The use of administrative data for the production of official economic statistics in Brazil - current situation and challenges for the future

Sidnéia Reis Cardoso
Fundação Instituto Brasileiro de Geografia e Estatística
sidneia.cardoso@ibge.gov.br

Ana Rosa Pais Ribeiro
Fundação Instituto Brasileiro de Geografia e Estatística
anarosa.ribeiro@ibge.gov.br

Abstract

The objective of the article is to present how administrative data has been used in the development of official economic statistics in Brazil since the second half of the nineties and the uses planned for the near future. In the beginning of the nineties, the Brazilian Statistical Office (IBGE) due to budget cut down had no perspective of developing economic census. IBGE has then realized that it was necessary to revise the organization of the Brazilian economic statistical system, in order to produce at a smaller cost an integrated and diversified set of economic statistics. The program of modernization of the economic statistics system then implemented by IBGE, was based in the transition to an integrated system of economic surveys, having as it's core the Central Business Register, which is systematically updated by administrative data. From this Register is extracted information for the drawing of the sample of each sectorial survey. It will be also reported another important experience of using administrative data, which is the use of the economic-fiscal information from tax authorities registers in the construction of the new series of the System of National Accounts published in March of 2007. This information complements the economic data about the institutional sector of private non-financial corporations that are not included in the scope of the economic surveys of IBGE. After ten years of the adoption of the current model of production of the economic statistics, IBGE considers necessary to enlarge the use of the administrative data with the aim of reducing the cost and the time for production of the surveys, as well as the response burden. In the conclusion of the article the problems and challenges for the increase of the use of administrative data by IBGE will be presented.

Key words: Administrative data, business register, economic statistics.

Background

The production of economic statistics, in Brazil, until the beginning of the nineteen nineties was based in censuses carried out every five years. The last Economic Censuses were referred to 1985. The reason why this program was interrupted was the unexpected budget cut imposed to IBGE at the end of the nineteen eighties and the beginning of the following decade. The lack of financial resources motivated IBGE

statisticians to make a deep reflection about the business economic surveys procedures in order to reduce costs and the time necessary to produce updated statistics and to better represent the recent changes in economic environment.

Two main concerns underlined these reflections. On one side, the questions related to standardization patterns, aiming to enlarge the articulations among national sources of data, including economic data coming from administrative sources and to promote the harmonization with international patterns, considering the growing need of international comparability. On the other side, the simplification of the data collection forms, aiming to lessen the burden to the respondents of statistical surveys.

Besides , it was also prioritized the supply of data to attend the needs of the Brazilian System of National Accounts, which up to that moment was not using the annual surveys statistics to compose its database.

Under these circumstances, by the middle of the nineteen nineties, IBGE has implemented a Program of Modernization of the Business Surveys replacing the Economic Censuses by an integrated system of annual sampling business surveys. On the same time, it has intensified the use of administrative data for the construction of a business register, periodically updated, to define the universe and the sampling design of each business survey. A new form of producing business statistics in IBGE has then started which has provided a coherent statistical database on the business sector of the economy, and more satisfactory answers to the demands of several users.

The Current system of business surveys and its articulation with administrative data

The first step in the implementation of the current system of business surveys was the consolidation of IBGE's Central Business Register – CEMPRE, regularly updated by internal and external sources, through the permanent use of administrative records. CEMPRE is the core element in the architecture of the current system of Brazilian business statistics.

At the same time IBGE has coordinated the development of a national classification of all economic activities with the cooperation of several governmental agencies, with the

objective of adopting it at national level. This standard classification has been adopted since 1995 by all the federal agencies dealing with registers of legal entities.

Nowadays the National Classification of Economic Activities - CNAE[1] has an extensive use: in all economic data produced in IBGE, as well as, in administrative files that require industrial classification. CNAE's structure was first developed according to ISIC Rev. 3, and in 2002, updated to version 3.1. A new version of the national classification, CNAE 2.0, derived from ISIC Rev.4 has started its implementation in 2007.

Another important step in the development of the current system was the deep revision of the annual business surveys. The main changes were:
  - the enterprise as the basic unit of investigation;
 - the coverage of each survey defined by the universe of enterprises in the segment;
- a sampling design including in the certainty stratum all large companies (the most complex units) and a sample stratum with representativeness for industries at class or group level for Brazil and regions.

