Introduction

1. OECD statistics on businesses

The OECD collects and develops a wide range of statistics on businesses and business activity. This publication presents the *Structural and Demographic Business Statistics (SDBS) Database*, *i.e.* the data collection by the Statistics Directorate relating to a number of key variables, such as value-added, operating surplus, employment, and the number of business units, broken down by industries and size classes.

Other OECD Directorates also collect important information relating to business activities. The Science Technology and Industry Directorate, for example, collect information on business research and development expenditure and innovation broken down by industrial sector and size class within its Analytical Business Enterprise Research and Development database (ANBERD). This database is part of the broader Structural Analysis Database (STAN) which additionally provides 2-digit industry breakdowns of economic activity (employment, production, value added, etc.) and input-output tables. The Directorate for Employment, Labour and Social Affairs maintains a large database of labour force statistics, which does provide information on employment in different industrial sectors as well as information on the numbers of self employed. The OECD's International Energy Agency collects information on carbon dioxide emissions and energy consumption by industrial sector for example. This is not an exhaustive list though; other Directorates also collect information related to businesses, such as financing and taxation.

Three datasets combine to form the *Structural and Demographic Business Statistics* (SDBS) database: Structural Statistics for Industry and Services (SSIS), Business Statistics by Size Class (BSC) and Business Demography (BD). The different variables in the SDBS are measured on an annual basis.

2. Structural and Demographic Business Statistics (SDBS)

The SSIS database provides information on economic variables broken down by 4-digit International Standard of Industrial Classification Revision 3 (ISIC Rev. 3) industries. The BSC database provides the same industry level of information as the SSIS database but this also broken down by size classes of businesses. In theory these two datasets are consistent and complementary, although in practice the datasets are considered separately because the time series for SSIS data are typically longer and the coverage of variables in the SSIS database is more comprehensive than the BSC database. Further on, the two datasets are not entirely consistent for some countries, reflecting partly different information sources and allocation methods and partly the different times when the datasets were compiled. As described in the country notes and in the examples that follow, for some countries it has been possible to make adjustments to the SSIS dataset to improve its internal consistency that cannot be replicated in the BSC database.

The BD database contains information on variables such as birth rates, death rates, survival rates, high-growth enterprises and gazelles for most OECD countries. As there is considerable policy interest in this type of information, the OECD Statistics Directorate has started to collect information on business demography data using harmonised questionnaires and definitions for variables. This is done in the framework of the OECD-Eurostat Entrepreneurship Indicators Programme (EIP), launched in 2006 by the OECD to build internationally comparable statistics on entrepreneurship and its determinants. As the BD database is recent, data are at present available for a few years only. In 2007 a *Eurostat-OECD Manual on Business Demography* has been published presenting different definitions of policy-relevant statistics on entrepreneurship and business demography and their empirical measurement.

3. Process of data collection

The United Nations Statistical Commission, at its 27th Session, recommended that the OECD, in collaboration with the United Nations Industrial Development Organisation (UNIDO), take responsibility for the collection and dissemination of world-wide industrial statistics. The collection of SSIS and BSC statistics for non-European Union countries has been organised through a joint OECD/UNIDO questionnaire.

For BD statistics, the OECD and Eurostat cooperate closely with National Statistical Offices for the collection of data through an annual questionnaire.

The OECD works with Eurostat in collecting the data for the SDBS database (SSIS, BSC, BD), as part of an agreement between the OECD and Eurostat designed to reduce the response burdens of EU member countries. Data for EU countries, including Norway, are sourced directly from Eurostat while data for non-EU countries are directly collected by the OECD.

The data received are checked and, if necessary, converted to SDBS standards (harmonised size class, ISIC Rev. 3). Quality checks are performed in order to be able to assess the consistency of the data, to flag potential outliers, and possibly, to remove or correct the data.

4. Characteristics of the SDBS database

Included variables

The SSIS and BSC databases provide, uniquely, comparable information on the following economic variables broken down by 4-digit ISIC Rev. 3 (where the coverage varies by country and database) and employment size classes in the case of the BSC database (1-9, 10-19, 20-49, 50-249, and 250+), for all OECD countries:

- Turnover or gross premiums written.
- Production at producers' prices and/or factor costs.
- Value added at basic prices and/or factor costs.
- Gross operating surplus.
- Total purchases of goods and services.
- Change in stocks of goods and services.
- Purchases of energy products.
- Gross investment in tangible goods.
- Gross investment in land.

