

**21<sup>st</sup> Meeting of the Wiesbaden Group on Business Registers - International Roundtable on Business Survey Frames**

**Paris, 23 – 27 November 2008**

**Bruno Erbisti Garcia  
Brazilian Institute of Geography and Statistics - IBGE  
27 November 2008**

**Parallel Session 7a - BIICS Countries Experience and Response**

1. Abstract.

The Central Business Registers (CEMPRE) of the Brazilian Institute of Geography and Statistics (IBGE) play an important role in all economic sectors in the Institute. It is mostly updated by two types of sources: The structural economic surveys conducted annually by the Institute (sectors of Manufacturing, Construction, Trade and Services), and administrative registers of the Ministry of Labor and Employment, made available annually through the *Annual Report of Social Information* (RAIS), and also the *General Register of Employment and Unemployment* (CAGED). The information stored on CEMPRE is cumulative and all historical data are preserved for enterprises and local units in activity, or inactivity.

This large database of information is processed and formulated to generate Surveys Frames, where the survey samples are selected. Each year, enterprises and local active units in the country are identified in order to generate statistics, which generate an annual basis called Annual Business Register (BR) Frame. Several statistics are generated from the Annual BR Frame information that are currently available in two publications: the Annual Statistics from BR, which examines all sectors of the country, and Annual Demography of Companies, which shows the movements of births and deaths of enterprises. In addition, other studies such as "Private Foundations and Non-Profit Organizations" (FASFIL) are released without regularity since the CEMPRE also includes the universe of non-profit organizations. At this time, a broad reform in computing systems is being held in software and databases, together with the final deployment of the new Brazilian Standard Industry Classification - CNAE 2.0 and the condition of activity, which defines a new statistical concept of economically active companies in the Institute.

2. The CEMPRE in Figures.

The CEMPRE is a bigger repository database containing about 11 millions of enterprises, including active and inactive companies. All economic data has been preserved over time, which has generated a historical database with more than 35 millions of records. Annually, we extract the annual BR frame containing the active companies. In 2006, base-year at last official publication, we had approx. 5.7 million active companies and 6.1 million local

active units. These companies contribute approx. 45 million employees to the whole country. This helps us to understand the complexity to manage CEMPRE.

### 3. History of the Central Business Registers.

The database of the Central Business Registers was established in 1991 and had its first update of records with the 1985 economic census. In 1995, aiming to capture a newly implemented economic classification and to provide updates of further details, an extensive survey was carried out that investigated about 400,000 companies. In parallel, the information from the records of RAIS began to compose the CEMPRE from the 1994 base-year. With the closer relationship between the Ministry of Labor and IBGE, it has been possible to receive data from RAIS annually since then. At that time, difficulties were noticed in sustaining the economic model that had the realization of the economic census as the basis for the production of economic statistics. Moreover, the fact that administrative records and CEMPRE use the same key identifier and the same activity, enabled better inter-relationships between records within and outside the institute. Since then, we have incorporated the results of Caged (other administrative records of the Ministry of Labor and Employment) and annual surveys of the Institute. In 1998, the first publication of 1996's results was made with this new reference model, which has been enhanced with improvements in treatment and data quality.

### 4. System's Model.

In the current model of the CEMPRE system, implemented last year, the database is basically composed of four different tables: the enterprises table, the local units table, and the two "hot tables". The enterprises and local units tables contain all historical records and all changes made in the companies so far. The "hot tables" of enterprises and local units have only the more current information. This prevents access to a database that is too large and unwieldy for use, and instead provides only the most current and relevant information. All economic Surveys Frames and the Annual Business Register Frames are extracted from these hot tables.

### 5. Sources of Update.

#### 5.1 Administrative Records.

The main source of updating is the 'Annual Report of Social Information' (RAIS). This form is filled in by all local units of companies in the country every year and aims to control labor and social security. Although this type of report contains all the information of companies and local units, it prioritizes the information of employees and salaries and does not provide information about taxes and revenues. Its advantage is its broad coverage of units in the country since it is mandatory for all companies in the country. The penalties for not

filling it in range from fines to the prohibition of the company to pursue activities. Caged, which is also used to update the CEMPRE, is a form that must be completed monthly by the companies with any movements of hiring and firing of staff. Because of its cyclical quality and because it contains fewer details, only new businesses and new local units are included in CEMPRE via Caged. Despite the frequency of Caged being monthly, the inclusion of new units in CEMPRE occurs every six months.

