**Statistics Canada International
Symposium Series - Proceedings**

Symposium 2005 :
Methodological Challenges for
Future Information needs

2005

**Statistics
Canada**   **Statistique
Canada**

Canadä

RE-DESIGNING THE UK'S BUSINESS REGISTER AND BUSINESS SURVEYS TO ENABLE PRODUCTION OF BETTER ECONOMIC STATISTICS

Mark Pont, Mike Hidiroglou, John Perry, Paul Smith, Pam Tate

ABSTRACT

The UK's Office for National Statistics (ONS) is starting a development programme for business surveys to meet the recommendations of a recent government report calling for improvements to economic statistics, in particular regional economic statistics. An important aim is to provide good quality Gross Value Added estimates for the UK's 12 *Government Office Regions*, and improved supporting statistics for small areas as part of an integrated system producing both National and Regional Accounts.

This paper describes some of the issues faced by ONS and describes our plans to meet these requirements through two channels. First, re-engineering the Inter-Departmental Business Register (IDBR), which is the sampling frame for ONS's business surveys. Second, integrating some of our business surveys and introducing common methods for them. The work will take advantage of new tools being developed in ONS's modernisation programme. The main areas of statistical development include a new stratification and accompanying domain estimation strategy; and improved methods of recasting data to business entities at different levels.

KEY WORDS: Business Register; Modernisation; Re-engineering; Regional Statistics

1. INTRODUCTION

1.1   Background

A series of initiatives, reviews and objectives over the last couple of years has affected the vision for the provision of economic statistics in the UK, particularly for statistics produced from the suite of business surveys. The main drivers are:

- the modernisation of ONS's statistical systems;
- the re-engineering and integration of business survey systems to take advantage of the new systems when they become available;
- the re-engineering of the ONS's Business Register to move it to a new computer platform and to introduce new features which will be needed to support other changes; and
- an external Review of Statistics for Economic Policymaking (Allsopp 2004); the *Allsopp Review.*

These four strands are all strongly interrelated in their effect on business surveys in the ONS, and it is not possible to disentangle the developments that are proposed to address them. Together, however, they promise to provide the impetus for a radical shake-up of the business surveys run by the ONS. There have been large changes in the system of business surveys before, but these have normally been restricted to certain sectors (for example from the introduction of the monthly and quarterly surveys in the service sector in the early 1990s) or to certain periodicities (for example the merging of annual surveys to make the Annual Business Inquiry in 1997-8).

The extent of the proposed changes is therefore wide, and requires a substantial period of time in development and implementation. The number of changes is perhaps no greater than a series of similar changes at times during the 1990s, but the difference here is the need to integrate the changes to make a coherent system while maintaining business as usual. In short the challenge is to make the changes into a whole programme, and to manage the dependencies appropriately.

In this paper we describe two main elements of the development programme – the re-engineering of the Inter-Departmental Business Register and the integration of surveys. The paper is structured as follows. A summary of Allsopp's objectives are provided in Section 2. The Business Register is described in Section 3 and the redesign required to support the Allsopp work is covered in Section 4. The proposed plan for integrating business surveys and its supporting methodology is discussed in Section 5. Some concluding remarks are made in Section 6.

2. THE ALLSOPP REVIEW

The Allsopp Review covered two main elements: the demand for and provision of regional information; and whether official economic statistics adequately reflect the changing structure of the economy. On the first of these, the broad findings were that there is a need for improvements to regional data, including: good quality and timely estimates of annual Gross Value Added (GVA) for the UK's 12 government office regions; a range of timely indicators to assess short-term regional activity; and a more timely measure of sub-annual regional GVA growth, based on short-term surveys. On the second element, the broad findings were that the changing structure of the economy has resulted in imbalances, particularly between the treatment of the manufacturing and services sectors, which need to be reversed. A principal requirement for industry sector detail is to support production of a coherent and consistent framework for meeting core demands for economic statistics, and from this perspective the imbalance arises not so much from any sectoral skewness in the main business surveys, but rather from a lack of product detail and incomplete price information on the service sectors. The Review recommended a review of the level of industrial stratification for business surveys, introducing more flexible procedures, and a statement of the national, regional and industrial priorities for production of statistics.

The consequent recommendations for the Business Register and business surveys include an extension of the IDBR and associated methodological techniques to play a more sophisticated role in managing inputs and outputs of the business survey process, as well as enlarging the ABI sample to provide better regional estimates. Sample sizes of the monthly turnover surveys may need to be increased to provide better advance annual measures. Estimation of product sales for all industries, with timeliness and detail are to be driven primarily by provider rather than user considerations. ONS should move towards a fully coherent business survey system in the long term and should assess: the relevant roles of survey and administrative data; the extension of surveys to broad industries not covered at present; and the combination and integration of surveys.

