

Detailed notes on “Size Analysis of Welsh Business, 2003” release

1 General notes

This document describes the process that has been used in calculating estimates for the First Release titled “Size Analysis of Welsh Business, 2003” published on 30 September 2004 (SDR 69/2004).

The release provides data on the structure of enterprises active in Wales in 2003, including estimates for the very smallest businesses that operate below the VAT threshold. It includes estimates of the number of enterprises active in Wales by size band (defined in terms of number of UK employees), and the aggregate employment and turnover in those enterprises. Results are presented broken down by size band, broad industry category and/or some of the different geographies within Wales.

1.1 Coverage of the analysis

Coverage of the analysis given has been restricted to what is referred to in the release as the “business sector.” This is defined as all private sector organisations, plus all public corporations and nationalised bodies. All central and local government organisations have been excluded. In terms of legal statuses as given in the following table, coverage relates to codes 1, 2, 3, 4 and 7 only.

Legal status code	Legal status description	Public sector?
1	Company	No
2	Sole proprietorship	No
3	Partnership	No
4	Public corporation / nationalised body	Yes
5	Central government	Yes
6	Local authority	Yes
7	Non-profit body or mutual association	No

All industries that are covered by these legal statuses are included in the analysis.

Only enterprises considered to be active in Wales are included, namely those with non-zero employment or turnover within their Welsh elements.

Estimates are included for the very smallest businesses that operate below the VAT threshold or that do not operate a PAYE scheme.

1.2 Key variables analysed

The analysis in this release is based on enterprises with some activity in Wales, regardless of where the enterprise is registered. In this context, an enterprise is taken to be an entire organisation, including its head office and all the individual UK sites, wherever they may be located. The region of registration of the enterprise is usually based on the location of its head office, although some of the individual sites may be located in other regions.

Three key variables are analysed, as follows:

- the number of enterprises which are active in Wales (or one of its sub-areas), which means that an enterprise is counted if at least one of its active sites is located in that area (an active site is one which has some employment or turnover attributed to it);
- the aggregate employment across all the sites within the enterprises that are located in that area; and
- the aggregate turnover allocated to these same sites.

The first of these variables **cannot** be summed across sub-Wales areas to give a Wales figure, as one enterprise may be active in more than one of the different areas and hence be double counted. The other two variables **can**, however, be summed, as the employment and turnover are allocated to individual sites, which can only be present in a single area.

Note that each employee in each enterprise is counted once, whether they are full-time or part-time, so the employment variable is a headcount measure, as opposed to the number employed on a whole-time equivalent basis. The employment variable is thus a count of jobs as opposed to a count of persons in employment (as the same person could be included as an employee of more than one of the enterprises), and clearly a workplace-based measure (as opposed to a residence-based measure).

Estimates for the turnover variable exclude figures for the financial and business services sector throughout this release. This is because the turnover information available for enterprises in the financial services sector also includes the amounts involved in the transactions carried out by these enterprises and is therefore not comparable to turnover information in the other sectors.

1.3 Size bands analysed

In all cases, the size band of the enterprise is based on the number of UK employees in the enterprise, so that an enterprise employing 10,000 UK staff but only a handful in Wales is categorised as a large, and not a micro, enterprise in the analysis.

1.4 Rounding and suppression

As all enterprise counts and employment aggregates in this release have been independently rounded to the nearest 5, and all turnover aggregates independently rounded to the nearest million pounds, the figures may not add up exactly.

In order to avoid suppression of data due to disclosure control (which must occur for any aggregates based on less than 20 enterprises), some of the size bands have been presented in an aggregated form. These aggregated size bands are shown in italics, and have been chosen to limit the amount of aggregation whilst ensuring no data are suppressed.

1.5 Summary of sources

The data in this release have been derived from a number of sources. The primary source is a detailed dataset based on extracts taken in March 2003 from the Inter-Departmental Business Register (IDBR) maintained by the Office for National Statistics (ONS). This dataset gives details of all enterprises active in Wales, including the UK employment and turnover and derived figures for the Welsh elements of employment and turnover (using information about the sites in each enterprise which are located in Wales). It also includes identifiers for industry (using the 2003 Standard Industrial Classification or SIC2003), geographical location and legal status of each enterprise, thus allowing the analysis given in the release.

The IDBR covers a very high proportion of the employment and turnover in the UK, but because it excludes most of the very large number of smallest businesses that operate below the VAT threshold, its coverage in terms of the number of enterprises is only partial (around 50%). In order to correct for this, the figures presented in the release include estimates for unregistered enterprises, together with their related employment and turnover. These estimates for unregistered enterprises are calculated using a combination of:

- IDBR data;
- information on persons who are self-employed in their main or second job, taken from the Labour Force Survey carried out by the ONS; and
- information on the numbers of sole traders and partnerships from the Survey of Personal Incomes data provided by the Inland Revenue.

