DIRECTION DES STATISTIQUES ECONOMIQUES
Département Répertoires et Statistiques des Entreprises
Division Statistiques Structurelles des Entreprises

Paris, le 27 septembre 1990 n° 250/E130

THE FRENCH SYSTEM OF ENTERPRISE STATISTICS

Paper to be presented at the
Fifth Meeting of the Voorburg Group
in Paris 1-5 October 1990

(Item 8 of the provisionnal agenda)

B. CAMUS

Describing exhaustively in a few pages the French system of enterprise statistics is quite impossible. This short paper attempts to give a general overview of the system with an emphasis upon its logical structure. Section I deals with the guiding principles as far as they seem to be original when compared with other countries'systems. A short presentation of three operations leading to future developments is given in section II. The references [1] and [2] in the bibliography provide more information about the genesis of the system and about all the statistical agencies and sources of data concerned with enterprise statistics. In particular, it is explained how the French system is both integrated in its design and decentralized in its organization.

I - THE GUIDING PRINCIPLES

The French system can be seen as a set of tools with strong interactions. In the sixties, the setting up of the first National Accounts pointed out the discrepancies between the different sources of data. It was then decided to establish a statistical system on enterprise with a strong emphasis on the consistency of the data. So, the aim of the statisticians has been to develop interactions between sources with a view to the reciprocical improvement of these sources and the reduction of discrepancies between statistics from varying sources. This purpose explains four main features of the system :

1 - Heavy investment in the creation of basic tools.
2 - Intensive use of administrative data.
3 - Development of statistical surveys in an integrated and coordinated way.
4 - Search for consistency leading to micro-economic syntheses.

These outlines are summed up in figure 1.

I.1 - Investment in basic tools

In order to get consistent data, one has to work from the outset, upstream the collection of data. There are three basis tools : the classifications (or nomenclatures), the register of firms and the general accounting plan.

The first item (standard classifications) is quite usual. What is noteworthy is the will to introduce the classifications into the administrative files and

the reporting system of the firms. In fact, the official classifications of activities and commodities are compulsory for governmental agencies and are only recommended for firms.

. The register SIRENE

The first condition to get consistent data is to have a clear definition and a complete enumeration of the collection units. Hence the creation of registers.

In France, the statisticians manage an inter-administrative register called SIRENE (acronym for computerized register of enterprises and establishments). It is an exhaustive list of units characterized by identification data (name, address, legal category, type of activity,...). As for now, there are two units : the enterprise (which corresponds to the concept of legal unit) and their establishments (more precisely, in the French context, the local units belonging to the enterprise). These are administrative units. Within the register, one tries to get real statistical units by removing "pseudo-units". There are two types of pseudo-units : those whose activity is irregular or non-productive and those that are transparent from the accounting point of view (activity recorded in the accounts of partners or in the units to which they belong). All in all, nowadays the register contains approximately 3 000 000 enterprises and 4 000 000 local units.

The register is a continuously operated system for identifying enterprises and their establishments. It allocates a non-significant nine-digit identification number to each new enterprise (called SIREN). The enterprise's establishments are identified by a "SIRET-number" with 14 digits, the first nine of which are the SIREN.

Morevoer, the register assign those units a code number describing their main (or principal) activity (called code APE). As for now, it is coded according to a 600 - headings classification.

Governmental agencies are obliged to use the SIRENE identification number in their dealings with enterprises and when they exchange data on and from enterprises among themselves. They must also include the main activity code in their files. This dissemination of codes is a necessary condition for getting compatible statistics by activities.

From a statistical point of view, the register is also a basis for sample surveys of enterprises and establishments. The register contains two statistical variables used for stratification : main activity and number of employees. The address allows stratification according to the region.

. The general accounting plan

The French statisticians can take advantage of a favorable background. In France, there is a long and old tradition of standardization of the business accounting. The first general accounting plan was issued in 1947 and revised in 1957 and 1982. This plan is quite detailed. The main items are compulsory for the enterprise and are used as references for the tax declaration on profits.

During the revisions of the accounting plan, statisticians have expressed their needs especially for National Accounts. On the whole, they succeeded. One of the main features of the French system is that in the profit and loss account, expenses and receipts are presented by nature of transaction and not by function (production, selling, administrative) as in the Anglo-Saxon countries. This is consistent with the National Accounts analysis. Moreover, the revision of 1982 has led to the introduction of notions drawn from the National Accounts standards : production, value added, gross operating surplus,....