ONS plans to address these objectives by re-engineering its Business Register, and by redesigning and integrating some of the main annual and sub-annual surveys.

3. THE INTER-DEPARTMENTAL BUSINESS REGISTER

3.1 Introduction

The Inter-Departmental Business Register (IDBR) is a list of UK businesses maintained by ONS. It is used for selecting samples for business surveys and to produce analyses of business activity. The IDBR plays an important strategic role in official statistics, and a good register is essential to underpin production of National Accounts.

The IDBR covers all businesses registered for Value Added Tax (VAT) or employers operating Pay As You Earn (PAYE) income tax schemes. As a result, the IDBR covers both the private sector, including non-profit organisations, and the public sector, including public administration.

3.2 Statistical Units

The two important statistical units in the present context are the enterprise and local unit (LU), each of which is defined precisely in the EU Regulation on Statistical Units (EEC 1993) as follows. The enterprise is 'the smallest combination of legal units that is an organisational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources. An enterprise carries out one or more activities at one or more locations. An enterprise may be a sole legal unit.' *The local unit* is defined

as 'an enterprise or part thereof (e.g.: a workshop, factory, warehouse, office, depot or mine) situated in a geographically identified place. At or from this place economic activity is carried out for which – save for certain exceptions – one or more persons work (even if only part-time) for one and the same enterprise.'

Sampling and mailout are carried out using the *reporting unit* (RU). The RU holds the mailing address for the business and is the unit for which businesses report their survey data to ONS. In general, the RU is the same as the enterprise. In some complex cases, enterprises are subdivided into RUs according to activity or geography, defined by grouping the appropriate LUs from within the enterprise.

IDBR business structures are updated from an annual survey known as the Business Register Survey (BRS). The BRS collects information from the largest businesses and a sample of smaller businesses about the business activity and number of employees at each site operated by the business. The BRS is designed, together with the Annual Business Inquiry (ABI), to meet users' requirements for the provision of statistics for small geographical areas.

3.3 Sampling and Data Collection Arrangements

Sampling is carried out using a Permanent Random Number (PRN) method (Ohlsson 1995), with stratified simple random samples taken. Most business surveys are stratified by industry and size, generally the number of people employed in a business, although some use the turnover of the business. The current selection system permits only one size variable to be used, and the system is not conducive to the definition of strata using non-consecutive codes from the Standard Industrial Classification (SIC(2003)) (ONS 2002). Therefore, the current stratification system, whilst meeting users' needs in the early 1990s when it was developed, does not offer sufficient functionality to meet today's wider information needs.

The unit chosen for sampling in most surveys is the RU. As mentioned earlier, the RU generally equates to a whole enterprise, but in some complex cases, enterprises are partitioned into two or more RUs. These arrangements mean that there is inconsistency between the levels used for selection, and restrict the scope for setting up different data collection structures for different surveys. The system is also sensitive to businesses restructuring, in which case potentially important linkages between the old and new structures may be lost. In addition, business structures are of mixed currency. Finally, employment values for each RU on the IDBR are mainly updated once a year, which can cause jumps in estimates.

We outline in the following section how these weaknesses will be addressed.

4. THE PROPOSED BUSINESS REGISTER

4.1 Sampling and Data Collection Entities

The changes that will be brought to the IDBR are in response to a range of developments as follows. First, ONS is migrating to new computer systems, which will require some reorganisation of IDBR data. The new Business Register will reflect changes in the European Regulation on business registers; an important change here is a better operational definition of the enterprise. Demand for statistics at the regional level and lower is growing. Part of this demand could be met by better use of data from administrative sources. The new Business Register will play an important role as the gateway for such data relating to businesses, and the hub to facilitate the linkage of administrative and survey data. Finally, national and international demand is increasing for data on business demography based on statistical business registers.

At the same time it is proposed to enhance the flexibility of data collection structures to meet changing user requirements. As noted from section 3.2, RUs serve two purposes; they are both sampling and collection units. The introduction of data collection entities (DCEs) will ensure that one set of units is dedicated to data collection. These units will be built to maximise the usefulness of data requested from surveyed enterprises, while ensuring that sampling takes place at a consistent level – the enterprise – for all businesses. The use of a standard unit for sampling is an important aspect of the co-ordination leading towards data integration. Because of the variety of ways businesses may be able or prefer to provide information, there is a need to identify for each survey, the arrangements

by which the data items are to be acquired for each sampled enterprise. These arrangements are the attributes of a DCE. The attributes are divided into three sets: (i) *coverage:* defining the relationship between the entity from which the data are being requested and the statistical entity for which the data are required; (ii) *collection instrument:* the means of obtaining the data, eg questionnaire, telephone interview, administrative record, etc; and (iii) *contact:* the respondent name, address, and telephone number within the business operating structure.

