

Methods explained

Methods explained is a quarterly series of short articles explaining statistical issues and methodologies relevant to ONS and other data. As well as defining the topic areas, the notes explain why and how these methodologies are used. Where relevant, the reader is also pointed to further sources of information.

Business Structure Database¹

The Inter-Departmental Business Register (IDBR) for Research

Peter Evans and Richard Welpton

Office for National Statistics

SUMMARY

The Interdepartmental Business Register (IDBR) is a comprehensive database of UK businesses, drawn from administrative data sources. This article initially provides a short explanation of the IDBR. The focus of the article is then directed to the development of the IDBR as a research tool by staff at the Virtual Microdata Laboratory (VML). The resulting Business Structure Database (BSD) is now routinely used for academic and government studies.

The IDBR

The IDBR² is the sampling frame for surveys of business activity conducted by the Office for National Statistics (ONS) and also by other government departments. The register contains records of over 2 million businesses from all sectors of the economy. The only exceptions are organisations generating turnover below the Value Added Tax (VAT) threshold (currently £68,000 annually) and/or those which do not pay employees via pay-as-you-earn (PAYE) where salaries of £100 per week and over must be paid via PAYE. A business may be included if it pays a salary to an employee of over £100 per week but does not generate sufficient revenue to be registered for VAT, and vice versa.

The IDBR draws upon the following administrative sources:

- HMRC (information about businesses which are registered for VAT or PAYE are provided to the ONS)
- Dunn and Bradstreet (information about business ownership links, provided annually)
- ONS surveys (other surveys supplement the above sources by identifying new, and maintaining existing, business structures)
- Companies House (received quarterly)

Box 1 illustrates the IDBR structure for simple and more complex organisations and the information collected.

The IDBR limits research by a wide audience for two reasons. First, access is highly restricted due to the inclusion of confidential HMRC data. Second, it is difficult to perform historical analyses on these data. The register is updated at regular intervals but a regular referenced set of changes is not maintained. Businesses may experience various demographic events (e.g. mergers and acquisitions) throughout their life, making it difficult to build up a longitudinal picture of businesses over their life cycle. To resolve these issues the BSD is an annual 'snapshot' of the IDBR which is deposited

within the VML for the purpose of micro-data analysis (see **Box 2** for a brief description of the VML). Security controls around access and procedures for using the VML protect the confidentiality of IDBR data. Considerable effort has also been dedicated to improving the micro-data integrity of the BSD, enabling researchers to undertake longitudinal analyses of data from the IDBR.

The BSD

The BSD 'snapshot' is taken every March and includes data on enterprises and local units. Two BSD files are created, and contain observations for enterprises and local units. The consistency of IDBR reference numbers throughout time enables these BSD files to form a panel data set.

Variables

The number of variables found in the BSD is small relative to other data sources. However, the BSD is virtuous by its extensive coverage, since any organisation registered for VAT or PAYE is recorded on the IDBR.

Box 3 provides an overview of the variables contained in the BSD for enterprises and local units.

In addition, the local unit files contain a variable that identifies reasons for inactivity of the unit. Examples include 'ceased trading', 'change of ownership', and 'liquidation'.

Employment and turnover figures are derived from HMRC administrative records - PAYE or VAT returns respectively. The BSD includes a marker identifying origination.

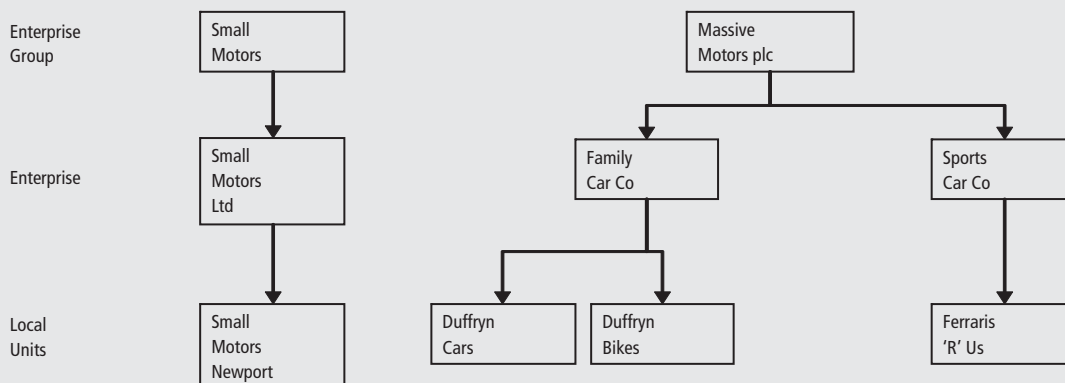
For various research purposes, researchers may use the extensive data available from the Annual Business Inquiry (ABI), including turnover. However, it should be noted that, particularly for smaller firms, turnover figures in the ABI are often imputed. For quality and consistency, all enterprise turnover data in the BSD are derived from at least one of the two administrative data sources.