This situation has two main advantages. First, in the statistical surveys, French statisticians follow the accounting plan as much as possible in order to limit observations errors. The main concepts are defined on the basis of the accounting headings. Very often, the questionnaire asks for economic breakdown of accounting headings. This is the case for measuring production, investment and subcontracting. Secondly microeconomic accounting data can be used to get aggregated data for National Accounts. Hence the possible use of tax data.

I.2 - Intensive use of administrative data

This is very specific to the French experience. There are historical reasons for that. The context was very favorable in the post-war period when first National Accounts were established. Moreover, the statisticiens adopted a deliberate strategy, using different tactics.

The first is to retrieve the administrative data files in order to make a statistical processing. One of the task of the Statistical Office, which was established in 1946, is to coordinate statistics using components provided by various government departments and agencies. This legal basis for the transfer of individual administrative data was reinforced by an ad hoc law in 1986.

The second is to work as upstream as it is possible. Statisticians tried to be involved in the definition and the processing of the administrative records. One way is to provide INSEE staff to the various government departments in order to design and carry out the data processing. For example, this has been done for the tax and the customs departments.

The third way is to manage directly some administrative sources. The best example is the register SIRENE managed by the statisticians under the overall direction of an interministerial committee.

Thus, there is a long tradition of utilization of administrative data. The annual reports on employees (acronym DADS in French) are regularly used since 1950. They cover about 1 300 000 employer's establishments and 15 000 000 employees. The annual reports on profits for the tax declarations are processed since the sixties. Nowadays, they provide accounting data for more than 2 000 000 legal entities.

The utilization of administrative data is not straightforward. It entails major problems : unwieldy and incomplete files (exemption, late returns,...), low quality for the statistical variables, inefficient editing,... But these problems may be overcome and the quality of the statistics can be satisfactory.

Moreover, the future extension of Electronic Data Interchange (EDI) will reinforce the interest of using administrative data. In fact, one of the main weaknesses of these sources was often the timeliness. For example, the statistical processing of annual tax data would start six months earlier if Electronic Data Interchange was used.

II.3 - Integrated statistical surveys

In this matter, there are several aspects. First, some statistical surveys have been designed as to supplement administrative data. As regards structural

statistics, it appeared, in the sixties, that tax data could be considered as a substitute for census because they provide some minimum data for all the legal entities. So statistical census on enterprises were abandoned. However, the administrative sources could not cover all the information requirements. For example, they provide no detailed data on production, investment or volume of work. Therefore, it was decided to develop light annual surveys with sampling for the smaller units. This is the origin of the annual survey on enterprises (acronym EAE in French). In this survey, the main sections cover data not collected in the tax data : breakdown of sales by economic classification, breakdown of the increase in fixed assets, measure of the volume of work,... (for more information about this survey see references [10] and [11]).

Secondly, the annual survey on enterprises was designed to be a source which could be merged with many statistical or administrative sources (see figure 2 where the center of the "daisy" shows the statistical units and the associated variables, the petals represent the other sources relating to the same variables). In fact, this enormous data set has never been compiled. It is partly done in a data base called ENEIDE (acronym for data base on manufacturing industry enterprises for economic development) which covers only the manufacturing industries.

Thirdly, there is a complex system of interpendancies between the different statistical sources (see figure 3). For example, results of the annual survey on enterprises are used to update the register, as regards the main activity of the enterprise (within the annual survey, the data on annual sales and number of employees for each activity allow to check the classification of the units). In the manufacturing industry, the annual survey on enterprises provide the breakdown of sales by economic activity. This is used to identify the units to be interrogated in the monthly and quaterly production surveys. There is a feedback : detailed information about products makes it possible to improve the quality of the codification of activities at the more aggregate level of the annual survey. These are just examples, there are many other interactions between the statisticals sources.

I.4 - Micro-economic syntheses

The preceding features allow to link records from administrative sources and statistical surveys. The register provide the definition of the statistical units and their identification through a common identification number. The

administrative data go through statistical processing. And some statistical surveys are designed in order to be matched with other sources. A micro-economic synthesis has two main advantages : it gives the opportunity for reciprocical validation and mutual enrichment, it allows to carry economic analyses crossing variables from different sources. There are many potential mergers of files.

In fact, there is only one systematic and regular operation of record-linking for enterprises : the unified system of business statistics (acronym SUSE in French). This system brings together information from the two major sources of structural business statistics : tax declarations and annual survey on enterprises.