DCEs are created only for selected businesses and are survey-specific. Ideally, we want to generate as many as possible automatically by an algorithm based on statistical entities on the IDBR, associated administrative links, and how data are to be collected from the respondent. They can be modified manually to take into account information fed back from respondents. Most enterprises operate at a single location, so the DCE structure in most cases will be very simple to reflect this. In the more complex cases, we expect DCEs to be defined either on a regional or an industrial basis, or by region and industry. In the most complex cases, more flexibility may be required in order to define DCEs along a mix of industrial, regional, organisational and other lines. Further work needs to be done to evaluate the extent to which it will be possible to automate the system, given the diversity of structures of businesses in the UK.

The increased importance of data collection structures places greater emphasis on the process of identifying and maintaining the structures of such businesses, known as profiling. Profiling needs to be sufficiently well-resourced and responsive to maximise the benefits arising from the new proposals.

4.2 Auxiliary Variables for Stratification and Estimation

The second major statistical benefit anticipated from re-engineering the IDBR is improved stratification. Stratification in ONS has through custom been restricted to rectangular strata, and stratification by a very restricted range of variables. As a result of the historical focus on activity and the proliferation of activity codes in coding systems such as NACE and its UK version – SIC(2003), very detailed industrial stratification has meant that for some surveys, the population of businesses in many strata is very small.

In order to implement the recommendations from the Allsopp Review, region becomes a more important stratification variable. Sample sizes are not large enough to support adding regional strata to the current detail, so we need a new strategy. We therefore need to ensure that the re-engineered IDBR is sufficiently flexible to accommodate the range of stratification options that may be required in the future, including the following points.
- Strata need to be sufficiently large. Cochran (1977, page 153) recommends around 30 units for large-sample methods to be valid, but in practice such large sample sizes are impractical within the current structure of ONS's business surveys. In recent redesigns (notably for the Retail Sales Inquiry) we have targeted an achieved response (that is, after expected non-response has been factored in) of 20 questionnaires per stratum, as an improvement over previous designs which sometimes had minima as small as 5 *sampled* (not responding) units. The use of a combined ratio estimator can ensure that the number of responses used to estimate ratios is more appropriate.
- Less detailed industry stratification and a more versatile use of the different levels of the SIC according to size of populations, homogeneity of SIC classes etc. Reduced industrial stratification is the price for additional control at a regional level, and this means increased variability in some estimates and the use of domain estimation techniques.
- The availability of additional variables for stratification and estimation. Ideally these will be easy to create, so that stratification is not limited to a small set of variables. Any information, or combination of information, should be available for stratification as long as it is available on or through the Register. This could include taking into account the complexity of businesses in stratification. Given the increased demand for regional statistics, we may want to have better control over the coverage of businesses that span more than one region or more than one industrial sector.

5. BUSINESS SURVEY INTEGRATION

5.1 Integration

In collecting data about businesses for statistical purposes, the eventual goal is to use data extracted directly from businesses' own accounting and management information systems, and/or from their published accounts, tax returns

and other data supplied to governmental bodies. However, various legal, administrative and technical barriers have to be overcome before this is possible, and it will take some considerable time to reach this goal.

In the meantime, there is a clear need to respond to the increasing demands for scope and quality of economic statistics, and ONS needs to find ways of meeting them, while keeping the costs to both itself and businesses acceptable. Hence, ONS has instigated a major programme of work that meets Allsopp's objectives as well as addressing some wider issues relating to the integration of the major business surveys.

5.2 Business Register and Employment Survey

The major part of the integration currently underway in ONS is to combine the existing BRS with the Annual Business Inquiry (part 1) (ABI/1), which presently have quite similar sample designs. BRS was described in Section 3.2. ABI/1 is the survey used to produce employment estimates at a range of levels of detail. Employment values are collected from RUs, and used to produce national estimates. Regional estimates are produced from values that are imputed for LUs based on the RU employment value being recast to the latest-known business structure according to a model.

The proposed Business Register and Employment Survey (BRES) yields improvements to both the Business Register and employment estimates. Employment estimates are improved since estimates will be produced from data collected directly from local units, thereby removing the errors associated with recasting data to an out-of-date frame. Both the Business Register and employment estimates benefit by having a larger sample under the combined scheme, whereas the cost to respondents for the combined surveys should be less than for the total of the individual surveys.