## 5.2 Annual Surveys of Manufacturing, Construction, Trade and Services.

The IBGE's economic surveys are conducted through the criterion of probability sampling and as such, have a lower coverage than administrative registers. The sample design of surveys is such that it contains a stratum in which all the companies are investigated. These companies are the largest of each sector of activity whose cut-off limit is determined based on the number of employees. The updates of surveys are done annually, and always take precedence over the other administrative sources. In addition, they contribute other attributes that are not present in administrative records as revenue of businesses. This process of feedback from CEMPRE allows the monitoring of the largest companies in the country and the consequent quality assurance records. The annual surveys from the base-year 2006 investigated about 211,000 enterprises that, although few in the numbers, are responsible for about 80% of working employees, of the total Brazilian economy in sectors of manufacturing, construction, trade and services.

## 5.3 Update Registers via SIMCAD

Since October 2007, telephone interviews with the enterprises that have had significant changes of address, economic activity, or any possible error in the fulfillment of administrative forms. This consultation conducted by telephone through Computer Assisted Telephone Interview methodology (CATI), aims to ratify or correct the information identified as suspected errors. This work, named SIMCAD, has been divided into projects with specific purposes of correction, and depends on the type of problem found in the worked database.

Source	Coverage	Frequency	Subject	Advantage	Drawbacks	Data Quality
RAIS	Wide (approx. 99.9% of active records)	Annual	All local units of the country	Mandatory, generates a broad coverage.	Economic classification self attributed, without descriptions. Doesn't have information about taxes and revenue	Regular. Focusing information about employees, not BR.
CAGED	Wide (more than 90% of registers)	Monthly, biannual receipt	All local units that have admitted or resigned staff	Mandatory and monthly.	Little information about BR.	Regular.
SIMCAD	From 0,4% to 3% of active records	It depends (project), varying from 3 to 8 months.	Enterprises and local units suspected of errors.	Targeted for criticism and correction of forms. Prioritizes most important economically	Contact with local units by telephone. Is not always in contact with the companies themselves, sometimes only with the accountant.	Good
Annual Surveys	Approx. 5% of active registers	Annual	Enterprises in manufacturing, construction, trade and services.	Biggest enterprises. About 5% of enterprises, representing approx. 75% of employment.	Most of them don't return information about local units	Good

**Table 1: The different sources that ensure quality of CEMPRE, and their characteristics.**

#### 6. Interactivity, technology and tools for use in database.

Since 2006, all systems and applications of CEMPRE have been going through a complete revision, involving the deployment of information technology for a project of creating a new ORACLE based bank which works in a dedicated machine server, connected to an SAS® server via high-speed connection. The software, now under Windows® platform, allows us to make updates, inclusions and exclusions of enterprises and local units, in addition to a fast-search system which retrieves information about any enterprise or local unit in the database. It also allows the generation of lists of records on demand, and helps all work of criticism and correction in the BR Section. This new system allows access to any of IBGE's users, in all States in the Country, and in all the

municipalities what the IBGE has an agency. We have been working to integrate the processes of annual surveys with CEMPRE, and must work on finalizing this system at least by the end of 2009.

## 7. CNAE 2.0 and Double Classification.

One of the key attributes of the Business Register is the classification of economic activities. The classification that has been used in the national statistical system and public administration as a whole is the *Brazilian Standard industry Classification - CNAE*, which was developed in 1994 in line with the UN classification, the ISIC Rev3. The CNAE is fully comparable with ISIC up to the division level, two digits compliance. Since was implemented, the CNAE has followed the structural revisions promoted in the ISIC. Thus, in 2002, CNAE 1.0 based on ISIC Rev3.1 was implemented and in 2007 the new structure of CNAE 2.0 was implemented, and has reflected the revision 4 of the ISIC. For the implementation of CNAE 2.0, a wide discussion was promoted in the *National Classification Commission*, which is headed by IBGE. A timetable was created in which the BR section would be implemented it at the beginning and the national accounts at the end, with a provision to incorporate it in 2010. To do that, it was necessary to develop a mechanism to register in double coding until all the surveys and the national accounts migrate their systems to CNAE2.0. We have been receiving the administrative registers encoded in CNAE 2.0 since last year, and have been working to improve the quality of these records. The system of self-attribution of CNAE in administrative registers actually is the biggest problem for us, because this system doesn't guarantee the best quality. Moreover, the changes in the CNAE structure caused a great impact on the classification, as we see in the table below:

	CNAE 1.0 (ISIC Rev 3.1)	CNAE 2.0 (ISIC Rev 4)
Sections	17	21
Divisions	59	87
Groups	223	285
Classes	581	673
Subclasses*	1183	1301

**Table 2 – Comparison between different economic classifications (CNAE 1.0 and CNAE 2.0).**

\* The lowest level of CNAE used by the IBGE is the class level. The use of the CNAE subclass today is limited to administrative records and was developed to meet the fiscal needs.

The details presented were responsible for divisions in the new classification, in up to nine different codes, as we see in the table following:

1code x N codes	Number of enterprises	Number of employees
1	50,9%	61,1%
2	13,0%	13,5%
3	16,1%	11,3%
4	2,4%	5,0%
5	12,9%	5,9%
6	0,7%	0,4%
7	<0,1%	<0,1%
8	0,4%	0,7%
9	3,6%	2,0%

**Table 3 – Relationship between CNAE 1.0 and CNAE 2.0, and impacts in number of companies and employees. (Source: IBGE, Annual BR Frame, 2006)**

Our current tasks are concentrated in the verification and validation of CNAE 2.0, which involve different processes, what includes:

- CNAE 2.0 verification of the most important enterprises through annual surveys.
- Verification of the codes involved in changes, through SIMCAD.
- Studies and analysis of CNAE 2.0 for the publication of the first results in 2009.

#### 8. Condition of Activity.

In the process of recasting the CEMPRE, a new attribute for the identification of active and inactive units was created: The '*condition of activity*'. In the old model, the concept of registration status was used. This old concept was based on combining RAIS of the base-year, with the registration status of the Federal Revenue Agency, and annual surveys of the IBGE. Because companies in Brazil are required by law to fill in the RAIS, Caged and Income Tax information, this requirement gives a broad coverage of companies and local active units in the country, but cannot guarantee that they are in fact working.

The administrative cancellation of registration of a company or a local unit in Brazil is very costly and bureaucratic, and in most cases, inactive companies are kept open, just to avoid the work and the cost of closing. Thus there are a lot of businesses and local units that although they are inactive, continue to deliver the RAIS, Caged and Income Tax, even without any economic activity. This problem causes difficulties for the evaluation of businesses and local units operating in the country. For this reason, it has developed a new concept to identify companies that are actually developing economic activity. This concept replaces completely the old concept of registration status, and takes into account a lot of information in considering the enterprise an active or inactive company. One of the pieces of information refers to a new question on the RAIS form, at the request of the IBGE. In this item, the local units answers to the following question: "Was this local unit

engaged in some activity last year?" This response, combined with others information determines whether a unit is active or not. The other combined details are, basically, the number of people employed, and if the company has responded to any of IBGE surveys. It is estimated that the reduction in the number of active companies in BR frames is around 25%. This reduction occurs basically on micro enterprises, i.e., those with up to 4 persons employed, with the greatest concentration in those who have no employees. It is noteworthy that while number of local units it represents 25% of total RAIS, in terms of number of persons employed the number is less than 0.03% of total RAIS.

#### 9. Future plans and projects in progress.

For the year 2009, we have the important task of disseminating statistical results in CNAE 2.0 for the first time. For this reason, we must finish the studies in CNAE 2.0 and ensure the quality of these results. Moreover, our task of implementing the new system was not totally completed. We are going to integrate CEMPRES and Surveys, and automate and standardize the entire process of generation of Survey Frames and the Annual BR frame. Our future plans include the creation of a history of succession to the companies, and introducing geodetic coordinates in BR. A project even more important, and not less ambitious, involves the creation of a table of Groups of Companies. This is a difficult task because the administrative registers included in the CEMPRES do not currently involve such information. This will require greater integration with all 27 states of the country.