- Gross investment in existing buildings and structures.
- Gross investment in machinery and equipment.
- Employment, number of persons engaged and number of employees.
- Employment, number of females' employees.
- Employment, number of employees in full time equivalent units.
- Hours worked by employees.
- Compensation of labour, all persons engaged and employees.
- Wages and salaries, all persons engaged and employees.
- Other employers' social contributions, employees.
- Number of enterprises and/or establishments.

The Business Demography database contains information on the following variables for most OECD countries. Indicators are broken down by industry using ISIC Rev. 3 and, for some of them, by employment size-class:

- Birth rate.
- Death rate.
- Survival rates.
- Employment by age of enterprise.
- Rate of High-growth enterprises.
- Rate of gazelles.

Industry Coverage

This publication provides data on industries according to ISIC Rev. 3. A full description of this classification, together with descriptions of concordance relationships between ISIC Rev. 3 and national and other international classification systems can be found at the United Nations Statistics Division's website.[1] Annex A provides a description of the 2, 3 and 4-digit industries for which the OECD requests information.

At its 37th Session, held in New York in March 2006, the United Nations Statistical Commission approved the revised International Standard Industrial Classification of All Economic Activities (ISIC), Revision 4 as the new international standards for activity. Details of this classification system can also be found at the UN website referred to above. The SDBS database will be updated to this new system in parallel with moves made by OECD member countries starting in 2010.

Size Classes of Industries Covered

The OECD Statistics Directorate requests breakdowns of industries into the following employment size classes, where the figures reflect in theory, although not always in practice, the numbers of persons engaged:

1-9, 10-19, 20-49, 50-99, 100-199, 200-249, 250-499, 500-999, 1000+

Unfortunately, it is not always possible for countries to provide information on this basis and, instead, countries provide data using the breakdowns relevant to the data sources available in their country. Data development in individual countries tends to follow policy requirements and the availability of information by size classes often reflects the needs of small and medium enterprise (SME) policy makers. Therefore, to improve the

usability of the database and to improve harmonisation and encourage country comparisons, this publication focuses on 5 size classes where the data across countries and variables can be most closely aligned. These are as follows:

1-9, 10-19, 20-49, 50-249, 250+

Not all country information fits perfectly into this system however; for the United States for example it implies a size class breakdown of 1-9, 10-19 and 20+, and, so, for some countries, there is necessarily some divergence from these target size classes. As such the database refers to 5 size classes that are country specific. To reinforce the point that there may be differences in the coverage of businesses within size classes across countries, these 5 size classes are referred to as NSC1, NSC2, NSC3, NSC4 and NSC5; where the NSC refers to national size class.