2. Detail behind the calculations

2.1 Developing a dataset for analysis from the IDBR

The analyses of enterprises registered on the IDBR is central to the analysis and is based on two extracts taken in March 2003, one covering all enterprises with some activity in Wales and one covering all Welsh sites (or local units) in these enterprises.

The enterprise extract includes a unique enterprise reference number as well as fields providing information about each enterprise such as UK and Welsh employment and employees, UK and Welsh turnover, industry, geographical location and legal status. The local unit extract includes a unique local unit reference number as well as a field linking the unit to its unique parent enterprise reference number, together with further fields detailing employment and geographical location.

The enterprise extract is developed into a detailed dataset by adding a field allocating each enterprise to an employee size band (based on the number of UK employees) and excluding all enterprises that are out of scope (see section 1.1). Also excluded are the small number of enterprises that are possibly double-counted on the IDBR (although note that ONS are currently carrying out detailed enterprise matching to identify the enterprises that are genuinely double counted which will remove the need for this step in future analyses of this kind).

Within this detailed dataset, the employee sizeband for companies (legal status 1) with only 1 employee is changed to the zero employee sizeband. This is because people who work on their own and register as a company automatically become an employee of the company when it is set up. Without this adjustment, the numbers of businesses with zero or one employee will change purely because of changes in legal status (which can be driven by changes in tax / liability laws for example), rather by any real change in the economy.

In order to carry out the geographical analysis within Wales, additional datasets are created from this first detailed dataset, with separate datasets required for each of the different geographic breakdowns of Wales given in the analysis. Each additional dataset is created by disaggregating those enterprises with multiple local units in Wales into the separate parts of each enterprise that are located in each geographic area being considered, using information about the local units of each enterprise.

2.2 Analysing the IDBR datasets

The datasets defined above are analysed to produce the required summaries of those enterprises registered on the IDBR. These analyses give enterprise (or part enterprise) counts and related employment and turnover aggregates, as follows:

- Figures for Wales by sizeband only
- Figures for Wales by both industry and sizeband
- Figures for West Wales and the Valleys and East Wales by both industry and sizeband
- Figures for West Wales and the Valleys split into the Valleys areas and the Rest by both industry and sizeband
- Figures for Welsh local authorities by sizeband only

The first two of these analyses come from the main dataset described in section 2.1, whilst each of the other three is based on one of the additional datasets created (as described in the last paragraph of section 2.1).

Whilst the employment and turnover totals within the different geographic areas will sum to the figures for Wales, the same is not true of the counts of enterprises active in each area. This is because a number of the enterprises are active in more than one of the areas and is the reason why additional datasets need to be created for each different geographic breakdown considered.

2.3 Calculation of estimates of unregistered enterprises

Note that the Annex gives an example of these calculations at the Wales level to help with the understanding of this section. The calculation of unregistered enterprises is set out in two parts below, with reference to the steps set out in the Annex.

2.3.1 Calculation of estimates of unregistered working proprietors

The IDBR covers a very high proportion of the employment and turnover in the UK, but because it excludes most of the very large number of smallest businesses that operate below the VAT threshold, its coverage in terms of the number of enterprises is only partial (around 50%). In order to correct for this, the figures presented in the release include estimates for unregistered enterprises, together with their related employment and turnover. These estimates for unregistered enterprises are calculated using a combination of:

- IDBR data;
- information on persons who are self-employed in their main or second job, taken from the Labour Force Survey carried out by the ONS; and
- information on the numbers of sole traders and partnerships from the Survey of Personal Incomes data provided by the Inland Revenue.

It is firstly necessary to calculate an estimate of the number of Welsh working proprietors (i.e. enterprise owners located in Wales) that are **not** included on the IDBR. To do this the number of Welsh working proprietors recorded on the IDBR is required and can be estimated as the difference between Welsh employment and employees for each enterprise using the detailed main dataset described in section 2.1 (see step 1 in the Annex).

This number can then be subtracted from the total number of working proprietors, which are estimated using self-employment data from the Annual Local Labour Force Survey (ALLFS). The data used are counts of persons (on a workplace basis) who were self-employed in either their main job or second job in the year ending February 2003. The ALLFS has recently been boosted to cover about 20,000 Welsh households.

It is assumed that all unregistered enterprises are either sole traders or small partnerships (i.e. with only 2 partners) with no employees, on the basis that bigger enterprises will have turnover above the VAT threshold and hence be included on the IDBR. Therefore both the IDBR data giving registered Welsh working proprietors and the LFS data giving total Welsh working proprietors need to be broken down into these two categories. This is achieved by making reference to the legal status of the enterprise for the IDBR data, but it is more difficult for LFS data.