The aim of the system is to improve the quality of the data by checking their consistency in all the directions :

- internal consistency : compatibility of accounting and economic data for one financial year.

- longitudinal consistency (between one financial year and the next) : This results in a follow-up in time of the largest firms. A special attention is paid to the mergers, the splits, the scissions...

- intersource consistency : for the common items, for example the loss and profit account. If there are discrepancies, an attempt is made to understand why and then to reconcile the accounts.

The originality of the system is that it entails processing at the individual level. It was observed that there is a cause and effect relationship between the errors relating to large units and the statistical discrepancies between the sources at the aggregate level. So the data are processed and corrected at the individual level with a special attention for the largest units, namely the enterprise with more than 20 employees (these 70 000 enterprises represent about 75 % of the total activity).

This central system of structured statistics on enterprises has been built in the seventies. It is now a major source for the enterprise side in the National Accounts.

As regards the local units, there is also a micro-economic synthesis : the File of Large Establishments (acronym FGE in French). It is an historical file of the large local units (with more than 100 employees). The aim is to follow precisely the evolution of the number of employees by using different sources (mainly administrative). This is the tool to analyze the demography of the large local units.

II - BRIEF OUTLOOK ON THREE OPERATIONS

In the first section, only the main outlines were presented. The figure 3 gives a more detailed overview of the system. Describing all the tools an their links would be very long. Reference [2] in the bibliography provides a minimum information about all the sources. Here, will be presented only three operations selected because they seem original and they draw some stimulating perspectives for the future. Two examples concern the field of structural statistics on enterprises. To be exhaustive, one should also deal with the short-term surveys and the sources about local units.

II.1 - The coordination of the annual surveys on enterprises

In fact, the "annual survey on enterprises" system is decentralized. There are six enterprise surveys carried out by different statistical services. All in all, more than 200 000 enterprises are interrogated every year. In the manufacturing industries, the survey is exhaustive for units with 10 employees or more (there is a sample survey under this threshold twice every five years). Elsewhere, exhaustivity begins above 20 employees and smaller units are sampled (even for units with 0 employee). It is a stratified sampling in which the criteria are main activity, size in terms of employees, and region. Half of the sampled units are retained from one year to the following in order to allow comparisons of results over time.

The register SIRENE is used as a basis. This is done in an intermediary tool of coordination called OCEAN (french acronym for coordination tool for the annual survey on enterprises).

This tool has many interesting features (see reference [8]). Firstly, it contains a sampling frame that can be used for other surveys than the EAE. So the coordination in the drawing of the samples can be extended in the future. Were it done, one could measure and monitor the statistical burden between units. Secondly, the sampling method (using a random number technique) allows different type of

coordination : positive or negative, coordination of samples according to different criteria, coordination between local units samples and enterprises samples. Thirdly, it has been designed to exchange information between the register and the different statistical services carrying out the surveys. This is done mainly by on-line procedures for updating and consultation.

This is only one aspect of the coordination of the annual enterprise survey. The development of this survey is managed by a committee called "comité SSE" (french acronym for business statistics system). As for now, this committee plan to launch a "fourth generation" survey. This entails reflections about the questionnaire, the editing and correcting process, the dissemination of data, the use of EDI...

II.2 - The enterprise group

In the French system the main statistical units are the enterprise and the local unit. They are managed by the register that can rely on administrative declarations. Economic analysis outlines the importance of the enterprise-group or cluster (all the enterprises controlled by another enterprise). They are the strategic units : the main decisions about financial strategy, external growth, and diversification are taken at this level.

In France, there is not yet a legal status for this unit. So the statisticians had to develop a system to define a statistical enterprise-group. Very briefly, an annual survey on financial links is carried out. The basic data collected are the percent of the held capital and the resulting voting rights. An algorithm allows then to determine different types of cluster or enterprise-group. Then, statistics can be elaborated by aggregating data on enterprises over different sets.

This situation will evolve with the development of the European Economic Community. The statiscal unit enterprise-group will be introduced in the register as it will be done in the other European countries. This raises many difficulties, some of them being methodological (see reference [9]). Certainly, there lies one of the big challenges for the European statisticians in the forthcoming years.