Table 1. **OECD BSC size classes: breakdown by number of persons engaged**

	ISIC sector	Variable	OECD BSC size classes: Number of persons engaged				
			NSC 1	NSC 2	NSC 3	NSC 4	NSC 5
Australia[1]	All	All	0-9	10-19	20-49	50-199	200+
Austria	All	All	1-9	10-19	20-49	50-249	250+
Belgium	All	All	1-9	10-19	20-49	50-249	250+
Canada	All	All	1-9	10-19	20-49	50-249	250+
Czech Republic	All	All	1-9	10-19	20-49	50-249	250+
Denmark	All	All	1-9	10-19	20-49	50-249	250+
Finland	All	All	1-9	10-19	20-49	50-249	250+
France	All	All	1-9	10-19	20-49	50-249	250+
Germany	All	All	1-9	10-19	20-49	50-249	250+
Greece	All	All	1-9	10-19	20-49	50-249	250+
Hungary	All	All	1-9	10-19	20-49	50-249	250+
Iceland	All	All	1-9	10-19	20-49	50-249	250+
Ireland	All	All	3-9	10-19	20-49	50-249	250+
Italy	All	All	1-9	10-19	20-49	50-249	250+
Japan	All	Gross investment in tangible goods, purchases of energy products, wages and salaries for employees	. .	. .	30-49	50-249	250+
	All	Other	4-9	10-19	20-49	50-249	250+
Korea	10-41	All	5-9	10-19	20-49	50-199	200+
	60-63	All	1-9	10-19	20-49	50+	. .
	Other	All	1-9	10-19	20+	. .	. .
Luxembourg	All	All	1-9	10-19	20-49	50-249	250+
Mexico	All	All	0-10	11-20	21-50	51-250	251+
Netherlands	All	All	1-9	10-19	20-49	50-249	250+
New Zealand	All	All	1-9	10-19	20-49	50-99	100+
Norway	All	All	1-9	10-19	20-49	50-249	250+
Poland	All	All	1-9	10-19	20-49	50-249	250+
Portugal	All	All	1-9	10-19	20-49	50-249	250+
Slovak Republic	All	All	1-9	10-19	20-49	50-249	250+
Spain	All	All	1-9	10-19	20-49	50-249	250+
Sweden	All	All	1-9	10-19	20-49	50-249	250+
Switzerland	All	All	1-9	10-19	20-49	50-249	250+
Turkey	All	All	1-9	10-19	20-49	50-249	250+
United Kingdom	All	All	1-9	10-19	20-49	50-249	250+
United States	All	All	1-9	10-19	20-99	100-499	500+

1. Size classes based on the number of employees

The table above shows for each country the actual size class of businesses within each of NSC 1 to NSC 5. As can be seen in most cases they align quite close to the 5 target size classes mentioned above, but not always. In fact a strict comparison across countries is only possible if the focus is on businesses with less than 20 persons engaged and businesses with 20 or more persons engaged; and even here the comparisons are restricted by the exclusion of some small businesses at the lower end of the less than 20 threshold; the extent of which varies across countries. Data for Ireland for example exclude businesses with less than 3 persons engaged, and the US data refers only to businesses with employees.

5. Issues in using the SDBS database

The Structural and Demographic Business Statistics (SDBS) database is designed to provide a description of the economic structure of businesses by their industrial classification and by their size. The primary advantage of the database compared to other datasets is that it provides detailed information on a wide range of economic variables at a very detailed industry level. The database therefore provides an ideal tool for users interested in international comparisons of the economic performance of industries at this detailed level. This is especially true where comparisons of business size classes are needed.

However, users should also be aware of some of the phenomena that need to be fully understood when using these datasets. Key amongst these concerns are comparisons over time. Since the scope and coverage of source information used in the SDBS database can vary over time, breaks in series often occur. Moreover, enterprises and establishments may change their industrial classification over time. Sometimes this will reflect the fact that the enterprise or establishment has change activity completely but it may also reflect small year-on-year changes.

This section presents some of the issues that can affect analytical applications of the database, such as comparability and changes in concepts. The metadata describing the sources and methods used in each country are provided at the end of the tables shown for each country. Detailed definitions of variable concepts are provided in Annex B.

Comparability of concepts and coverage across countries

Effort has been made to ensure that the variables presented are comparable across countries but on occasions this is not possible. Three areas, worthy of particular mention, where some issues of incomparability arise are:

- *Value-added*: The definition of value-added used in this publication is consistent with that used in the 1993 System of National Accounts (for a given price basis, factor costs or basic prices). However not all countries are able to provide estimates on this basis and typically, for these countries, the estimates are based on production minus the intermediate consumption of goods and industrial services only. This differs from the SNA definition which excludes the intermediate consumption of all services.
- *The coverage of businesses:* For most countries the main sources of information used in the compilation of structural business statistics are economic censuses, business surveys and business registers. These typically cover establishments and/or enterprises above a certain size limit (usually based on employees or turnover above a certain level) but these thresholds vary across countries. This is of particular relevance in the context of business statistics by size class.
- *The statistical unit used for businesses:* Many measures are available for the statistical unit used to define businesses and these vary across countries. Two key measures prevail in this publication. The first is based on the *enterprise* and the second on the *establishment*.

