UK Sources and Methods

3 July 2013 – Philipp Kern, Gerhard Schnyder, John F. Wilson

**Firm samples**

The 1904 and 1938 benchmark years are from Scott/Windolf; 1976 is from Stockman. Data for all three years has been used unaltered. The 1997 and 2000 data is from Conyon and encompasses over 2100 firms ranked by market capitalisation. As Conyon’s dataset does not include information on total assets, a list of the largest 250 companies by total (net) assets was retrieved from Thomson One and the corresponding firms subsequently extracted from the Conyon dataset.

For the remaining years, we created the sample from scratch using various sources:

The 1958 firm sample presented significant challenges. A source listing all 250 (or more) firms required could not be located. Therefore, the sample was compiled from multiple lists with often overlapping firms. To minimize the overlap, two sources employing different metrics were used:

1. National Institute for Economic and Social Research (1955): “A Classified List of Large Companies Engaged in British Industry.”
2. Fiedler, M. and Gospel, H. (2010): “The Top 100 Largest Employers in UK and Germany in the Twentieth Century,” Cologne Economic History Paper.

The NIESR list includes the 100 largest listed and unlisted British companies by net assets in 1953/54. Fiedler & Gospel list the 100 largest firms by employment. Accounting for overlaps, the two sources yielded 149 unique non-financial firms for 1958. Without other lists for the mid-50s available, there was no straightforward way to increase the sample size to 200 firms. We considered completing the sample by manually searching The Stock Exchange Official Year-Book (StExYB) 1960, which incorporates data for 1958. However, this proved prohibitively time-consuming, as the thousands of firms included in the StExYB are not listed by size, making it impossible to easily identify the next largest companies, i.e. the largest firms not yet in the sample. The sample size for non-financial firms therefore remains at 149, but financial data was collected from the StExYB to calculate total assets and rank them accordingly.

The 1958 financials sample faced similar difficulties, as the NIESR list includes no financials and Fiedler & Gospel only list a handful. In this case, however, manually collecting the remaining large financial firms from the StExYB was a viable option as it lists financial firms in a separate section. It is important to note, though, that some large financials, particularly unlisted ones, may still have slipped through the cracks. As before, total assets were calculated using StExYB data.

The 1983 sample was much easier to define thanks to the Times 1000 list. After removing firms with a country of control outside the UK, the largest 200 non-financial firms ranked by turnover were used. The Times 1000 does not include data on total assets, but it does list capital employed, which is defined as total tangible assets less current liabilities other than bank loans and overdrafts and future tax. Capital employed data was collected for all firms in the sample.

The Times 1000 only names the top financial firms in separate, narrowly-defined lists, e.g. clearing banks, non-life insurance companies, building societies, etc. The sample of financial firms was thus compiled as follows:

* Top 5 finance houses (by capital employed)
* Top 5 accepting houses (by total assets)
* Top 5 discount houses (by total assets)
* Top 9 clearing banks (by total assets)
* Top 5 other British banks
* Top 5 insurance companies - life (by life etc. funds)
* Top 5 insurance companies - non-life (by premium income)
* Top 5 investment trusts (by investments at market value)
* Top 6 building societies (by total assets)

Some of these category lists are rather short (often only top ten), which means that they might leave out firms that are larger than the top firms in another category. Several firms in the sample are subsidiaries of other firms in the sample, e.g. Coutts is a subsidiary of National Westminster Bank. In order to obtain a sufficiently large sample for financial firms, we included these subsidiaries in the sample (contrary to the prescriptions of the ‘white paper’). Ultimately, the number of subsidiaries is limited to the financial sector and is small.

The 1993 sample was compiled using Thomson One according to whitepaper definitions, i.e. the top 200 non-financials and top 50 financials by total assets. The sample includes both private and publicly listed firms incorporated in the United Kingdom.

The 2010 non-financials sample was defined using Bureau van Dijk’s OSIRIS database of publicly listed firms. The search steps used are:

1. Listed/Unlisted companies: Publicly listed companies
2. World region/Country: United Kingdom
3. Type of entities: Industrial companies
4. Total Assets: 2010, Top 250

For the 2010 financials sample, search step 3 was modified to include “Banks and Financial Institutions, Insurance Companies”, step 4 was changed to include the top 100 firms by total assets in 2010. As OSIRIS does not include private companies, the 2010 sample is limited to publicly listed corporations.

**Director Data**

For 1904 and 1938, director data from Windolf/Scott and for 1976 from Stokman was used as-is. In the case of 1997 and 2000, only relevant director data was extracted from Conyon’s dataset, i.e. directors of the firms in our 1997 and 2000 samples.

Director data for the 1958, 1983 and 1993 samples was collected from The Official Stock Exchange Year-Book (StExYB) for the corresponding years. Each person was assigned a unique Director ID. Names, titles and roles were collected as printed in the StExYB. The StExYB is inconsistent in using full first names and initials. Informed guesses were made in cases where there was the suspicion that a person with the same last name, but full first name in one case, and abbreviated first name in the other, was actually the same person. With very common last names, for example Clark or Smith, it was generally assumed that they were distinct persons. However, when the last name was very uncommon (e.g. Wansbrough-Jones) and when titles or other additional information largely matched, it was assumed to be the same person and initials were replaced with the full name, and the Director ID merged.

Director data was collected for all firms that are in the StExYB, which are generally only listed firms.

For the 2010 sample, director data was gathered from the BoardEx database. A list of current directors, i.e. in February 2012, was retrieved using the batch-download feature. This list highlights directors who joined or left the board since the last report date (31/12/2010), which was used to reconstruct the board composition at year-end 2010. As the data exported from BoardEx did not contain a male/female flag, we entered this information manually for familiar first names. For ambiguous or unfamiliar names, the gender information was retrieved from BoardEx by searching for the individual in question (BoardEx’s profile page indicates gender).

Additional information on board appointments in 1997, 2000 and 2010 was retrieved from BoardEx. This includes specific information regarding the individual’s appointment on the board in that year—executive/non-executive director, role description, independence—as well as general information on the individual, including date of birth, nationality, gender, education, achievements and employment history. As there is no unique identifier common to both BoardEx and the respective sources of our director data, individuals were matched using a ‘fuzzy lookup’ technique. To reduce the number of false positives, the similarity threshold was set 0.7 (1 representing a perfect match, 0 no likeness), which was found to produce very few bad matches without omitting too many good matches.

BoardEx does not offer data on board appointments before the late 1990s, so no appointment-specific data could be extracted for 1976, 1983 and 1993. However, because a significant portion of directors in those samples may still have held the same or other board appointments in the late 1990s, general data on those individuals was extracted from BoardEx, including demographics, education, achievements, and employment history. Again, a fuzzy match technique was used as described above.