Descriptive statistics

Not all of the enterprises observed in the BSD are 'active' - defined as an enterprise with at least one local unit for which live data is available. Live data may not be available if a company is not currently trading, or turnover is so low that the company is not liable for VAT. It should be noted that records of inactive enterprises may be kept in the IDBR database for up to two years, until confirmation that an

Box 1

The IDBR structure and information collected



The 'enterprise' can be thought of as the 'company'. Local units are 'plants', for example, a retail outlet or factory. The plant is the source of business activity. It may be a factory that produces finished goods or an accountancy office, for example. The left-hand side shows a single-site enterprise (the enterprise is the local unit). The right-hand side provides an example of a more complex company: the enterprise group 'owns' two companies, Family Cars and Sports Cars. Family Cars controls two local units. These are factories which produce cars and bikes respectively.

- Name
- Address including postcode
- Standard Industrial Classification (SIC) 2003 and 2007
- Employment and employees
- Turnover
- Legal status (company, sole proprietor, partnership, public corporation or nationalised body, local authority or non-profit body)
- Enterprise group links
- Country of ownership
- Company number
- Intrastat marker for goods and services traded (imports and exports) between the EU member states and the UK.

The following information is collected at local unit and enterprise level:

Box 2

The Virtual Microdata Laboratory (VML)

The Virtual Microdata Laboratory (VML) is a depository of economic and social firm data which ONS makes available to researchers, across academia and government, and has already supported a large volume of academic and policy work³. It is a

secure technical environment adhering to strict disclosure control principles to ensure full confidentiality of all data. Responsibility for running the laboratory rests with the Microdata Analysis & User Support (MAUS) team⁴.

Box 3

Variables in the BSD (enterprise level)

Variable

- Enterprise reference
- Enterprise group reference
- Postcode
- Immediate foreign ownership
- Ultimate foreign ownership
- Employment
- Turnover
- Standard Industrial Classification
- Birth date
- Death date
- Number of local units
- Number of reporting units
- Status

Description

- IDBR identifier code
- Reference of parent organisation
- Enterprise postcode
- Foreign ownership marker
- Foreign ownership marker
- Employment for enterprise
- Turnover for enterprise
- SIC92, SIC03 or SIC07 (depending on year)
- Year business began trading
- Year business stopped trading
- Number of local units making up the enterprise
- Number of reporting units making up the enterprise
- Legal status of the enterprise (e.g. company, partnership, sole proprietor etc.)

enterprise has ceased trading is received. **Figure 1** depicts the total number of enterprises for each year of the dataset, and how many of these are active.

It is clear that the total number of enterprises included in the sample increases over time, significantly so from 2003 onwards. However, the number of active enterprises do not rise in proportion.

Trend growth in the number of enterprises by type of sector is now examined. Using Standard Industrial Classification (SIC) information available for each enterprise in the BSD, we classify enterprises as either manufacturing, service or other non-service organisations (agriculture, mining and energy). **Figure 2** shows the growth rate of each sector, in comparison to the baseline year of 1997, and illustrates the decline of the manufacturing sector and the relatively high level of growth by the service sector which appears to accelerate from 2003.

Further descriptive analysis explores the number of enterprises by sector with respect to the number of local units and employment. **Table 1** reports that the largest proportion of (active) enterprises contain only one local unit. These are single-site enterprises, where all the operations of a particular business are located in the same premises (see Box1). Single site enterprises also account for the largest proportion of employment. By contrast, enterprises with the highest number of local units (typically large nationwide organisations) are fewest in number. However owing to their size, they account for the second-highest proportion of employment in

the IDBR. Enterprises with between 51 and 100 local units account for the smallest proportion of employment.

In addition, the average employment per local unit with respect to the number of local units per enterprise is found to follow an inverted U-shaped curve: average employment per plant is smallest at single-site enterprises, increasing with more local units. However average employment per local unit falls for the very largest enterprises (by local unit count).

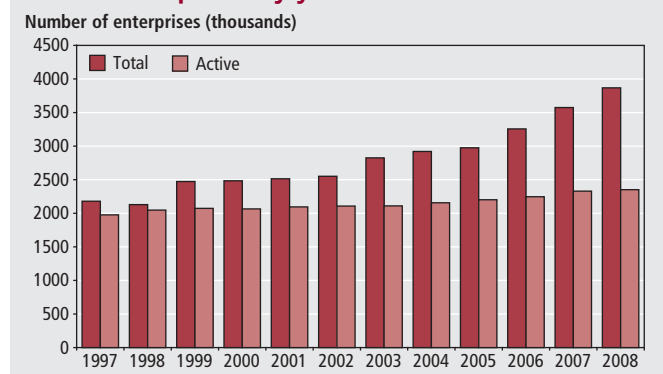
Table 2 displays some characteristics for enterprises in each of the three defined sectors. The panel nature of the observations allows a comparison of characteristics in different time periods. For example, in 2007/2008 a greater proportion of enterprises were single site compared to the previous decade, for all sectors. It is also interesting that activity in service enterprises is not as concentrated on a small number of local units compared to enterprises in the manufacturing or other non-service sectors – a higher percentage of service enterprises derives activity from over 1000 local units (these are likely to be large retailers and chains).