II.3 - Combined use of tax data and direct surveys

Among the administrative data are the declarations filled by the firms for the payment of the Value Added Tax (VAT). It provides monthly data about sales and investment expenditures. This source has been used in France since 1976. It proved to be quite reliable, providing that some statistical processing is made. One difficulty was that the field covered was not large enough in some service activities. So it was planned to carry some supplementary surveys. This operation has already been presented (see reference [7]). This seems to be a promising direction for the future and this French experience remains quite original.

BIBLIOGRAPHY

. *In French*

[1] - "Les statistiques d'entreprises - Réalités observées et méthodes d'observation" - M. QUELENNEC - Collections de l'INSEE - E101 - Juin 1986

[2] - "Les sources statistiques sur les entreprises" - Collections de l'INSEE - E117 - Septembre 1988

. *In English*

[3] - M. QUELENNEC - The business statistics system in France and consistency between sources - Statistical Journal of the United Nations ECE 2 April 1984.

[4] - B. CAMUS, M-H. TAMISIER - The French Unified System of Business Statistics - Statistical Journal of the United Nations ECE 2 - 1984.

[5] - J-M. CALLIES - The management of identification registers by statistical offices - Statistical Journal of the United Nations ECE 2 - 1984.

[6] - M. QUELENNEC - Adaptating statistical machinery to meet increasing demand for statistics on the production system : sources to supplement surveys of enterprises - EUROSTAT Seminary - April 1989.

[7] - ACTIF, DEPOUTOT, HONNIBAL - The use of VAT files and of supplementary direct surveys in service industry - Paper prepared for the "Voorburg group" - October 1989.

[8] - F. COTTON - Use of SIRENE for enterprise and establishment statistical surveys - 4th Round Table on Business Survey Frames - Newport - October 1989

[9] - V. THOLLON-POMMEROL - The integration of clusters in registers and business statistics. Paper prepared for the 5th Round Table on Business Survey Frames - Williamsburg - October 1990

[10] - P. TROGAN - A survey on service industry in the French West Indies - Paper prepared for the Voorburg Group - October 1990.

[11] - V. MAILLARD - The French annual survey on service industries - Paper prepared for the Voorburg group - October 1990.