Besides these changes, the definition of the content of the questionnaires has prioritized the needs of national accounting, specifically for the estimations of value added and it's composition, employment and gross capital formation.

The Annual Survey of Trade, Mining and Manufacturing, Construction and Services, carried on by IBGE since the year of 1996, are sample surveys, with the sampling frame stratified into two strata: a certainty stratum in which all the enterprises are investigated and another stratum with enterprises randoming selected, according to probability sampling approaches. Up to the reference year of 2006, the certainty stratum was defined considering the total number of employees in the enterprise - 20 for trade and services surveys and 30 for mining and manufacturing and construction surveys.

The implementation of this pattern of business surveys was possible due to the existence of the business register - CEMPRE as a central frame for the design of the samples of the business surveys. CEMPRE is regularly updated according to internal sources and administrative files.

---

[1] The detailed structure including the explanatory notes of the National Classification of Economic Activities - CNAE is available at : http://www.ibge.gov.br/concla/default.php

Each unit in CEMPRE is identified according to four basic elements: the location (address), unique legal identification (CNPJ[2]), the economic identity (CNAE's code) and the size indication (total number of employees and salaries paid). CEMPRE covers all legal entities formally constituted in the national territory.

CEMPRE is updated merging information from the following sources:
- two administrative files of the Labor and Employment Ministry[3]: the Social Security Administrative Record (RAIS), used for annual update and the General Register of Employees and Unemployed (CAGED), used for every six month updates;
- the National Register of Legal Entity - CNPJ from of the Ministry of Finance;
- annual business surveys from IBGE.
 The information from IBGE's business surveys prevails over the others.

 Agreements with the Labor and Employment Ministry since 1995 guarantees IBGE's access to the two main administrative databases used in CEMPRE's update: RAIS files are received in IBGE with gap of approximately eight months in relation to the year base and CAGED files are received monthly.

Administrative data originated from RAIS come exclusively from the establishments that answered the form named RAIS-EMPLOYER in the year base, including basic register information (name, address, unique legal number) and economic information on number of employees and salaries paid. CAGED's data used in CEMPRE refers to the new enterprises and local units created throughout the current year and not included in RAIS file received in the same year, due to the above mentioned gap. Register information from CNPJ come included in RAIS files, and are used as indicators of the activity or inactivity of the business.

Data treatment is systematically undertaken for the maintenance of the quality of CEMPRE comprising the editing of the register data and the checking on the homogeneity of concepts and the activity code, in a time approach (long term series) and in analysis for each year data. The nature and quality of CEMPRE's content –

---

[2] It's a numeric key with 14 digits of the National Register of Legal Entity, a register managed by the Federal Revenue of Brazil, Ministry of Finance.

[3] More detailed information on RAIS and CAGED can be found in the site of the MTE http://www.mte.gov.br/rais/default.asp and http://www.mte.gov.br/caged/default.asp.

updated register's information on name, address, legal number, activity code, employment and salaries paid for each of all legal units in the country – allow it's multiple uses as it supports data reports at the lowest level of economic activity defined by the national activity classification and at local level. Statistics based on CEMPRE's database are yearly published as well studies on business demography.

At the present time CEMPRE is updated to 2006 with information from the administrative registers and from business surveys. It covers a universe of 5,7 million of active legal units, among which 89,8% are enterprises, 9,9% are non-profit entities and 0,3% are public administration agencies.

Starting at the reference year of 2007 CEMPRE's set of data has been enlarged to include a new variable, the revenue range of each enterprise. Data on the revenue range together with the number of employees and salaries paid will allow a better classification of the enterprises by size (small, medium and large). This information comes from the business income tax files received after a cooperation agreement between the Tax Authority and IBGE.

The use of administrative data in the new series of the Brazilian System of National Accounts

The latest revision of the Brazilian System of National Accounts, 2000 year base, was announced by IBGE in March of 2007. The previous version of Brazilian National Accounts , 1985 year base, was based in the 1993 System of National Accounts, revision 3. The Production Account estimations of the previous version were based on data from the 1985 Economic Censuses. For the following years they were done through extrapolation of the 1985 data using volume and price indices.

The 2000 year-base revision included the updated of concepts and definitions according to the latest international recommendations on National Accounts, and the use of a new structure for it's activity categories based on the National Classification of Economic Activities - CNAE 1.0, a derived classification of the International Standard Industrial Classification – ISIC version 3.1. The industry categories used in the 2000 year-base revision presents a higher detail for activities and products in the Supply and Uses Table. The previous version presented results for 43 industries and 80 products. The present version presents results for 56 industries and 110 products. In

relation to the Institutional sectors of the Integrated Economic Account the present version presents a separate account for the Non-profit Institutions Serving Households.