Average employment for enterprises in each sector have been calculated. The average employment is relatively small, which is not surprising given that the majority of enterprises are based in one location and, as seen in Table 1, employ small numbers of workers. However, it can be inferred that on average, manufacturers employ more workers. These results may be explained by the fact that manufacturing plants are larger than, for example, small retailers or practices classified as service industries.

Finally, two types of entry and exit rates for enterprises in each sector (using the method adopted by Disney et al (1999)) are presented. The first 'panel' rates refer to movements of enterprises into and out of the BSD panel per se. The second set of rates are calculated using the IDBR enterprise birth years and death years available in the BSD.

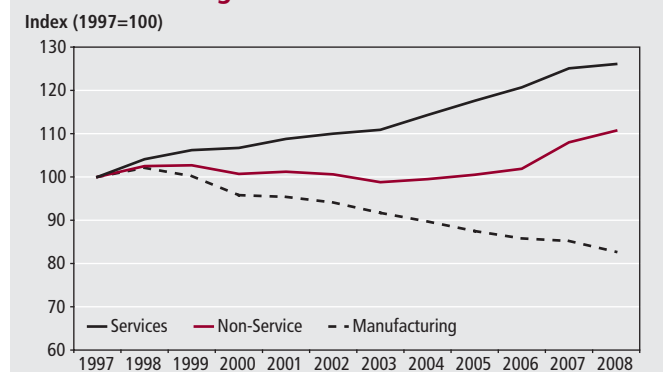
There are two advantages for using the latter method. First, accurate rates may be calculated based on real trading activity by an

Figure 1
Active enterprises by year



Source: ONS Business Structure Database

Figure 2
Relative sector growth 1997–2008



Source: ONS Business Structure Database

Table 1
Constitution of Enterprises

Number of local units	Average number of enterprises (per cent of total in brackets)	Average count of employment (per cent of total in brackets)	Average employment per local unit, 1997–2008
	1997–2008	1997–2008	1997–2008
1	2,075,187 (96.91)	11,202,219 (42.02)	5
2 to 5	57,178 (2.67)	3,522,001 (13.21)	25
6 to 10	4,302 (0.2)	1,302,017 (4.88)	41
11 to 20	2,127 (0.1)	1,291,142 (4.84)	42
21 to 50	1,414 (0.07)	1,558,587 (5.85)	35
51 to 100	497 (0.02)	1,112,781 (4.17)	32
100 plus	670 (0.03)	6,672,331 (25.03)	26
Total	2,140,705	26,661,078	

Source: ONS Business Structure Database

Table 2
Characteristics of enterprises by sector

Period	Percentages					
	Manufacturing		Non-Service		Services	
	97/98	07/08	97/98	07/08	97/98	07/08
No. of local units						
1	82.1	86.2	94.5	96.2	73.4	76.6
2 to 5	11.6	8.3	2.3	1.5	7.7	4.8
6 to 10	1.7	1.5	0.5	0.3	1.6	1.3
11 to 20	1.5	1.1	0.6	0.4	1.4	1.3
21 to 50	1.5	1.1	0.6	0.4	2.1	1.8
51 to 100	0.5	0.6	0.3	0.3	1.7	1.5
101 to 500	1.1	1.1	0.9	0.7	5.2	5.5
501 to 1000	0.0	0.3	0.2	0.1	3.1	2.7
1000 plus	0.0	0.0	0.1	0.0	3.8	4.5
Mean Employment	24	20	5	5	14	13
Entry Rate Panel	8.4	8.6	6.3	10.7	8.3	14.7
Exit Rate Panel	10.2	11.6	7.5	11.4	10.6	14.6
Entry Rate	7.5	7	4.7	8.8	8.1	12.5
Exit Rate	9.4	8.8	7.1	7.2	10.2	10.8

Source: ONS Business Structure Database

enterprise, not simply when entry or exit into or out of, or inclusion or exclusion of, a survey occurs. Secondly, calculating an entry rate using the first year of the data series will result in a 100 per cent entry rate, and likewise when calculating an exit rate using the last year of data. This problem can be solved by using the actual recorded year that a company begins or ceases to trade.

