CBR DATABASE METHODOLOGY

Companies in the Cambridge region

This document describes how the database of over twenty thousand companies operating in the Cambridge region was created. The database contains only companies and limited partnerships (LLPs) and therefore excludes sole proprietorships and other forms of unincorporated businesses. The database provides data for turnover and employment of Cambridge based companies, but it must be noted that these figures represent their employment and turnover in Cambridge and in other parts of the country and abroad. Further work is being carried out to enable us to provide an estimate of the local content of their employment. The source of the data is the FAME database produced by Bureau Van Dijk which includes over 3 million active UK-based companies and a further 6 million that are dormant or who have died.

Cambridge Ahead area and the meaning of Cambridge companies

This is defined as a twenty-mile radius from the centre of Cambridge. All postcodes (about 20,000) that fell within this area were selected. Each postcode was assigned to an electoral ward.

The **Cambridge based** companies (e.g. Arm Holdings) were then selected as companies with either their registered office, or their primary trading address, in one of the Cambridge Ahead postcodes. The location of each company for mapping purposes was taken to be the postcode of its primary trading address. In many cases the primary trading address was not provided, In general, the registered office address was then used as the location postcode. However, in many cases local knowledge and internet searches enabled us to identify a primary trading address despite it not being found on FAME.

Another group of businesses was identified as **Cambridge active** companies (e.g. Smith & Nephew). These companies do not have their registered office, or their primary trading address, in one of the Cambridge Ahead postcodes but do operate within the Cambridge

Ahead area. Most of these were identified since they gave a non-primary trading address within the Cambridge Ahead area. Others were identified from local company network lists, but were included only if the postcode of their trading location in the Cambridge Ahead area could be identified. Whilst the financial data for these companies are collected they are not reported since their global figures would be misleading for the Cambridge region.

Period covered and data collected

The 2015 draw gathered data for five years: 2014/15; 2013/14; 2012/13; 2011/12 and 2010/2011, thus providing four years of growth. The year 2014/15 includes data for companies for accounting year ends between 6th April 2014 and 5th April 2015. It was decided to select any company in the Cambridge Ahead area that had been active in any of these five years.

The data gathered for each company included: Company name; Registered number; Registered accounts type; Legal form; Current Market Capitalisation; Trade Description; its Sector according to the Standard Industrial Classification (SIC) 2007; Peer Group; IPO date; current Immediate Shareholder information; current Domestic and Global ultimate owner; Registered office address and postcode; Primary trading address & employment when provided; and Other trading address in Cambridge area and employment when provided.

The financial information included: Accounting date; total assets; employment; and turnover for each of the five years if available. All changes of accounting dates were identified and adjustments were made where necessary to annualise the accounting information.

Treatment of parents and subsidiaries

This is probably the most difficult aspect of building the database and requires judgements to be made. First, Cambridge based companies with no parent and no subsidiaries on the database can be retained in the dataset. Second, for the remaining companies, all parents and their subsidiaries are gathered together. Third, if the parent is on the database and includes the information about its subsidiaries in its consolidated accounts, then the subsidiaries are dropped and only the parent is retained on the database. Fourth, for those companies with a parent that is a holding company and does not report the consolidated accounts for itself and its subsidiaries, the parent company is removed from the database and the subsidiaries are

retained. Fifth, the same procedure is carried out for Cambridge active companies. This means that it is possible to have a parent as Cambridge active (e.g. AstraZeneca) and its subsidiaries (e.g. Medimmune) as Cambridge based since this will not involve double-counting of financial data.

There are also cases where a company ceases to trade because it transfers its activities to a new company registration. This would show up on the database as a death and a birth whilst no real change has occurred in its economic activity. We have attempted to identify such cases and to show the company as continuing. Another problem is the acquisition of a Cambridge based company by a company from outside the area. In some cases the Cambridge based company may continue to trade and report as before, but in other cases it may be wound up despite its activity in Cambridge being unchanged. We continue to work on this latter problem which will be helped if this database is maintained over time so that these acquisitions can be tracked.

Marshall of Cambridge is a special case and is treated as such. Part of the business, Marshall Motor Holdings PLC has a stock market listing and so this is separated from the rest of the business (largely Marshall Aerospace). We treat the former as non-KI and the latter as KI.

Sector reclassifications

The SICs for the companies are self-reported. Several hundred companies had not reported their SIC. In these cases their activity descriptions and their names (and occasionally internet searches) allowed us to assign SICs to them.

Another group of companies were assigned to generic activities such as 'head offices' or 'holding companies'. These were reassigned into SICs that represented their economic activities.

We decided to have 'Life Science and Heathcare' and 'Information Technology and Telecoms' as two of our sectors owing to their importance in the Cambridge region. This involved identifying these businesses across a wide variety of SICs. We were helped in achieving this (and in other decisions) by an expert panel of local business representatives and economic consultants.