Data from the Survey of Personal Incomes (SPI) are used to breakdown the LFS data into the two categories. A ratio of the number of Welsh sole traders to the sum of Welsh sole traders and partners is derived from the SPI data and applied to the LFS data in estimating the number of **sole** working proprietors from the LFS. The remaining working proprietors in the LFS are then assumed to be partners (see step 2 of the Annex). The SPI data used are an average of figures for 2000-01 and 2001-02, the latest available.

Thus it is possible to calculate raw estimates for the number of sole traders and partnerships **not** included on the IDBR. Note that this calculation is broken down by industry and can also be broken down by local authority (see step 3 in the Annex). However, the SPI ratios used to breakdown the LFS data are only available at the Wales level by industry, and so in calculating the local authority figures, the Welsh ratios for each industry are assumed to hold for all local authorities. This assumption is probably a reasonable one, based on some analysis of the limited SPI data by industry at a sub-Wales level which indicated that differences in the SPI ratios by industry were more significant than differences by location.

2.3.2 Conversion into unregistered enterprises and related employment and turnover

The calculation of the raw estimates of unregistered Welsh working proprietors yields some negative numbers for the agriculture industry, probably due to some classification inconsistencies in the various datasets in respect of agricultural partnerships. It is assumed therefore that all working proprietors for agriculture are included on the IDBR and so the estimated unregistered figures for agriculture are set to zero (step 4 in the Annex).

In order to convert the estimate of unregistered working proprietors into an estimated count of unregistered enterprises in Wales, the number of partnerships is halved and added to the number of sole traders. This is based on the assumptions that there are no employees in either category of unregistered enterprises and that there are only two partners in each unregistered partnership (see step 5 in the Annex).

The number of unregistered working proprietors is equal to the estimated Welsh employment aggregates for unregistered enterprises. Again this is based on the assumption that there are no employees in either category, the Welsh employment thus being equal to the Welsh working proprietors themselves (see step 5 in the Annex).

Finally in order to calculate Welsh turnover aggregates for each industry grouping, estimates of turnover per employment for those sole traders and partnerships with no employees which **are** registered on the IDBR are applied to the total number of Welsh working proprietors derived above. In doing so, the turnover per employment estimates for these registered enterprises are halved, (based on the assumption that they will have higher turnover than their unregistered counterparts). Resultant turnover figures are also constrained so that the turnover per unregistered enterprise does not exceed the VAT threshold in each industry (see step 5 in the Annex).

Note the Wales turnover per employment ratios for registered enterprises are used when calculating turnover aggregates for the unregistered enterprises for the different geographical breakdowns of Wales.

2.4 Calculation of estimates of all enterprises

The figures for unregistered enterprises are added to the datasets that emerge from the analysis of the IDBR extracts as detailed in 2.2. The unregistered figures are added to the zero employee size band for each industry and geographical area to be consistent with the assumption that all unregistered enterprises have no employees.

The unregistered enterprises account for around 50% of the total number of enterprises active in Wales (in the business sector), 10% of the related Welsh employment and 3% of the related Welsh turnover.

Annex 1

Example calculation of unregistered enterprise estimates

The notes against the column headings below are explained in full at the end of the Annex.

Step 1: Take registered working proprietors (legal status 2 and 3) from IDBR

		<i>Registered working proprietors</i>		
SIC Code	SIC Description	Sole Proprietors (LS2)	Partners (LS3)	Working Proprietors
		(a)	(b)	(c)=(a)+(b)
AB	Agriculture	6,263	18,759	25,022
CDE	Production	1,443	1,875	3,318
F	Construction	4,232	3,174	7,406
GHI	Distribution, Hotels, Restaurants and Transport	11,368	16,658	28,026
JK	Financial and Business Services	4,027	4,724	8,751
LMN	Public Administration, Health and Education	1,189	2,592	3,781
OPQ	Other Services	2,086	1,640	3,726

Step 2: Take LFS estimates of total working proprietors, broken down using Wales level SPI data

		<i>LFS estimates of total working proprietors, broken down using Wales SPI data</i>		
SIC Code	SIC Description	Sole Proprietors (LS2)	Partners (LS3)	Working Proprietors
		(d)=(f) x SPI	(e)=(f)-(d)	(f)
AB	Agriculture	7,231	13,637	20,868
CDE	Production	7,937	3,486	11,423
F	Construction	28,096	5,970	34,066
GHI	Distribution, Hotels, Restaurants and Transport	23,483	19,508	42,991
JK	Financial and Business Services	16,815	7,988	24,803
LMN	Public Administration, Health and Education	11,352	5,995	17,347
OPQ	Other Services	14,947	3,454	18,401

Step 3: Subtract registered figures from LFS estimates to get raw estimates of unregistered enterprises