The BRES also gives an opportunity to investigate the possibilities of collecting a wider range of data at local unit level that may subsequently be used in regional estimation in other surveys. The first step in capturing and using more data on local units of multi-site businesses is to establish what data the businesses have available. During early 2005, ONS conducted about 100 interviews with a range of businesses to study the availability of data at local unit level, for the following variables: employment; turnover; capital expenditure; inventories; purchases; employment costs; profits and the level and cost of production. Quite surprisingly, the study found no real differences between economic sectors or the size of business. The main determinant was the availability and use of computer software, particularly accountancy software. In general, the study found that turnover, employment and employment costs were fairly widely available from businesses, or were easily obtainable, with the ease of access related to the level of computerisation. Data were available from a minority of local units relating to purchases; profits; inventories and capital expenditure. Again, availability depended on the level of computerisation. Finally, information about the costs of production was largely unavailable for local units since records tended to be kept centrally.

This study was quite small, and didn't cover the largest businesses. Furthermore, the study didn't go as far as collecting data from businesses and the practice may be different from the theory. However, this work gives a reasonable platform upon which to base a system for the collection and production of regional statistics. Since data appear to be largely available at the lowest level, they can be aggregated to any appropriate regional definition. Further work is needed in this area to get a clearer understanding of these issues, particularly for the largest businesses.

5.3 Productivity estimates

The integration of BRS and ABI/1 requires that the two present parts of ABI are separated. Part 2 of the survey produces annual estimates of financial variables, and data are collected from a common sample in both parts of the survey. One concern in separating these two parts was that the quality of 'per head' estimates would deteriorate because the two surveys were not linked, and the correlation between the numerator and denominator would be lost. However, our study showed that at the very worst, this worsening would be very small. This arises mainly because most of the correlation is already accounted for in the use of the ratio estimator to produce both the output and employment estimates. The remaining correlation is in fact then only a correlation between residuals after the fitting of the ratio model. Splitting the surveys means that each can be allocated more optimally, thereby recouping some or all of the small quality loss associated with the split.

5.4 Updating employment values on the IDBR

The present strategy for updating the IDBR from the BRS is to update employment only for those businesses that responded to the survey during the current year. Therefore, some auxiliary information is out-of-date; very much so in some cases. It is proposed to create a model from the relationship between returned employment values in BRES with auxiliary information already held on the IDBR in order that a new modelled auxiliary variable can be created that is consistent across all records on the Register.

5.5 Other areas of integration

The other two main areas of business survey integration will be taken forward on a longer timescale. These are a redesigned Annual Business Survey to meet the needs for annual data from businesses contributing to GDP estimates, at both national and regional levels, and a Monthly Business Survey to meet the needs for sub-annual indicators of economic performance. This will involve merging a range of existing monthly and quarterly surveys, and making more use of data from administrative sources in order to reduce respondent burden.

6. CONCLUSION

The Allsopp review has provided a powerful motivation for reviewing ONS's procedures for business surveys. The first step in the chain of processes within business surveys is of course the Business Register. In this paper we have set out a range of developments, some building on features which have been incorporated previously but under-utilised, and some which are new but which are motivated by challenges in using the existing register to support the surveys. The survey designs that are based on a newly redesigned Business Register must profit as much as possible from the data that reside on the register, in terms of consistency of approach (it is already a maxim for statistical offices that integration and standard methods are the only ways in which a cost-effective survey system can be produced with ever-diminishing resources, while savings are diverted into new developments and outputs), sampling efficiency and provider-friendliness through flexibility in data collection.

The flexibility in business structures for data collection proposed here should allow us to develop more provider-friendly methods, as well as potentially increasing the range of geographic (and in some cases other) detail to support estimation. It is clearly better to base estimation processes on real data, and we have yet to evaluate the trade-off between making provision easy and estimating with precisely the data we want. But the framework outlined in this paper will allow us to evaluate this trade-off and provide a setting within which the additional methods that are needed can be developed. The last plank of the structure on which the business surveys system will be built is administrative data. These already underpin the register, but there is the potential to use them in certain circumstances to replace parts of the survey collections, presumably with some loss of quality (perhaps definitional, perhaps timeliness), but with the potential to reduce the respondent burden and perhaps to increase the amount of data available. In some cases it appears that legislation may be needed to increase the availability of such data to ONS, but at a time when we are rebuilding ONS's survey systems, now is the opportunity to make sure that those systems are sufficiently flexible to permit the easy incorporation of administrative data and associated methods.

REFERENCES

Allsopp, C. (2004), *Review of Statistics for Economic Policymaking*, Norwich: HMSO.

Cochran, W. (1977), *Sampling Technique*s, New York: Wiley.

EEC (1993), *Council Regulation No 696/93 on the Statistical Units for the observation and analysis of the production system in the Community*.

Ohlsson, E. (1995), "Co-ordination of samples using permanent random numbers", in B.G. Cox et al. (eds.) *Business Survey Methods*, New York: Wiley, pp. 153-169.

ONS (2002), *UK Standard Industrial Classification of Economic Activities 2003*, London: The Stationery Office.