

**BERR SUMMARY: ONS - UK COMPANY STATISTICS
RECONCILIATION PROJECT**

A - INTRODUCTION

1. This research was commissioned by the Department of Business, Enterprise and Regulatory Reform (BERR) in October 2007 to contribute to strategic priorities in connection with better regulation and reducing the administrative burdens within the Corporate and Insolvency Activity Framework. The ONS was chosen to carry out this research due to the expertise required for this report with additional input coming from the BERR analysts. The project was undertaken from October 2007 to March 2008.

2. This project had three main objectives. The first objective was to develop a clearer understanding of the numbers and types of companies in the UK. This information could then be used by BERR and other stakeholders in a variety of policy debates and to support the production of Impact Assessments¹. The second objective was to analyse trends in the UK company population. Finally, this project wanted to find out if different databases (FAME, IDBR and BSD) would give a consistent view of the UK company population. In the past, different databases have been used in the policy arena, and the question has arisen; should a certain database be used in order to answer a certain question or can any database be used to give a transparent and accurate answer?

B - BACKGROUND

3. For the purpose of this report "Companies" are defined as corporate bodies that are registered at Companies House.

4. Companies can be grouped into size categories (small, medium and large). In many published statistics this is based on an employment criterion alone but definitions for regulatory purposes are often more complex. The Companies Act (2006) defines size using three measures - employment, gross assets and turnover. To be classified as a certain size the company must meet the threshold for that size band on at least two measures. The specific values for each measure are shown in table 1.

Table 1: Company size thresholds up to 31 March 2008

Variable	Small	Medium	Large
-----------------	--------------	---------------	--------------

¹ An Impact Assessment is a process used by the UK government to evaluate any regulatory and non regulatory change the government is considering

Number of employees	≤ 50	$50 < x \leq 250$	>250
Turnover (£ million)	≤ 5.6	$5.6 < x \leq 22.8$	> 22.8
Gross Asset (£ million)	≤ 2.8	$2.8 < x \leq 11.4$	> 11.4

Note: The symbols ' \leq ' is the less than or equal sign

5. In the case where a company varies on all three measures, for example, turnover (small), number of employees (medium) and gross assets (large) the company will be classified as a medium company.

6. Small companies that have turnover not exceeding £5.6 million and a balance sheet total not exceeding £2.8 million will also qualify for audit exemptions (a company's annual accounts will not have to be audited).

7. In the first year a company's size is determined in the static definition mentioned above. However, over time there is a dynamic element to company size. For a company to change size categories, the company must meet the size category they are moving to, for two consecutive years to become that size. In this transition period, the company will be classified in their original size category.

8. All limited (private and public) companies must send their annual accounts to the registrar (Companies House). Full accounts will include information on employment, gross assets and turnover. If eligible and the companies choose to do so, small, medium and dormant companies may send abbreviated accounts. Abbreviated accounts for small and dormant companies only include information on gross assets. Before the Companies Act 2006, abbreviated accounts for medium companies only included information on employment and gross assets². This creates problems, as all three size measures are required to determine company size, section E below will discuss potential solutions.

C - DATA SOURCES

9. There are a variety of different data sources for businesses and companies in the UK each of which differ according to their own purposes. In this section we will briefly mention the main data sources for UK Companies; this is not a comprehensive list but rather the data sources BERR use most often.

10. **Companies House (CH)** is the official register for companies in the UK. The CH database captures basic information on each company including name, address, date of incorporation, legal form, legal status, type of

² After the Companies Act 2006, abbreviated accounts for medium companies will include employment, gross assets and turnover information

accounts and date of incorporation. CH makes available to the public and commercial database providers annual accounts and all other documents filed with them. CH also produces an annual report, 'Companies Register Activities'³. This is useful for high level statistics including the number of companies on the Register, numbers joining and leaving the register and analyses by legal form and region of registration. It does not itself produce information on the size of individual companies or statistics based on size, which as mentioned is extremely useful for policy makers.

11. **FAME** contains detailed information on all public and private companies currently registered in the UK and includes information on company profiles, profit and loss accounts, balance sheets, cash flow statements, directors' details, industrial activity etc. FAME is a commercial data source compiled by Jordans and Bureau Van Dijk (BvD), and distributed by BvD, with the main source of the information being Companies House. FAME provides easy access to the raw data needed to determine size categories mentioned previously. When a company has a missing value for employment or turnover, which is required to determine company size, FAME has the option of imputing the missing value(s) based on the gross asset value and the Standard Industrial Classification (SIC) of the company.

12. **HM Customs and Revenue (HMRC)** maintains a register of all tax payers in the UK. All businesses and companies that are tax payers (for example, through VAT or PAYE) to HMRC will have a record at HMRC. Records at HMRC will be kept alive as long as there are outstanding tax amounts due. Due to legal reasons the information held at HMRC is not publicly available. However, some of it is used to compile the Inter-Departmental Business Register, described below.