		<i>Raw estimates of unregistered working proprietors</i>		
SIC Code	SIC Description	Sole Proprietors (LS2)	Partners (LS3)	Working Proprietors
		(g)=(d)-(a)	(h)=(e)-(b)	(j)=(g)+(h)
AB	Agriculture	968	-5,122	-4,154
CDE	Production	6,494	1,611	8,105
F	Construction	23,864	2,796	26,660
GHI	Distribution, Hotels, Restaurants and Transport	12,115	2,850	14,965
JK	Financial and Business Services	12,788	3,264	16,052
LMN	Public Administration, Health and Education	10,163	3,403	13,566
OPQ	Other Services	12,861	1,814	14,675

Step 4: Calculate final estimates of unregistered working proprietors, adjusting to zero for agriculture (i.e. assuming IDBR covers all of agriculture)

Final estimates of unregistered working proprietors

SIC Code	SIC Description	Sole Proprietors (LS2) (k)	Partners (LS3) (l)	Working Proprietors (m)=(k)+(l)
AB	Agriculture	0	0	0
CDE	Production	6,494	1,611	8,105
F	Construction	23,864	2,796	26,660
GHI	Distribution, Hotels, Restaurants and Transport	12,115	2,850	14,965
JK	Financial and Business Services	12,788	3,264	16,052
LMN	Public Administration, Health and Education	10,163	3,403	13,566
OPQ	Other Services	12,861	1,814	14,675

Step 5: Calculate unregistered enterprises, employment and turnover (see notes overleaf)

SIC Code	SIC Description	Enterprises: Unregistered Sole Traders (n)=(k)	Enterprises: Unregistered Partnerships (o)=(l)/2	Enterprises: All Unregistered Enterprises (p)=(n)+(o)
		(n)=(k)	(o)=(l)/2	(p)=(n)+(o)
AB	Agriculture	0	0	0
CDE	Production	6,494	806	7,299
F	Construction	23,864	1,398	25,262
GHI	Distribution, Hotels, Restaurants and Transport	12,115	1,425	13,540
JK	Financial and Business Services	12,788	1,632	14,420
LMN	Public Administration, Health and Education	10,163	1,701	11,865
OPQ	Other Services	12,861	907	13,768

SIC Code	SIC Description	Employment: Unregistered Sole Traders (q)=(k)	Employment: Unregistered Partnerships (r)=(l)	Employment: All Unregistered Enterprises (s)=(q)+(r)
		(q)=(k)	(r)=(l)	(s)=(q)+(r)
AB	Agriculture	0	0	0
CDE	Production	6,494	1,611	8,105
F	Construction	23,864	2,796	26,660
GHI	Distribution, Hotels, Restaurants and Transport	12,115	2,850	14,965
JK	Financial and Business Services	12,788	3,264	16,052
LMN	Public Administration, Health and Education	10,163	3,403	13,566
OPQ	Other Services	12,861	1,814	14,675

SIC Code	SIC Description	Turnover: Unregistered Sole Traders (t)	Turnover: Unregistered Partnerships (u)	Turnover: All Unregistered Enterprises (v)=(t)+(u)
		(t)	(u)	(v)=(t)+(u)
AB	Agriculture	0	0	0
CDE	Production	158,684	26,912	185,596
F	Construction	741,708	68,773	810,482
GHI	Distribution, Hotels, Restaurants and Transport	465,336	75,244	540,580
JK	Financial and Business Services	311,476	57,482	368,958
LMN	Public Administration, Health and Education	236,320	58,601	294,921
OPQ	Other Services	324,191	33,065	357,256

Notes (see column headings in above tables for references):

(a) Extracted from IDBR

(b) Extracted from IDBR

(c) = (a)+(b)

(d) LFS WPs (see (f)), multiplied by Wales SPI percentage of sole traders to sole traders plus partners for that industry

(e) LFS WPs (see (f)), less estimate of sole traders

(f) LFS WPs measured by reference to main and second self-employment jobs, by workplace work

(g) LFS sole traders estimate less IDBR sole traders figure

(h) LFS partners estimate less IDBR partners figure

(j) = (g)+(h)

(k) For agriculture, assume full coverage on IDBR, therefore = 0.

(l) For agriculture, assume full coverage on IDBR, therefore = 0.

(m) = (k) + (l)

(n) Number of unregistered sole trader enterprises = derived number of unregistered sole traders (assuming no employees)

(o) Number of unregistered partnership enterprises = half of derived number of unregistered partners (assuming two partners and no employees)

(p) = (n) + (o)

(q) Employment in unregistered sole trader enterprises = derived number of unregistered sole traders

(r) Employment in unregistered partnership enterprises = derived number of unregistered partners

(s) = (p) + (r)

(t) Turnover in unregistered sole trader enterprises = assumed turnover per head of employment at **Wales and SIC level** x assumed employment (see (q))

(u) Turnover in unregistered partnership enterprises = assumed turnover per head of employment at **Wales and SIC level** x assumed employment (see (r))

(v) = (t) + (u)