure 3 shows how the physical integration system operates exclusively on the RSBs, as it is the system that is used to uniquely identify the basic units. The identification of statistical variables also determines the definition of the statistical populations of the RSBs. Conversely the conceptual and statistical integration systems allow to identify the units and the statistical variables that intervene on the whole SIR.



In the Data Collection Systems the source “Extra-national Data Bases” specifies the need for interoperability with the RSB in the other EU countries (i.e. relations with EGR-IS). With reference to the BR, such integration is already operational through the current EuroGroups Register (EGR) and, in progress, through the European System of Interoperable Business Registers (ESBRs). This is a requirement now highlighted and currently in careful assessment also for individuals. Figure 4 shows the relationships between the various parts of the system, focussing on the present and future possible occurring elements of the SIR.

**Figure 4**



Some relevant aspects should be highlighted:

1. *The close relations that exist between different RSBs.*

If those between places and both individuals and business units is evident, that between the RSB of individuals and the RSB of business units is less intuitive. Even if the identification of the statistical units in the two RSBs can proceed through separate production processes that use different methodologies, nevertheless an “evidence” in one register can be an input – for evaluation – for the other register. An example is the identification of a statistical unit in the RSB of business units, when it is characterized as “sole proprietorship with the owner coming from a non-European country”.

This characterization may be acquired in the RSB of individuals to update the population of “habitually dwelling in the country”.

Some problems of consistency of the SIR may occur, regarding both units and variables. There could be units in one base register that do not have a correspondence in another base register to whom they are linked (even if they should have it) and the reasons could merely be the use of administrative input sources that have a different time lag in updating information (i.e. the self-employed persons performing an activity having the characteristics to be included in the BR, should be contained also in the Population Register as individuals).

It must be explicitly planned how the system should align itself in presence of the inconsistencies of data over time. This also include the decision on how and when the system should implement new units when these are caught by surveys but not yet registered in the SIR.

Coherence also regards variables, as they could have different definitions in the registers. When considering employment from the two supply and demand points of view, some problems of measurement arise (i.e. heads vs jobs) and may happen that one register is an input to adjust the measure in another one (i.e. from part-time to full time), thus bringing about the necessity to align them again with economic and budgetary data and data on labour cost, with a big impact on productivity and other business statistical indicators.

2. *The possibility of defining derivative registers starting from the integration of two or more registers.*  
Making a suitable use of data coming from the RST Labour and the RST Education and training, there will be the possibility of establishing a register on “Training-Job paths”.
3. *The circularity of information within the system.*  
The RTSs can be input for some RSEs, e.g. the development of the RSE Job supply and the RSE Job demand could take input data from RST Labour. In some cases the RSEs could be even the input for the RSBs. For instance, in case of the SBR, data on employment from the RSE Job demand is used to estimate the variable “size of the enterprise”.

## **5 The new role of the SBR within the new integrated system**

In order that the integrated system works, actually an extension of the RSBs must be considered, to include all units that are necessary to the system. Especially the two registers “RSB of business units” and “RSB of individuals, families and households” shall contain all the units, and also admit overlap.

Taking the 2008 SNA as a reference, an SBR should record all institutional units that are engaged in productive economic activities.

In Figure 5 an example scheme is shown where an attempt has been made to describe the coverage of the current register and its future foreseen extension, in terms of institutional sectors.

All desirable units are included in what can be called the sector of “*producers*”. They can be divided into three subsets, according to the extent of their registration in the administrative archives that are main input sources for the SBR, combined with the “purpose” of their production.

In the first subset, that usually contains units that are registered in the Tax register or in the Chambers of Commerce, the producers can be distinguished according to the type of their production, if it is *market* or *non-market* oriented. Market producers are enterprises that sell all or most of their output at prices that are economically significant and make decisions about what to produce and how much to produce in response to expected levels of demand and expected costs of supply and are exposed to the risks associated with this production. Thus, Government units and most of the non-profit institutions (NPIs) that provide most of their services at prices that are not economically significant, are considered nonmarket producers.