Comparability over time

For some countries the source data used to estimate the variables in this publication are not always consistent over time and neither are the concepts nor the coverage of businesses. For example, data may be based on survey information in one year and exhaustive census information in another, or the coverage of the same data source itself may change over time, for example a census or survey may include all establishments with more than 20 employees in one year and then all employing establishments thereafter.

Changes in concepts can also complicate the comparability of data over time. It is not always obvious however where concepts change but, on occasions, especially for sectors, such as the distribution sector (wholesale and retail) where turnover and production can be especially affected by valuation changes in estimates of changes in stocks, it is possible to identify these occurrences.

Users should also note that changes may occur for specific sectors only, reflecting the different data sources often used by statistical offices to measure activity in sectors as diverse as manufacturing, services and construction. This means that in some cases breaks in series occur in certain sectors only.

Time series analyses – Classification changes and consolidation

The points raised above also relate to time series applications but, fundamentally, users need to be aware that the database provides snapshots of enterprises or establishments at points in time. Both enterprises and establishments can change their industrial classification over time; sometimes this reflects the fact that the enterprise or establishment has changed activity completely but it may also reflect small year-on-year changes.

Additionally, if enterprises in two different industrial sectors merge, the new entity may re-appear in one single industrial sector only, even if the overall activities of the new entity are unchanged. Similar issues arise if establishments in different sectors merge or when other demographic events occur such as take-overs or break-ups for example.

Gross measures of activity such as turnover, production and intermediate consumption of goods and services within an enterprise may be recorded net of the enterprise's intra-transactions or gross; depending on the country or enterprise's practice or available accounting information. If, therefore, an enterprise breaks-up into two new enterprises but both new entities remain in the same industrial sector, gross production, turnover and intermediate consumption may rise, although, all other things equal, no changes will be seen in net measures such as value-added, operating surplus or labour and investment related variables.

Analyses of time-series data related to business size classes are also of considerable policy interest and similar issues of time series consistency can arise when businesses move from one size class to another as they grow or contract.

Confidentialised data

Sometimes data are 'confidentialised' by statistical offices at a relatively aggregated (ISIC Rev. 3, 2-digit) level. This can occur because a statistical office might desire to show all but one of the 4-digit groups within a 2-digit sector leading to the secondary suppression of the 2-digit figure to preserve anonymity in the confidentialised 4-digit entry (otherwise the confidentialised 4 digit cell can be determined). Additionally, because statistical offices will typically confidentialise industries according to the classification systems used in their countries and then use this information as the building blocks to arrive at ISIC

estimates, it is often the case that the numbers of industries that are confidentialised are proportionally higher when shown in ISIC than when shown on a national basis. This reflects cautionary prudence on the parts of statistical institutions. The process of ensuring that data are not disclosive when shown in a national classification system can be burdensome, particularly when legal penalties arise for accidental disclosure.

Coherence of sources

Statistical offices use a variety of sources to produce business statistics data and this variety can often cause comparability problems. In many countries, for example, different sources of information are used to estimate activity in different sectors. For example, manufacturing may be based on a survey of businesses with one or more employees whereas service sectors may be based on administrative data for businesses above some turnover threshold say. But the different sources can also relate to variables. Investment information for example may be sourced from dedicated investment surveys that may not have the same coverage, sample size or stratification as used for other variables. Another important example relates to information on employment, which, for some countries is based on labour force and not business surveys, which can compromise labour productivity calculations. Differences in sources are highlighted in the country notes.

Comparisons of BSC and SSIS statistics

The data requested for the BSC and SSIS databases are the same but differences exist in the coverage of variables and years. Typically, less information is available on the BSC database reflecting the greater data requirements involved in providing information at this level. Other differences exist however. For some countries the estimates of 4-digit totals for the same variable differ, reflecting the often different data sources used in the compilation of the BSC and SSIS datasets and also the coverage of these data sources. Timing is also a factor as for some countries the returns were provided at different times and so more up-to-date estimates may be reflected in one dataset than another.