One of the main improvements introduced in the new version of Brazilian Account was related to the quality of the database. The latest revision has incorporated the results of the business surveys for mining and manufacturing, services, trade and construction, as the main sources for the estimates of the current values of the annual production account. The new version also started to use data from income tax files received from Tax Authorities as the main source for the construction of a complete set of accounts (from production to capital formation) for the Non-financial Business as well as for the Non-profit Institutions Serving Households sectors. Data from income tax were used also for the estimates related to activities not covered by IBGE's business surveys, such as electricity, health and education activities.

An agreement between IBGE the Tax Authorities, since 1998, permits the access to annual business and other legal entities non-individualized economic data coming from income tax files. These data are received with a gap of approximately 15 months in relation to the year of reference.

For 2000, the base-year of the current Brazilian National Accounts, an extensive analysis of data consistency between statistical sources (IBGE's business register and business surveys) and administrative data from Income Tax files covering Non-financial Business, was developed in two approaches: one focusing business demography and another, comparing values of National Accounts aggregates constructed according to data from each source.

The study on the business demography has matched the set of enterprises in IBGE's register with the ones in Income Tax files. Three groups of enterprises have been identified: the enterprises belonging to both sources; the enterprises only included in CEMPRE; and the enterprises included only in Income Tax files. Comparison between these two databases has allowed the inclusion of enterprises out of IBGE's surveys in the estimates of Brazilian Gross Domestic Product guaranteeing that the whole universe of business was considered in National Accounts estimates of the Non-financial Business aggregates

The study comparing values of National Accounts aggregates based in IBGE's business surveys database, on one side, and in Income Tax files database, on the

other, at a detailed level of activity, has pointed out no substantial diference between the two sources. The new born enterprises included in the Income Tax's database and not in IBGE's surveys, due to CEMPRE's temporary gap, is the main reason to explain them.

For a group of large enterprises, selected according to the value of annual gross revenue, the study has compared the aggregated value from the Production Account and Income Generation Account and has concluded about the fitness of Income Tax files data for the purpose of the construction of the aggregates of the Non-financial Institutional Sector.

Besides Income Tax files, other administrative records are also used in the construction of the National Accounts database, such as the Central Bank's records on Financial Sector and data on Pension Funding activities, from Ministry of the Social Welfare's files.

The last revision of the Brazilian National Accounts System has introduced an upgrade in the comprehensiveness of the system database and consequently in the quality of the estimates of National Accounts aggregates through an enlarged use of administrative data in articulation with IBGE's statistical data. But to guarantee the effectiveness of this performance it is necessary to assure the systematic access to the administrative data being used. The absence or delay in receiving data from these sources may cause serious discontinuities in the works of the Brazilian National Accounts mainly the ones related to the Integrated Economic Account.

Final considerations

In the last fifteen years IBGE has stepped toward a more intensive use of administrative data which has contributed to increase the rationality of it's business statistics system and the quality of National Accounts estimates. However, it is necessary to enlarge the use of administrative data in order to go further towards reducing costs, time and the response burden of it's business surveys.

In relation to the Income Tax files IBGE aims to have access to individualized data concerning the complete set of financial and economic data of the business units. This would allow IBGE to go further in the rationalization of its system of business surveys.

Data of the small companies would come from the Income Tax records instead of being collected by IBGE, reducing, in this way, costs, time and response burden. More resources could then be allocated to improve the collection of data in big and more complex companies with multiple activities and locations.

However the achievement of that goal faces the restrictions of fiscal secret determined by the Federal Constitution of 1988. IBGE intends to go on promoting actions to receive individualized data on business units from Income Tax files, based in the argument that the statistical secret law guarantees the secret of any individual information.

IBGE has also been promoting articulation with other governmental agencies with the purpose of going further in the articulation between statistical database and administrative data. For instance, with National Regulatory Agencies in order to obtain financial and economic data of the enterprises under their control and with External Trade Agencies; in order to obtain records of imports and exports by enterprise.

The growing demand for high quality and up-to-date economic statistics will be better attended by IBGE, at a smaller cost and with more efficiency, with the strengthen of the integration of statistical and administrative data.