13. The **Office of National Statistics (ONS)** maintains the **Inter-Departmental Business Register (IDBR)** and the **Business Structure Database (BSD)**.

14. The **IDBR** is a live register of all active businesses, including companies, but the unit of focus is businesses rather than company. The sources for this database are HMRC, CH, the ONS's own surveys, interactions with large companies and commercial sources on company ownership and structure. The IDBR covers 99% of economic activity in the UK. The IDBR does not hold information on company's gross assets. This is one weakness of this database for work on company size, as gross assets are needed. The IDBR does not include the very many small businesses operating without PAYE⁴ or VAT⁵ schemes and some not for profit organisations. Access to the IDBR is carefully controlled.

³ Previously known as "Companies In"

⁴ All businesses / companies have a legal obligation to operate a PAYE scheme on the payments they make to their employees if their earnings reach the National Insurance Lower Earnings Limit (LEL). For the tax year 2008-09 this is £90 a week, £390 a month or £4,680 a year.

15. The **BSD** is a series of annual snapshots of the IDBR. It aims to create a longitudinal version of the IDBR for research purposes, taking full account of changes in ownership and restructuring of businesses (including companies). This information is made available to approved researchers within the ONS's secure Virtual Microdata Laboratory environment⁶.

16. **London Stock Exchange (LSE)** produces detailed statistics on UK and international companies listed on the main London market, secondary markets and the Alternative Investment Market (AIM). This information is publicly available.

D: POPULATION

17. Most of the analysis undertaken in this report comes from the May 2007 FAME database. The total population of companies on this database was 3.56 million. Removing Irish and foreign companies reduced the population to 3.34 million.

18. A further 1.07 million companies were removed from the analysis because they had never produced annual accounts and hence there was a lack of information on these companies and more specifically no information on gross assets, employment and turnover. It is possible that most of these companies that have never produced annual accounts are new firms⁷.

19. The final population size used in this report is thus 2.27 million. Companies in this population can be further analysed down by company activity. Approximately 68% of companies are active-live (1.55 million), 7% of companies are inactive-live non-trading (0.15 million) and 25% of companies are inactive-other (0.58 million)⁸. BERR's own analysis linking FAME to ONS BSD (section K) suggests that most of these Inactive - Other companies are dissolved.

E: METHODOLOGY TO DETERMINE SIZE OF THE COMPANY

⁵ All businesses /companies must register for VAT if their turnover for the previous 12 months is over a specific limit - currently £67,000 - or if their turnover may soon go over this limit. Businesses / companies may register voluntarily at any time.

⁶ <http://www.ons.gov.uk/about/who-we-are/our-services/vml/about-the-vml/datasets-available/dataset-downloads>

⁷ New companies have their first accounting period set as the first anniversary of the last day in the month in which the company was incorporated. The company then has a further 9 months in which to submit their accounts. In total a new firm will have up to 22 months to submit their first annual accounts.

⁸ **Active live company** – A company that has trading activities and produces annual returns/ accounts
Inactive-live non-trading company – A company that is not doing 'business', but may have other accounting transactions going through its books.

Inactive other company – include Dissolved (that ceases to exist), In Liquidation, and In Receivership i.e. dead or in course of removal from the register

20. As mentioned in section B, company law does not require all companies to report on all three measures, which are needed to determine company size. 99% of companies file a figure for gross assets, 22% file employment and 6% file turnover. In total 97,000 companies (4% of total population) provided information on all three measures. There are several approaches to estimating company size when one or more of the three required measures are not available⁹.

- I. Missing data assumed small** - If information is missing for any of the measures, assume the size for that measure(s) will be small. For example if there is no information on turnover, assume turnover will be small for that company. The rationale for this method is that if a company has not provided specific information on one of the measures (or two) it is because they are not legally obliged to do so (small or medium abbreviated accounts).
- II. Imputed and known data from FAME** – For any missing data, FAME will impute a value based on information provided by the other measures (if available) and the standard industry classification (SIC) of the company.
- III. Known data only** - This method focuses only on the information provided by the company. For example if a company has provided information for gross assets (large), the company will be determined as large if no other information has been submitted. Further, if the company has provided information on two of the measures, which disagree when determining the size of the company, then the smallest company size has been chosen.

