



Catalogue no. 11-522-XIE

**Statistics Canada International Symposium
Series - Proceedings**

**Symposium 2003: Challenges
in Survey Taking for the Next
Decade**

2003



Proceedings of Statistics Canada Symposium 2003
Challenges in Survey Taking for the Next Decade

GETTING THE MOST OUT OF THE PORTFOLIO OF UK GOVERNMENT HOUSEHOLD SURVEYS: THE OFFICE FOR NATIONAL STATISTICS' PLANS FOR AN INTEGRATED SOCIAL SURVEY, THE CONTINUOUS POPULATION SURVEY (CPS)

Tony Manners¹

ABSTRACT

ONS will be developing and trialling the Integrated Social Survey (ISS) over 2003-06, with fieldwork starting in January 2007. The ISS aims for integrated outputs from the major government surveys, based on an integrated source. It will combine collection from a single household sample of data for all the topics currently collected in our labour force survey, household budget survey with diaries, food and nutrition survey and two multi-purpose surveys (one detailed, with annual data; the other for fast turn-round for monthly data). The aggregate mean interview length for these surveys is currently some 5 hours + the diaries. There will be a "whole-sample" module of key census-type variables from which we will be able, given the aggregate sample size, to provide more frequent estimates for smaller areas than at present; and more precise annual estimates than the current national sources. These variables will serve also as the major classifiers for the detailed information on income, expenditure, housing, family growth, labour-market related matters, health etc which will be collected in viable "interview combinations". The interview combinations will each comprise the whole-sample module and "part-sample" modules of data which must for analytical reasons be collected on the same sample cases. We anticipate some 6 or 7 interview combinations: we will seek to reduce respondent burden and improve response compared with the current set of 4 independent surveys. There is a major programme of methodological work on the data collection, weighting and estimation issues. Operationally, new case management and management information systems will be developed to handle the output of individual modules from within as well as between "interview combinations" to their own optimal timetables. There are significant challenges for interviewer training to meet the demands of the integrated survey; solutions will be based on successful trials in the 1990s.

KEYWORDS: Coherence of Official Statistics; Household Survey Integration; Inter-Censal Statistics; Modular Instrument Design.

1. INTRODUCTION

1.1 There is a need for National Statistics (NS) in the UK to provide better information on key social and socio-economic variables such as housing, employment, ethnicity, education and health. This will improve inter-censal monitoring of key variables for a range of policy purposes and small area statistics for the Neighbourhood Statistics Programme (NeSS). At present, the Labour Force Survey (LFS) is the largest single source of survey data available for monitoring purposes and produces estimates to a level of precision not matched by any other UK household survey. Together with local LFS boost samples in England (LLFS) and Wales (WLFS), the LFS sample provides annual estimates for a range of indicators (mainly concerned with the labour market and adult training targets), down to Local Education Authority (LEA) level (similar in size to county). To provide a wider range of information at (the smaller) Local Authority District level (mean population approx. 130,000 but wide variation) would be a big step forward for NS, but to achieve this a much larger sample of households is needed. ONS is in the process of developing its ideas for the integration of the ONS-led continuous household surveys to meet this and related objectives. The integrated survey is now under development. It is known as the Continuous Population Survey (CPS).

¹ Tony Manners, Office for National Statistics, United Kingdom, Tony.Manners@ons.gsi.gov.uk

The ONS surveys and the contribution they could make each year to the CPS are:

| | Approx. no. of GB households ('000s) for annual independent sample | Approx. no. of GB adults 16+ ('000s) for annual independent sample |
|---|--|--|
| Labour Force Survey (annual) | 96 | 190 |
| Local Labour Force Survey (LFS annual boost) | 40 | 70 |
| Welsh Labour Force Survey (LFS annual boost) | 15 | 28 |
| Omnibus | 14 | 14 |
| General Household Survey (GHS) | 9 | 18 |
| Expenditure and Food Survey (EFS) | 7 | 14 |
| <i>Sub-total</i> | <i>181</i> | <i>334</i> |
| Additional survey (Annual Population Survey, APS) | 65 | 113 |
| Total | 246 | 447 |

In addition to the current ONS surveys, there will be a further addition to the sample to meet the needs of the Neighbourhood Statistics Programme. This addition, called the Annual Population Survey (APS), will boost the LFS sample in selected local areas to meet minimum requirements from January 2004; it will report from the total combined sample. With the contribution of the LFS boosts and the APS, an integrated survey might achieve interviews with more than 240,000 households and over 400,000 adults. If this development is successful, the addition at a second stage of other NS continuous household surveys currently conducted outside ONS would increase the sample size.