**Figure 5 - Current coverage and future enlargement of the Base Register of productive units**

Type of Units			Additional information	Coverage	
PRODUCERS	Traditional coverage of SBR	Market	Non-financial corporations and quasi corporations - private and public (S.11)	Legal form profit	Included in SBR
			Financial corporations and quasi corporations - private and public (S.12)	Legal form profit	Included in SBR
			NPISHs (non-financial S.11 and financial S.12) that behave like corporations	Behave like corporations - Legal form non-profit	Included in SBR
		Non Market	Unincorporated household enterprises (S.14)	Registered (proxy = VAT code) or with greater than a given number of employees - Distinguishing Agricultural and Non Agricultural	Included in SBR
			Government (S.13)	Public producers	Included in SBR
			Non-market NPISHs (not controlled by government S.15)	Non market non-profit institutions serving households - private producers	Under construction
	Extended coverage	Market	AGRICULTURAL household enterprises (S.14)	Without VAT code, market, selling most or all production, eligible according to threshold in Regulation (EC) No 1166/2008 (FSS)	Planned
			Own-account non-registered NON AGRICULTURAL household enterprises (S.14)	Including also self-employed. Outworkers without VAT code and a with a minimum threshold of turnover	Planned
			Other own-account, unincorporated household enterprises (S.14)		NO
	Household non market enterprises		Households only undertaking production for own final use		NO

As regards NPIs, they are diverse in nature, even if are mainly nonmarket producers, in that they are producing goods and services not for the generation of income or profit. Some behave like corporations and some undertake activities similar to government but independently of it. The NPIs that behave like corporations should thus be classified under S.11 or S.12. Other NPIs that are controlled by government are to be classified under S.13. The remaining NPIs, those that produce goods and services but do not sell them at economically significant prices and are not government controlled, are classified in a separate institutional sector (S.15) called non-profit institutions serving households (NPISHs). They are mainly non-governmental social institutions, such as churches, social clubs, charitable associations, etc.

Other producers in the same subset are also the financial and non-financial corporations and quasi

corporations and the registered household enterprises, that include also agricultural enterprises. This last group, the household unincorporated enterprises, that also includes agricultural ones, contains households that in the SBR are considered enterprises, they have a VAT code or have more than a given number of employees. For establish a parallel with 2008 SNA, it contains partnerships not recognized as independent legal entities and sole proprietorships. Those having the characteristics of quasi-corporations are in S.12 or S.11 depending on if they are financial or not financial, and the rest is in S.14. This first subset of units is considered traditionally as part of the desirable SBR coverage.

Of course the SBR has not reached the full coverage from the moment of the first set-up. The BR was historically focussed on business units, within the SBS scope, and its role was primarily to be a support for business statistics – a business register for the business statistics. As expected, the core coverage of the SBR will continue to be composed of all active economic units that are market producers and whose principal activity is the production of goods and services. Registered market producers are very important from an economic point of view and in principle are not difficult to cover based on available administrative registers.

However, for some years now the need has already been felt to widen the scope to non-market units and therefore for first all the public producers have been covered (S.13), as a first step.

As regards NPIs, while the market NPISHs that behave like corporations have already been covered in the SBR, the non-market NPISHs (S.15) are now subject to an enlargement program using new input sources and data from the last Census, thus they will be included very soon, to set-up the new Register of NPIs. This can be considered the second step of the enlargement, that is already widely shared by all users. Nevertheless, since the SBR is now included in an integrated system it needs an enlargement to include all units that are necessary to the system. The next enlargement will be searched in the household sector, that is in the second subset of units in figure 5.

The second part of the “producers” are units that are not immediately recognized as enterprises. They belong to the household sector, they do not have a VAT code and are not present in the typical input sources used in the SBR updating process to consider a legal unit as an enterprise.