Classification of sectors as knowledge intensive (KI)

The classification of sectors is based on the London Analysis, Identifying Science and Technology Businesses in Official Statistics that can be found at:

www.ons.gov.uk/ons/rel/regional-trends/london-analysis/identifying-science-and-technology-businesses-in-official-statistics/index.html

We rely primarily on the science and technology indicator provided in this publication and their analysis of how each SIC07 code is classified according to the OECD and Eurostat "High-tech statistics" publication available here:

http://ec.europa.eu/eurostat/statistics-explained/index.php/High-tech_statistics

We classify KI sectors as Information technology and communications; Life science and healthcare; High-tech Manufacturing; and Knowledge intensive services.

The OECD definition classifies knowledge intensive services (KIS) into four groups: KIS financial services; KIS market services; other KIS; and KIS other high-tech services. As explained below, in general we take only the last of these to be KI in our definition.

The final classifications of sectors adopted for the CBR database are shown, along with the number of companies in each sector, in the workbook called CBR Dataset sector classifications & counts Jan 2016.

In general, the KI sectors are: High technology and medium-high technology manufacturing and KIS high-tech services. Low and medium-low technology manufacturing, KIS market services, KIS financial services and Other KIS are generally classified as non-KI.

There are a few exceptions to these general rules and these are discussed below.

SIC 42220

This concerns construction projects for telecoms and electricity and we have taken these seven companies to be KI.

SICs 53100 – 53202

These are postal courier services and despite them being classified as KIS high-tech services by the OECD, we do not regard them as KI.

SICs 59111 – 59120

These sectors involve TV and video production and despite being classified by the OECD as Other KIS, we have classified them as KI.

SICs 59200 – 60200

These sectors involve TV and radio broadcasting and despite being classified by the OECD as Other KIS, we have classified them as KI.

SICs 69109 & 77400

These sectors cover intellectual property protection and despite being classified as KIS market services, we have classified them as KI.

SICs 71121 – 71200 and 74901

Despite these being classified as KIS market services, we have classified them as KI. In support of this decision, their science and technology category is given as 'Other scientific/technological services'. However, to be consistent we should also have classified quantity surveyors as KI, but we did not.

SIC 75000

This sector covers veterinary activities. After some debate it was decided to not classify as KI despite the importance of the racing industry in Newmarket.

SICs 86101 – 86900

These sectors involve human health activities and whilst some of the companies within these sectors may be classifiable as KI, we decided to classify these sectors as a whole as non-KI.

A summary of this classification and the number of companies in each sector on the dtabase can be found in http://www.cambridgeahead.co.uk/wp-content/uploads/2016/02/CBR-Dataset-Sector-classifications-&-counts-Feb-2016.pdf

Of course, it should be noted that we have classified sectors, not companies. There will be KI companies operating in non-KI sectors and vice versa.

Estimation of global employment and turnover figures

The vast majority of our companies are SMEs and take the full exemption in filing their accounts. This means that not much more than balance sheet figures are presented. This means that we need to adopt a variety of approaches to the estimation of employment and turnover. First we tackle employment, using accounting information for the approximately 1500 companies that provide it. The second approach begins by estimating employment from

total assets for the remainder. To do this we collect employment, turnover and total assets for all companies in the 50-249 size range on FAME. From these we calculate the median employment to total assets ratio and the median turnover to employment ratio for each SIC (amalgamating adjacent SICs when the number of observations is too small). We then use the appropriate employment to assets median to estimate employment for each company that has not reported it.

For about half of the companies we use this estimate of employment but not before it has gone through several checks. We have employment at the primary trading address provided for over 9,000 companies that did not report employment in their accounts. Where this is roughly consistent with the estimate from the median employment to asset ratio, we use the primary trading address employment as given. The next check is to compare our estimate of employment with the employment size range estimated by FAME, which led to about 1500 companies having the mid value of this size range used in preference to the asset-based estimate. For about 1000 companies the estimate of employment exceeded 50, the level at which companies are obliged to provide employment figures. In each of these cases internet searches led either to a more accurate figure being identified, or to the estimated employment being set to a maximum of 49 (152 companies). Occasionally, companies reported turnover figures but not employment and in these cases employment was also estimated on the basis of turnover.

For those without employment figures in the earlier years we use the assets figures for the previous years for that company as ratios to their current assets to calculate employment for the previous three years (with missing values when there is no asset figure for that year (ie births).

We now have employment best estimates that are robust and these can be used to estimate turnover figures by using the turnover to employment ratios. This was done for all companies with turnover not provided and calculated for all five years (with missing values where the employment figure is missing).

Andy Cosh, Assistant Director

Centre for Business Research, University of Cambridge

February 2016