Table 3: Breakdown of company size according to the three different approaches

Size (000's)	Missing data is small	Imputed and known data from FAME	Known data only
Small	2,229	2,209	2,187
Medium	31	44	60
Large	17	24	31
Total	2,277	2,277	2,278

21. Table 3 shows a breakdown of company size by the three different methods. As expected method 1 gives the smallest number of medium and large companies, similarly method 3 gives the largest number of medium and large companies. These can be regarded as placing lower and upper limits respectively on the number of medium and large companies. An interesting point to note is that method 2 gives a number almost in the middle of method

⁹ To determine Company size in this report the static definition has been used.

1 and 3 for the number of medium and large companies. Most companies are small, whatever method is used.

22. BERR analysts have suggest that method 2 may be more appropriate for determining company size, as the missing value will be estimated using known data and the SIC code; however more work in this area is still required. The main ONS analysis uses the third method and is the basis of the results reported below. This method was chosen for two reasons by the ONS researchers. Firstly, imputed values (method 2) are derived from gross assets and SIC code, implying difficulties in attempting to analyse, for example, employment compared to gross assets as if they were two independent variables (section F). Finally, drawing relationships between FAME and other datasets such as the IDBR requires reviewing actual results, which does not occur under the other methods. The remaining analysis in this report will not be sensitive to the method chosen to determine company size. However, when looking at some regulatory changes the calculation of costs and benefits will be sensitive to the number of medium and large companies used.

F: DOES MISSING DATA MATTER IN DETERMINING COMPANY SIZE?

23. To determine company size, information on all three measures is required, however, due to reporting requirements we do not demand that companies provide all the necessary information. Almost all (99%) companies provide information on gross assets, this then leads us to the question; are the measures good proxies for each other in determining company size? For example if a company has gross assets which are large, would you typically expect the company's turnover to be large?

24. Where information exists, analysis has shown that the size band for each measure tends to agree¹⁰. For example if a company has turnover which is small, you would expect the company's number of employees to be small. This is especially true for small and to a lesser extent large companies; medium companies are least likely to have agreement on both measures.

G: COMPANY SIZE – STATIC VERSUS DYNAMIC

25. Company size as mentioned in section B is based on static and dynamic elements. In this project, to determine company size the static definition has been used (the dynamic element has been ignored due to the complex calculation required), but the question arises; how good a proxy is a single year's data for company size over two or more years?

¹⁰ Results show that employment and turnover will allocate companies to the same size band on 76% of times from the fully reported population

26. To answer this question partially, analysis has been undertaken comparing company size for each measure in static periods (2003 or 2005) to that of a dynamic period (2002-2003 or 2004-05) of two years. For example, comparing a company turnover size band in 2003 to the lower of its size bands for turnover in 2002 and 2003.

27. The results show when a measure is of a particular size, it is highly likely to be that size in the future i.e. if turnover is small in year 1; it is likely to be small in year 2. For instance, turnover in 2003 and turnover in 2002-03 will report the same company size band in 97% of companies. Therefore, a static definition of company size can be a good proxy for dynamic company size. This analysis was only undertaken, where information was observable (not imputed or implied), and the population used for this analysis was 129K, hence this relationship may not be a true representation of the UK Company Population¹¹.

H: COMPARING ACCOUNT TYPE, LEGAL FORM TO COMPANY SIZE

28. Analysis was undertaken on the relationship between account type, legal form and company size. The results from this analysis are expected and summarised below;

- 96% of all companies in this population are private limited companies.
- 96% of all companies in this population are small (company size).
- Public quoted companies are almost always large.
- 64% of all accounts submitted have small company audit exemptions.
- Around 10% of small companies provide audited full accounts, even though they may be eligible for abbreviated / exempted accounts.
- 10,000 medium companies (17% of all medium companies) provide small company accounts, probably due to the dynamic nature of determining company size (mentioned in section B)

I: MICRO COMPANY ANALYSIS

29. The UK government is committed to promoting new and small businesses / companies, as result the 'Think Small First' approach will lie at the heart of any future proposed changes. Therefore it is important to gain an understanding of the micro companies' population. Micro companies are not defined in the Companies Act 2006 but a commonly used definition is that they have ten or less employees. ONS analysis for this project using the IDBR (December 2007) suggests that 85% of all companies are micro companies. Furthermore, 61% of all companies have three employees or less. Micro companies make up the majority of UK companies.

J: CHURN ANALYSIS

¹¹ The information for this analysis was taken from March 2006 FAME database

30. Churn analysis compared the FAME databases for two periods: March 2006 and May 2007. This analysis should reflect the flow of companies onto and off the Companies House register. This information is important as it can provide insight into the number of new companies and the number of closures by legal form, by size or by industry.

31. New companies are required to register with Companies House therefore the number of new companies in any churn analysis should be accurate. Companies which cease to exist do not have to inform Companies House, hence may not have their status changed accordingly on the register¹².