Comparisons with other data sets

As mentioned, the OECD produces many other statistics related to businesses and their activities. This information can be used to supplement the SDBS database, but some care is needed in bringing the different OECD business datasets together. Although every effort is made using international standards, such as ISIC Rev. 3, to harmonise the various datasets, users should note that some differences still arise. On occasions this will be because different concordance relationships are, necessarily, used for different indicators to bridge the gap between national classification systems and ISIC Rev. 3. Sometimes it may reflect differences in the way business units are defined or sampled and on others it may reflect differences in the way similar concepts are measured.

One important dataset in the context of business statistics is national accounts data, where significant differences in concepts can occur. Economic indicators provided by the structural and demographic business statistics database typically provide more detailed information on industrial sectors than those provided by national accounts based statistics. This occurs because adjustments are made for the national accounts which, typically, can only be made at the more aggregated industry level. Four types of adjustments are commonly made.

- The first is driven by the need to achieve “exhaustiveness” in the coverage of economic activity such that estimates of gross domestic product (GDP) cover all market and

non-market activity. Examples of this activity are: imputed-rent for owner-occupied dwellings and the production of agricultural products used and consumed by households.

- The second reflects the coverage of structural business statistics. Typically these are sourced from census data, surveys or business registers, but these sources are not always exhaustive in their coverage of businesses. Since they often only collect information on businesses above a certain size threshold (which are often based on turnover or employment criteria), micro and small enterprises or establishments are not always covered in structural business statistics. Additionally, structural business statistics generally only provide information that is based on business returns. In countries with large informal economies, businesses often under-declare their real turnover, profits and employment say, or may evade registration on administrative sources altogether. National accounts estimates try to adjust for this. Finally, some countries do not include all state owned, public corporations or non profit institutions, in their structural business statistics data.
- The third reflects conceptual adjustments made by national accountants, for example "tips" paid to waitresses and waiters in a restaurant say are recorded as income received by the employing enterprise in the national accounts. Larger adjustments are made to correct estimates that do not conform to SNA definitions. For example, some countries do not include intermediate expenditures on services in their structural business statistics estimates of total intermediate consumption of goods and services which are, in these cases, instead included in value-added. Another source of difference in industry estimates of value-added reflects adjustments made for financial intermediation services indirectly measured (FISIM).[2]
- The fourth reflects more general adjustments that embody, in practice, all of the three above and are driven by the need to achieve coherence between business statistics and other sources of information that can be used in the compilation of GDP, such as household budget and labour force surveys.

National Accounts based estimates, therefore, provide for more comprehensive coverage of economic activities and should be used where the focus is on total GDP or the contribution of aggregated industries to economic activity, irrespective of the size of any particular establishment or enterprise. The SSIS and BSC databases should be used where the focus is on: a specific industrial sector or sectors; information on the sizes of businesses is important; and/or where the focus is on market activities.

These are not the only areas where differences between SDBS and national accounts based statistics occur however. Another important difference arises when the business units used to measure activity differ. Typically, in most OECD countries, estimates of activity by industry sector are based on the most homogeneous production unit available. The unit used varies across countries depending on the available data sources but most countries will try to measure activity for the national accounts using a local kind of activity unit (used in many European countries) or the closely related establishment unit (used in many non-EU OECD countries). Enterprise data, which are how most European countries present SDBS data, reflect a combination of one or more sub-enterprise units such as establishments or local kind of activity units, and, so, data presented by industries on this basis may be different (higher or lower) than comparable national accounts data because different groupings of sub-enterprise units will be used across industries. Indeed, even if the same groupings of sub-enterprise units exist in a given sector, differences can arise in

gross measures, such as production and intermediate consumption, reflecting the differing degrees of consolidation that exist within these sub-enterprise groups.

6. Examples of uses of the SDBS database

Examples of use of the SDBS database include analyses of structural change and industry dynamics, which are key for effective government policies on national competitiveness, entrepreneurship and innovation. The SDBS database helps policy makers and analysts to assess the structural changes in business activity and industrial dynamics over time. Another area of analysis for which the SDBS database can be efficiently used concerns the study of business size over time, more specifically the contribution of individual size classes to aggregate measures of economic performance.