FIGURE 1

OVERVIEW OF BUSINESS STATISTICS SYSTEM

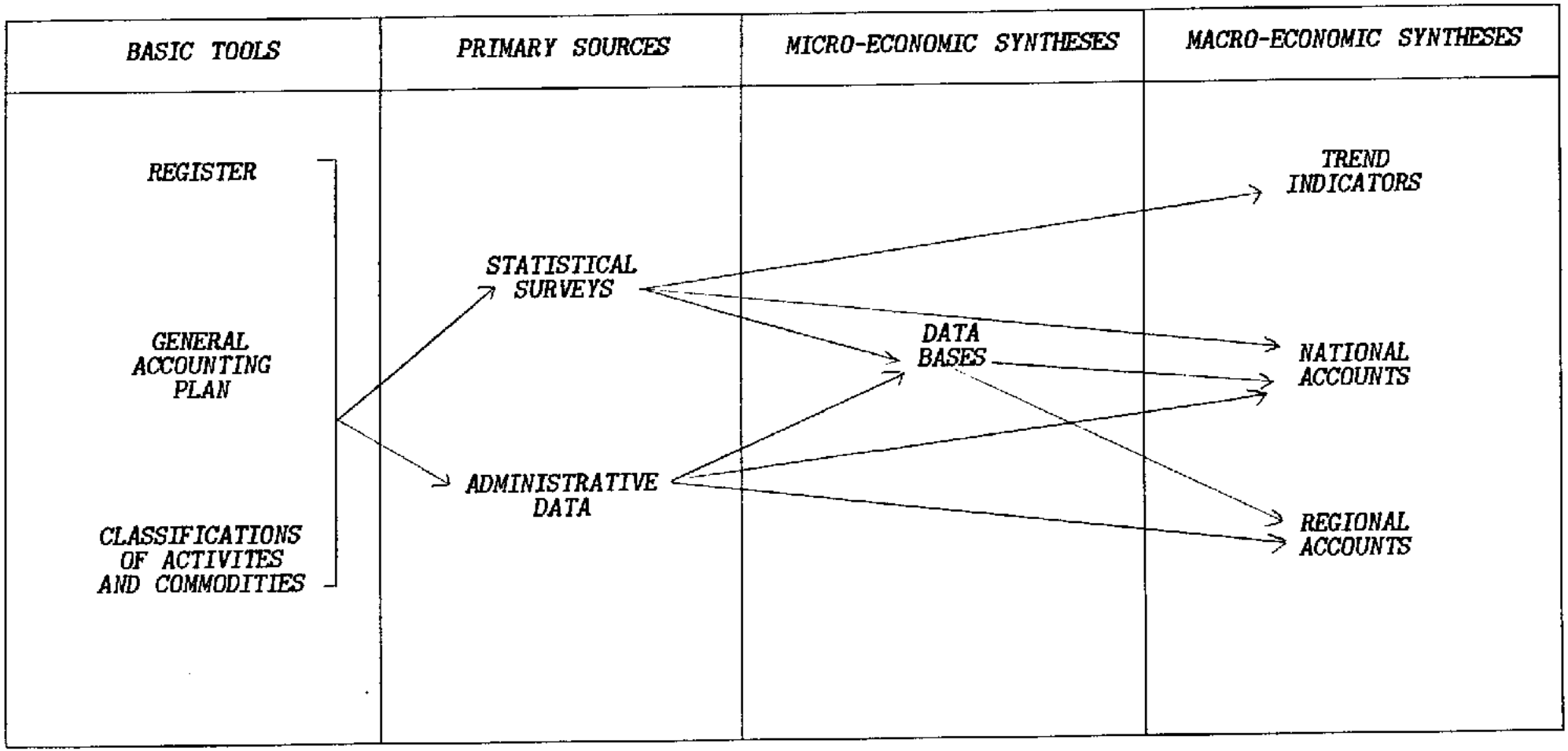
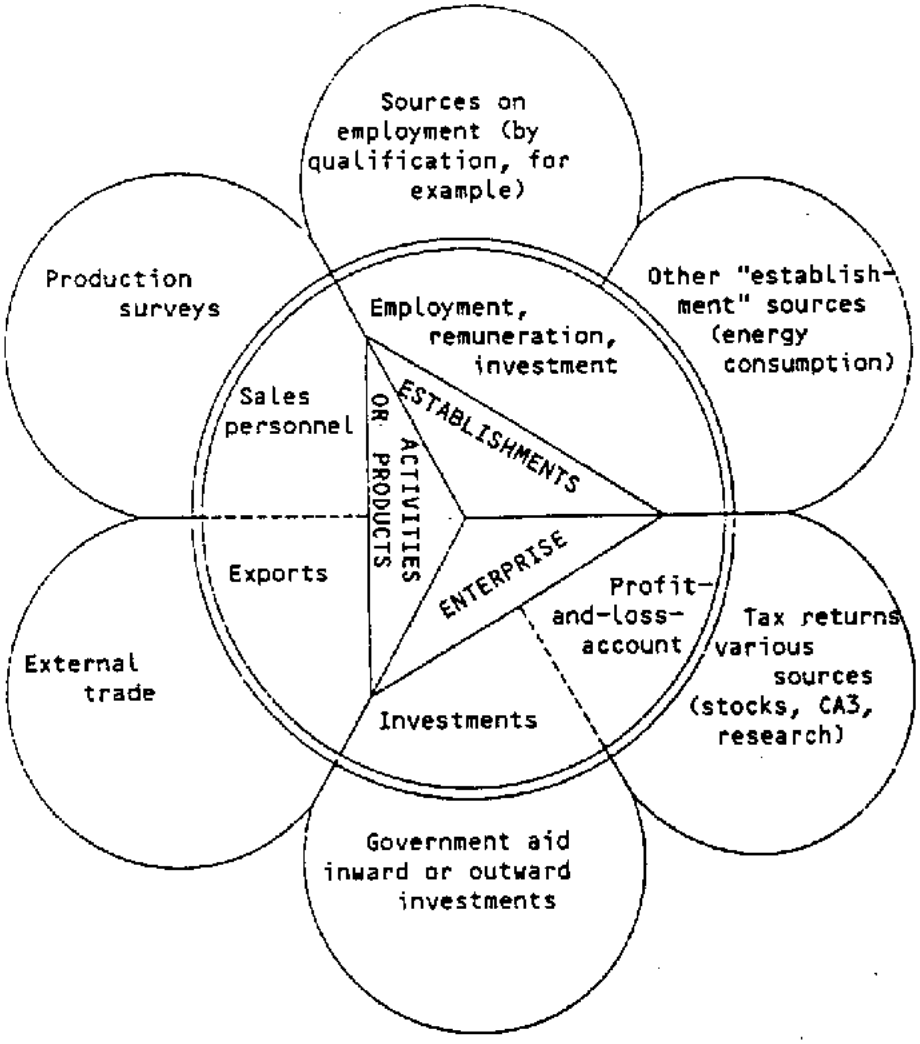


