25th Meeting of the Wiesbaden Group on Business Registers - International Roundtable on Business Survey Frames	
Tokyo, 8 – 11 November 2016	Andrew Allen ONS, UK Session No.2
	Name of Session
The redevelopment of the UK Business Register	

Introduction

The UK business register was initially developed in 1994, and since that time has been continually enhanced and improved. However, the technology used is now old and this is a critical part of the economic statistical infrastructure. So in 2016, as part of a wider data transformation program, ONS decided to embark on the redevelopment of the register.

The redevelopment project covers more than just a new statistical business register. Working with other government departments, ONS has identified the need for a new administrative database of businesses for the UK. This will lead to an expanded role for the business register within the UK.

Since the UK does not have a unique business number, this will be the first time the business population has been fully identified – the aim will be to produce a canonical list of businesses for wide use across government and beyond. The UK has a company registration system, but un-incorporated entities are only identifiable through tax records. This definitive list will then be used as the building block to populate a new statistical business register.

Both the new administrative database and the statistical business register will include new administrative data sources to improve coverage of very small businesses, compared to the existing statistical business register.

This paper provides an update on progress so far – we are a long way from the finish line.

New UK System

A new definitive list of businesses, called the Business Index (BI) is being developed. This will be used for more than statistical purposes, will be administrative unit based, and will not contain any statistical units. This will be used for a variety of administrative/policy functions, many of which are not statistical. The BI development will create a new unique number for UK business population, which will be the first time that the UK has had a business numbering system that will cover the entire population.   Initially,

this will only be deployed for the BI, but it will be freely available to cross government users, so it is possible that over time it will be more widely adopted.

The BI will also be used as the source data for a new Statistical Business Register (SBR). The aim is to extend the coverage of SBR, compared to the old Inter Departmental Business Register (IDBR) , by accessing new tax data sources. The main new data sources that will be required are Self Assessment and Corporation Tax. Self Assessment is used to collect tax in a variety of circumstances, but we are applying for access to the specific sections used by small sole proprietors and partnerships to help populate the new SBR. Legislation is in progress that should enable ONS to get wider access to this data for the first time.

It is important to note that the missing entities are by definition very small and have been previously accounted for by ONS in estimation . However, there is now a user demand to identify these individually for wider policy /administrative purposes.

Since there is no unique business number in the UK , a key function of the BI , will be the de-duplication of administrative records. This has always been an important function of the IDBR, but in the past this was part of the statistical unit enterprise creation. With the new index, de-duplication of administrative data will take place in the BI, and then this de-duplicated administrative data will be used as a feed into the SBR. We are currently reviewing the matching techniques used , with the aim of improving quality and minimizing clerical intervention for multiple matches.

The SBR will then convert this administrative data into enterprises, by using a combination of VAT group reporting records, business profiling and other survey contact with businesses. This method of creating enterprises is already in place in the IDBR, but the business index adds a preliminary match at administrative data level.

The new SBR is likely to hold more variables than the current register, the exact requirements will be determined during the discovery phase, but we already know that there is a strong user demand to hold assets data, and to develop the new sampling system to be able to stratify using assets.

The uses of the SBR have changed a lot since the IDBR was set up in 1994. There is now a strong user requirement to produce economic analysis directly from the register, which will require greater consistency between variable reference periods and changes to the way data are prioritized and time stamped. This also puts pressure on data timeliness, with employment and turnover variables being required to be up to date as possible, and all variables have to be fully time stamped.

SBR Boundary

The project team includes enterprise architecture colleagues, who are looking at the scope and boundary of the SBR.

Currently the UK IDBR covers the whole register process, but also includes some survey respondent management systems. We are also responsible for regular production of sampling frames, managing the business survey sampling and the coordination and overlap control between surveys. Finally we produce publications from the register about the business population e.g. business demography. It would be interesting to get feed-back on where the SBR boundary is drawn in other countries.

Agile Delivery Method and Progress to date

The projects are being delivered by a commercial IT company, using the well known Agile delivery mechanism based in three stages: discovery, alpha and beta. This follows UK Government Digital Service procedures. The discovery phase is about scoping the project, understanding users, and preparing for development. At Alpha stage a prototype to prove the concept is developed in close collaboration with users. The operational system is then built in beta, again in close collaboration with users. Key to this delivery methodology is user story driven development and regular delivery in small increments.

Both products are being delivered using the latest technology, this includes Hadoop and Cloudera. There is potential for great improvement in register management and data handling from the use of these and other new technologies which we are only just starting to explore.

The BI is in the Beta phase. A working prototype was produced in the alpha phase , which proved the concept and was showcased to many external government and ONS stakeholders.

The SBR is in the discovery phase , which is about scoping the development and engaging with users on their requirements for a new system.

The new systems will also lead to changes within the Business Register Division organizational structure, so these are also being planned to enable the operational running of the new systems.

Timescale and Transitions

The aim is to get the BI in service by spring 2017 and the SBR by early 2018. A key issue will be the transition by survey users from the old IDBR and onto the new SBR. Experience of making changes to the IDBR indicates that this transition will need to be handled carefully. A team will be set up containing business register and business survey members to plan and manage the switch over. This will include measuring any discontinuities, and assessing how to treat these in survey outputs.

Next Steps

Delivering a production quality Beta for the BI is the important next step, along with identifying customers who are keen to trial the new BI via an Application Programming Interface (API).

In parallel the user requirements for the new Statistical Business Register will be gathered, with the aim of building an alpha prototype early in 2017.