In Table 2, the second set of entry and exit rates are smaller than the first, which reflects the fact that enterprises may be entering or leaving the data, but for administrative reasons, valid birth or death dates have not been confirmed and entered onto the IDBR. The entry and exit rates are consistent with the story portrayed in Figure 2. In both time periods, exit rates in the manufacturing sector are higher than entry rates – more enterprises cease to trade than begin, and this highlights the decline in manufacturing. The reverse is true for enterprises in the service sector.

Why use the BSD?

Using the BSD by itself or in conjunction with other data offers the following advantages:

- The ability to analyse data at the local unit level
- Historical data allows researchers to examine business structure, performance and behaviour throughout time.
- Provision of a large sample of enterprises – particularly useful

when concentrating on specific five-digit SIC sectors, when other surveys would only cover a handful of enterprises. The large sample can also be used to construct more accurate entry and exit rates, which are often utilised by industrial and labour economists.

- Turnover and employment are derived from administrative sources, not inferred. This may be useful when examining smaller organisations which are infrequently targeted by other ONS surveys, such as the Annual Business Inquiry.
- Possibilities for detailed spatial analysis.

Box 4 provides a brief description of two completed projects using the BSD at the VML.

Current research using the BSD includes analysis of growth rates in Northern Ireland, and turnover analysis of 'creative' industries.

In addition, many researchers use information from the BSD to extend sample size and therefore coverage of business organisations included in other ONS business surveys. An example is linking the BSD to the E-commerce Survey – businesses included in the latter may serve as a counterfactual when examining turnover in the BSD.

Future work on the BSD

It may be possible to define and trace 'demographic events' experienced by organisations throughout time. Based upon Eurostat guidelines (see European Commission (2003)), these include:

- takeovers
- mergers
- change of ownership
- break-ups

Preliminary work has already been undertaken in this area. For example, using a combination of enterprise references and enterprise group references, the case where an enterprise changes its enterprise group reference number from one period to the next may infer that the enterprise has become part of a new enterprise group (i.e. either because of a takeover or merger). The ability to implement the Eurostat methodology is currently being examined.

It may be possible to extend coverage of the BSD backwards in time, by linking observations to the Annual Respondents Database (ARD) which is primarily formed from the ABI. Such an exercise ought to allow a selected number of enterprises in the BSD which are also in the ARD to be traced back to the early 1990s and beyond.

Box 4

Research using the BSD at the VML

Foreman-Peck J and Nicholls T (2008) 'Mergers and movement: periphery and the impact of SME takeovers'

This research examines whether takeovers of small and medium-sized enterprises (SME) by larger organisations leads to an increase in productivity. Takeovers and mergers can potentially be identified in the BSD by examining changes to links between organisations. The authors found that takeovers do not lead to an increase in productivity, post acquisition.

Hijzen A, Upward R and Wright P (2008) 'Job creation, job destruction and the role of small firms: firm-level evidence for the UK'

Pioneering research estimating the quantity of jobs created and 'destroyed' by different types of business organisations from 1997 to 2007 (this research is currently being updated with the latest data), and reasons to explain the magnitudes of these rates. For example, analyses are undertaken by business size, exposure to international competition, and the role of job relocation.

Final remarks

The BSD has been created to allow VML users access to IDBR data for the purpose of research. The data can accurately depict the life-span of enterprises, and also the local units associated with an enterprise. Furthermore, changes to company structure can be analysed.

The inclusion of IDBR reference numbers on each observation allows the potential for linking to other data sources held in the secure environment of the VML. This is designed to create new opportunities for research, and increase the impact of research using ONS data.

Notes

1. This work contains statistical data from ONS which is Crown copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

We would like to thank Andrew Allen at the IDBR team for helpful advice on the IDBR and continued support to the VML, and also Dr Felix Ritchie, ONS, Rhys Davies and Robert Gilhooly (both formerly of the ONS) for guidance.

2. Comprehensive information about the IDBR may be obtained from:
www.statistics.gov.uk/idbr
3. Detailed information about the data sets held in the VML can be found at:
www.ons.gov.uk/about/who-we-are/our-services/vml/about-the-vml/datasets-available/dataset-downloads/index.html
4. Contact the MAUS team at:
maus@ons.gsi.gov.uk

CONTACT

✉ elmr@ons.gsi.gov.uk

REFERENCES

- Davies R (2006) 'Business Structure Database user guide version 1', Office for National Statistics mimeo.
- Disney R, Haskel J and Heden Y (1999) 'Entry, exit and establishment survival in UK manufacturing', Centre for Research on Globalisation and Labour Markets Paper 99/9.
- European Commission (2003) 'A general overview of demographic events: business register', Recommendations Manual, Chapter 12.

 Office for
National Statistics

Related titles from the Office for National Statistics



www.palgrave.com/ons/