In the household sector can be identified some units engaged in the production of goods or services with the primary objective of generating employment and incomes to the persons concerned. These units typically operate at a low level of organization, with little or no division between labour and capital as factors of production and on a small scale. They may operate without a fixed location, or in homes, small shops or workshops. Many of these enterprises are operated by an individual working alone, as a self-employed entrepreneur (own-account worker), or with the help of unpaid family members, while other unincorporated enterprises may engage paid workers.

They could be unincorporated enterprises or own account enterprises, including self-employed, that are market and selling most or all production. They are less than a specified size in terms of the number of persons engaged, or of employees or of employees employed on a continuous basis. It is important to distinguish whether it is agricultural or not.

Giving a measure of this part of the economy could lead to an extension of the market economy beyond the traditional enterprises. Of course there are some difficulties in identifying such units for placing them and keeping up-to-date in a base register and above all to have enough information for estimating the main characteristics. The limit could be to rely simply on the size of the enterprise, defined either in terms of turnover or number of employees.

In the above scheme it is shown which part is likely to be included in the “RSB of business units” to enlarge the coverage of the existing SBR and to provide frames for a wider variety of surveys in different sectors of the economy. An operative approach has been chosen that takes into account the difficulties in undertaking the maintenance of an extended register and the availability of suitable sources.

Since integration of agricultural statistics into a comprehensive and coherent system of economic statistics is one of the goals for the next future, the first step is to include in the base register that part of the Farm Register that is now excluded from the SBR. They are units having an organizational structure according to the thresholds suggested in the Regulation on farm structure surveys and the survey on agricultural production methods, that is the EU criteria for including eligible units in the Farm Register<sup>3</sup>. In this case thresholds are very low, they are physical minimum thresholds on both surfaces and livestock, with the aim of a 98% coverage.

Another effort will be to include the own account workers, namely outworkers non-registered in the administrative archives that are main input sources for identifying the unit “enterprise” in the SBR (specifically the Tax register and the Chambers of Commerce), even if they are recognized in some other secondary administrative sources. They do not have a VAT code and until now they were considered as an input of labour for an enterprise in the traditional sector.

Up to now they are included in the SBR as registered household enterprises just in case they have employees (as employers workers) or if they have a VAT code, since this was the criteria to consider it a business unit.

Thanks to these new inclusions the new *Base Register of Productive units* will be not only a business register for the business statistics, but also the benchmark for National Accounts.

As regards the third set of units that can be considered “producers” in figure 5, they are households having a different purpose for their production (own-use production), the so-called “household non-market production for own final use”. There is no plan in future to take into account a possible inclusion in the SBR. To decide which are the households that belong to this part and are excluded from the

---

<sup>3</sup> Regulation (EC) No 1166/2008 of the European Parliament and of the Council of 19 November 2008 on farm structure surveys and the survey on agricultural production methods and repealing Council Regulation (EEC) No 571/88.

previous sector the only suggestion seems the purpose of the production – specifically services produced by households for their own consumption – even if there could be interest in measuring these activities for some forms of analysis. It could be interesting, for example, to measure the “subsistence agriculture” (the results of any agricultural production which are used entirely by those responsible for the production). However in the same way in which it is considered excluded from the Farm Register, also in the extended register will not be included.

While the services provided by owner-occupied housing are surely excluded, a doubt remains whether to consider or not the services provided by paid domestic staff, or the caretakers of condominium buildings<sup>4</sup>. The relevance is a consequence of the existence of the SIR, since in an integrated system it would be desirable a connection with employee or household-based surveys such as the labour force survey where employees of these units are likely to be included. However there is a complete coverage guaranteed by the system, since all the units will be included at least in one base register: the third subset in the above scheme, household non-market enterprises will be included only in the “RSB of individuals, families and households”, without any overlap with the RBS of productive units, and its entities will not be treated as economic units.