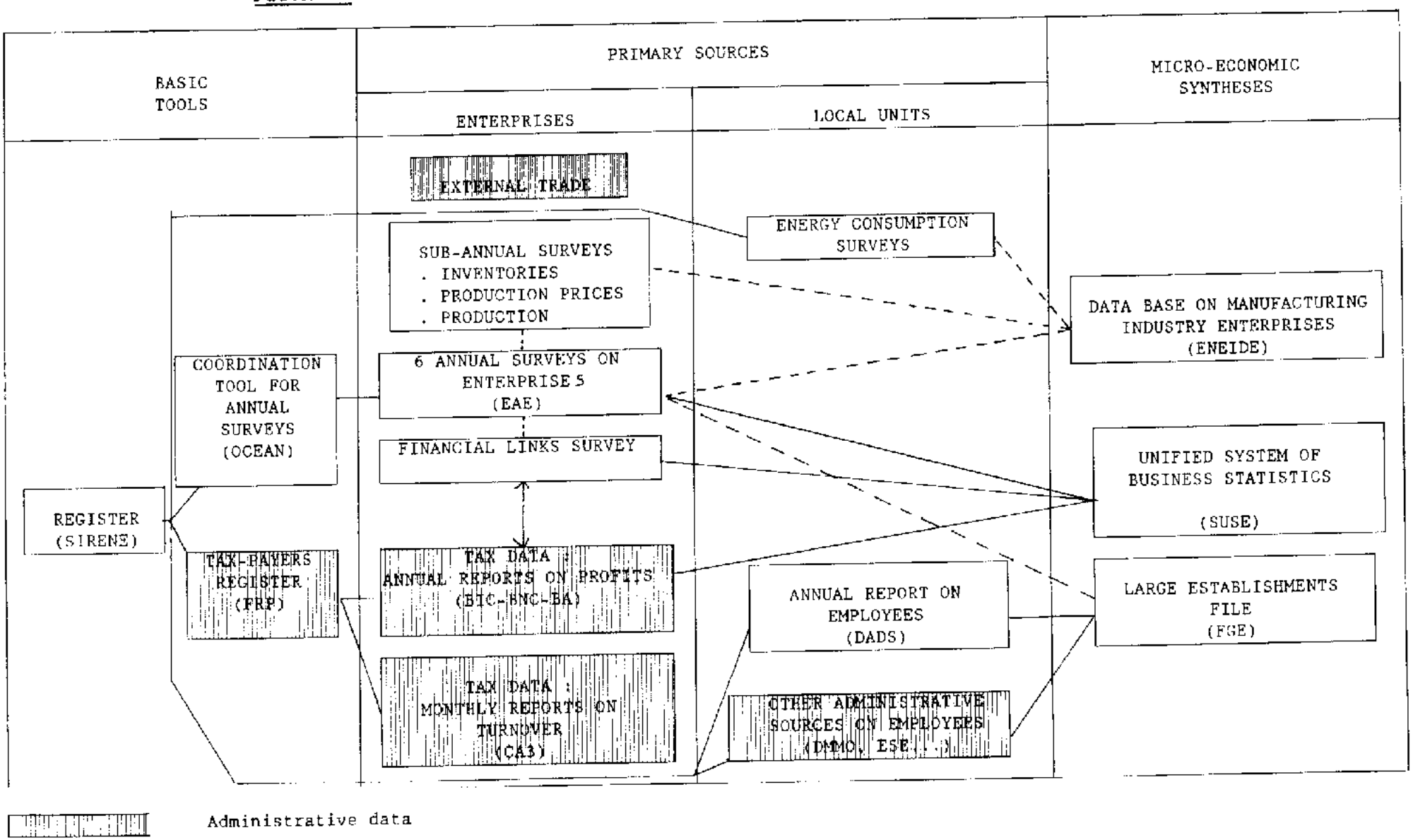
FIGURE 2

THE ANNUAL SURVEY ON ENTERPRISES (EAE) AND CONNECTED FILES



NOTE: The disk of the daisy describes the EAE (statistical units and associated variables); the petals represent connected sources.

FIGURE 3 - MAIN TOOLS AI ARTICULATIONS OF BUSINESS STA STICS SYSTEM



Administrative data