32. The difference between the two FAME databases is 867,000. To a large extent this will reflect the number of new companies and the number of dead companies in this period, however it may also include companies which died in a previous period being removed from the register in this time period. Furthermore, the accuracy of this churn analysis is uncertain as the two extracts were not produced to the same specification.¹³

33. As previously mentioned (section C), Companies House holds its own record of the number of companies flowing onto and off the register. The number of companies registered at Companies House has increased dramatically in the past eight years. On the 31st March 1998, there were roughly 1.3 million companies registered, by 2006 this figure had increased to 2.3 million companies. However, this is not the whole story; as the number of companies on the register is very dynamic. In 1998/99 there were approximately 200,000 new companies incorporating annually, by 2006/07 this number had increased to 450,000. On the other hand there has been a steady increase in the number of companies becoming dissolved, in 1998/99 and 2006/7 the number of companies dissolved was 100,000 and 200,000 respectively. The net gain (i.e. registrations minus dissolutions) of companies on the companies register has been growing gradually in this time period, from 100,000 companies in 1998/9 to 200,000 in 2006/07

K: RECONCILIATION BETWEEN FAME AND BSD

34. The penultimate part of this project attempted to reconcile the May 2007 FAME database and the 2007 BSD. Reconciling these two databases would bring significant benefits, as both databases combined would contain all the relevant information needed to determine company size. FAME holds information on gross assets for 99% of companies and the IDBR/BSDB contains information on turnover and employment for at least 95% of companies.

¹² Companies are not actually removed from the CH register for some years after they have ceased activity.

¹³ This inconsistency will be overcome when 2007 can be compared with the next (2008) extract.

35. The linking of the two databases was more complicated than first imagined as only half of the observations on the FAME database matched onto the BSD. This was primarily due to timing differences between the two databases and to the two databases having different definitions of inclusion, for example for non-trading companies.

L: IDBR ANALYSIS

36. ONS conducted similar analysis to much of that carried out on the FAME database using the IDBR. The results from the IDBR are broadly similar to the FAME analysis, which is a further means of verifying the analysis undertaken in this research project. The only major difference between the databases was in terms of the number of medium and large companies. The IDBR reported in September 2007 that there were 17,000 medium and 4,000 large companies. These figures differ substantially to the numbers reported in section E, further work in this area is required to find out the reason for this difference. However it should be noted that size bands derived from IDBR are based purely on turnover and employment data.

M: CONCLUSIONS

37. This research project has investigated and analysed the number and types of companies in the UK using the IDBR and the FAME databases. Companies can be grouped by legal form, account type and level of activity from the FAME database easily. However, to determine a company's size information is required on employment, gross assets and turnover, which is more complicated, as not all companies are legally obliged to provide information on these measures.

38. The BERR analysts and the ONS researchers have suggested that there are at least three solutions to the missing data problem. This has created a range for the number of medium and large companies. All analysts are confident that the true population of medium and large companies will be within this range and that the second method (imputed and known data) will generally be the most appropriate method for determining company size. This method suggests that there are 44,000 medium and 17,000 large companies. Further work was carried out on whether the individual measures are good proxies for each other. Results from the FAME databases and the IDBR (where information is available) have indicated that the measures are relatively good proxies for each other and this indicates that the missing data issue is not as problematic as first thought.

39. The UK company population is dynamic. In the last eight years the number of new companies has grown year on year, but so has the number of dissolved companies. The net effect has been an increase in the number of

companies in the UK over the last eight years from 1.3 million in 1998 to 2.3 million in 2006.

40. This research project attempted to link the BSD and the FAME database. If this were possible it would bring significant benefits; the missing data problem to determine company size would almost disappear and information held on the companies could be double checked.

N: FUTURE WORK

41. From this research we have identified a number of areas that need further investigation, these are summarised in the following points:

- Further reconciliation of the IDBR / BSD and the FAME databases would bring significant benefits. The following questions could be tackled:

When the BSD and the FAME database provide the same piece of information (e.g. on turnover), do the two databases agree?

Do companies which link across the BSD and the FAME database have specific characteristics which are less likely to be shown by companies that do not link across the two databases?

The numbers of medium and large companies in the IDBR and FAME do differ. It would be useful to uncover the possible reasons for these differences.

Does the level of matching between FAME and the BSD change over time?

- BERR analysts considered that method 2 (imputed from FAME and known data) is the most accurate method for determining company size and could be more easily replicated in future. Further work in this area could confirm this conclusion.
- Future analysis on the companies that have never produced annual accounts (in the original FAME population) could be undertaken. This research project indicated that these companies are new and in the future we could test out this hypothesis more thoroughly and to find out other characteristics of these companies.
- The size thresholds for turnover and gross assets changed on 1st April 2008. What impact has this change in size thresholds had on the number of medium and large companies?

Corporate Law and Governance

January 2009