## 6 Conclusions

The current Italian modernization programme is based both on a strong and structured governance and on production lines modelled on registers, where the BR will have an enhanced central role. The new Integrated System of Registers is a core element of this project of modernization.

The SIR is a single logical and functional environment to support the consistency of statistical production processes in Istat, whose main components are the Base Statistical Registers, the Extended Statistical Registers and the Thematic Statistical Registers.

The SIR should provide for a system of governance of both the management and the data supply processes of the registers, that will guarantee stability and consistency for the whole system, taking into account some basic principles of the overall governance of the system. Governance will be regulated by the Committee for the SIR Management, having specific duties, however each “Variable Responsible”, who ensures the certification of the variable, will manage the processes that impact precisely on that variable. They have to manage homogeneous groups of variables, such as those on labour or those on the income statements. An organised governance mechanism is needed so that there are clearly defined responsible persons for the system variables, both the core variables that define the population of the base register, and the thematic variables (mainly in satellite registers).

An absolute priority is given to the problem of data security, although it will be important as well to develop methods for integrating surveys and administrative sources and to manage temporal aspects.

The SIR will be built incrementally starting from what is already operational in the current situation. However, some problems have to be faced, also from a theoretical point of view. The close relations that exist between different RSBs could produce an increase in information but at the same time could be a source of coherence problems.

Within the new system, the Business Register will change its role to include all units that are necessary to the system. Some new sectors are likely to be included in the existing “RSB of business units” to enlarge its coverage and to provide frames for a wider variety of surveys in different sectors of the economy. An operative approach has been chosen that take into account the difficulties in undertaking the maintenance of an extended register and the availability of suitable sources.

In this way the new *Base Register of Productive units* will be not only a business register for the business statistics, but also the benchmark for National Accounts.

---

<sup>4</sup> Of course the enterprises that comprise economic activities of households that are employers of domestic and other personnel should be included in the SBR if they are registered as employers.

## References

**AMBROSELLI S, GAROFALO G.** (2015), *Reversing the flow: from an integrated system of administrative microdata to an infrastructure for the users*, proceedings of NTTS 2015 - New Techniques and Technologies for Statistics, Bruxelles, 9-10 March 2015.

<https://ec.europa.eu/eurostat/cros/system/files/NTTS2015%20proceedings.pdf>

**AMBROSELLI S, DI BELLA G.** (2014), *Towards a more efficient system of administrative data management and quality evaluation to support statistics production in Istat*, Q2014, Wien, 2014.

**ALLEVA G.** (2015), *Adding Value to Statistics in the Data Revolution Age*, ISI2015, Rio de Janeiro, July 2015. [http://www.istat.it/it/files/2015/10/GAlleva\\_IPS\\_ISI\\_WSC\\_2015\\_AddingValueStat\\_PAPER.pdf](http://www.istat.it/it/files/2015/10/GAlleva_IPS_ISI_WSC_2015_AddingValueStat_PAPER.pdf)

**LENZERINI et al.** (2014), *Ontology-based Data Management for the Italian Public Debt*, proceedings of the Eighth International Conference (FOIS 2014), pp. 372-385 - Formal Ontology in Information Systems, Rio de Janeiro, September 2014.

<http://www.dis.uniroma1.it/~graphol/documentation/FOIS-2014.pdf>

**UNECE** (2015), *Guidelines on Statistical Business Registers*, New York and Geneva, 2015.

[http://www.unece.org:8080/fileadmin/DAM/stats/publications/2015/ECE\\_CES\\_39\\_WEB.pdf](http://www.unece.org:8080/fileadmin/DAM/stats/publications/2015/ECE_CES_39_WEB.pdf)

**EC, IMF, OECD, UN & WORLD BANK** , *System of National Accounts 2008 (2008 SNA)*, European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations and World Bank, New York, 2009.

<http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>