Point in time comparisons

A strength of the SDBS database is its ability to provide a detailed description of the industrial structure of economies at a very detailed level, at fixed points in time, facilitating cross country comparisons. The database therefore lends itself to answering questions such as: What proportion of total employment is provided by small businesses? What is the contribution of a particular 4-digit industry to total value-added and how does this compare across countries? What were the unit labour costs by 4-digit industry and size class of OECD countries in 2006? Are small firms more or less profitable than large firms (where profit can be measured using gross operating surplus as a per cent of turnover)? How does the decomposition of value-added into its labour and capital components (gross operating surplus) vary by industry size? Are workers in small firms better paid than in large firms in some industries?

The database can also provide useful inputs into the development of other economic databases. For example, the OECD's input-output database contains information on intra-industry transactions at a relatively aggregated industry level but it is possible to incorporate SSIS and BSC data into the input-output framework using relatively simple assumptions and without distorting the overall input, labour and value-added structures of these SSIS and BSC industries. As such, inter dependencies between detailed industries and business sizes can be investigated, for example the dependencies of large businesses on small and *vice-versa* or alternatively the total direct and indirect dependencies of small businesses on export markets.

Changes over time

An area of considerable interest to policy makers and analysts concerns changes in business activity such as production, value-added and employment gains and losses over time. The SDBS database can provide an assessment of these changes but generally the comparisons should be performed over relatively short periods to avoid issues of incomparability arising over time. Demographic events like mergers, take-overs and break-ups may result in reclassifications of businesses, or parts of businesses, from one industry to another. The SDBS database accurately reflects these changes and users should familiarise themselves with the country notes to ensure that their analyses are not affected by breaks in series.

Another area of analysis for which the SDBS database can be used concerns the study of business size over time. Because it is not longitudinal (*i.e.* it does not track individual firms over time), the database cannot show the contribution small businesses make to

economic and employment growth: as (smaller) firms go from the start-up phase to some optimal size, they typically change from one size class to another.

Structural indicators

The information in the SDBS database can also be used to analyse more thoroughly the economic structure by describing the structural characteristics of industries and/or size classes. Combining the different variables in the database, the structural performance of individual industries can be illustrated by structural indicators such as:

- average firm size: total employment/number of enterprises, indicating the importance of scale (and scope) in industries;
- average wage cost: wage costs/employment, indicating the unit labour cost in industries;
- labor productivity (in nominal terms): value added/employment as an indicator of economic performance in terms of productivity and value creation;
- vertical integration: value added/production showing how much value of the production is created within the industry.

These indicators at a very detailed industry level additionally allow for a more fine-grained analysis of industrial dynamics. Productivity growth analyses for example focus often on relatively aggregated industries, thereby generally assuming that homogeneity exists within pre-defined industry groups. But this assumption might not hold, as changes in 2-digit labour productivity could be due to changes in the share of activities of the 2-digit industry's 4-digit components or because of variances in the contributions made by different sizes of firms (as well as changes that might be driven by entries and exits into the population).

7. Online data

The Structural and Demographic Business Statistics online database can be accessed *via http://dx.doi.org/10.1787/sdbs-data-en*.

8. Symbols and abbreviations used in the publication

. . Not available

| Break

ANZSIC: Australian and New Zealand Standard Industrial Classification

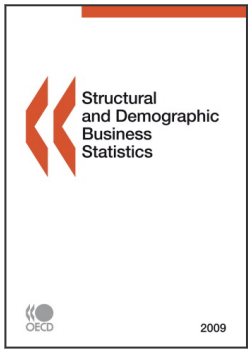
ISIC: International Standard Industrial Classification of All Economic Activities

NACE : European Classification of Economic Activities

NAICS: North American Industry Classification System

Notes

1. *http://unstats.un.org/unsd/cr/registry/regct.asp?Lg=1*.
2. FISIM reflects the value of services provided by financial intermediaries for which no explicit charges are made. For example, interest is payable by borrowers and by lenders to financial intermediaries, and so, put simply, FISIM captures the value of the property income receivable by financial intermediaries less the interest payable by them.



From:

Structural and Demographic Business Statistics 2009

Access the complete publication at:
https://doi.org/10.1787/9789264072886-en

Please cite this chapter as:

OECD (2010), "Introduction", in *Structural and Demographic Business Statistics 2009*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/9789264072886-2-en

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.