Proposals for integration were investigated previously in the 1990s². There are now more business drivers. In brief:

- there is increasing demand for small area statistics which cannot be met with current arrangements: there is particular pressure for information on ethnic group information and measures of income but other variables are also needed;
- the Neighbourhood Statistics work adds to this demand: larger samples would also improve model-based estimates;
- the investment in surveys should be exploited more fully between censuses;
- there is a need for greater coherence in NS statistics, which integration of sources will promote;
- there is already a programme of harmonisation of questions and classifications across government; integration would take advantage of, and further develop, such standardisation;
- there is increasing demand for more surveys: in particular, possible EU Regulations obliging member states to conduct a survey of income and living conditions (SILC) and to make design changes to the LFS;

2. PROPOSED DESIGN

The proposed design concentrates on integration of the continuous surveys on which ONS leads: LFS, EFS, GHS and NS Omnibus. The design is compatible with continued delivery of all current outputs of these surveys, including those which are not annual estimates; this is particularly important for the LFS and its boosts. The title for the project is the Continuous Population Survey (CPS).

² K Deacon & T Manners (1997) Integrated Household Survey: Report on Field Trials in Survey Methodology Bulletin, 40, January 1997, ONS, London; pp.37-41

The CPS will cover Great Britain and allow for the addition of Northern Ireland data to provide UK results as at present.

The aim is to redesign the questionnaire content in terms of modules, covering *all* the content of the existing surveys. A fixed core of key variables, e.g. demographic information, household composition, economic status (employment, unemployment, inactivity), ethnicity, tenure and banded household income, will be included in every interview to provide a very large sample. Other modules will be asked of only part of the sample, e.g. the labour market module would be very similar to the current LFS preserving the current panel design required by Eurostat and other current users. The income and expenditure module would be very similar to the existing EFS with diary component. A simplified illustration of how modules might be combined is shown diagrammatically at Appendix A.

The CPS design has taken note of relevant developments in other NSIs, in particular in Statistics Netherlands, and of earlier work by ONS towards an integrated household survey which has demonstrated its operational feasibility.

The design will allow for addition of other continuous surveys at a later stage, if their sponsors choose.

The form of integration proposed for the initial stage of the CPS involves:

- a single sample of addresses drawn from the Postcode Address File;
- a survey instrument comprising all the topic modules to meet the information needs currently met in the separate surveys;
- *whole-sample* topic modules which provide, for all CPS households and persons, information on key variables which require samples of that size;
- *part-sample* topic modules which provide information on variables for which sufficient precision to meet policy needs can be obtained from just part of the CPS sample;
- combination of the *whole-sample* topic modules with selected *part-sample* topic modules to make a viable *interview combination*, i.e. making up an interview which flows acceptably for respondents; there will be a small number of different *interview combinations* so that all *part-sample* topic modules are covered; a diagram illustrating the idea of interview combinations, with hypothetical examples, is at Annex 1;
- interpenetration of the *interview combinations* in interviewer workloads (*mixed workloads*). It is essential to a cost-effective CPS that each interviewer carries out the range of different combinations in a local area, to minimise travel costs.

3. BENEFITS

The potential improvements from the CPS approach will include a greater precision of estimates without increasing costs, resulting from:

- a larger independent annual sample for the variables in the *whole-sample* (core module) than is currently available from the LFS and boost surveys alone
- an unclustered design, which would be an improvement in the precision of estimates for all variables except those particular to the LFS (which are already based on an unclustered design) without the need for increased interviewing costs.

The CPS also provides the opportunity for more regular, coherent reporting of estimates from a single source. At present, even with harmonised questions across surveys we have problems with coherence. For example, a key pioneering role of the GHS was to monitor growth in one parent families and cohabitation. Over time, the LFS has been used for the same purpose, given its larger sample size. Although the differences in estimates are small, annual results from the two surveys give different results.

The CPS could provide more regular estimates for *whole sample* topics because of the sample size on which estimates are based (the exception being the labour market variables). At the household level, indicators relating to tenure, accommodation type, and amenities, for example, could be produced on a quarterly basis with greater

precision than the annual estimates from current 'best sources'. Likewise, at the person level, for indicators relating, for example, to marital status, cohabitation, health, and provision of care.

For more regular reporting of key labour market indicators - economic activity, employment and unemployment rates - even a large boost to the current sample size would have a very limited impact on the variances of monthly change unless the design was changed to a monthly panel.

The CPS *whole sample* module will also enable users to analyse a number of key variables in combination, with a considerably larger sample than is possible from existing surveys, such as the GHS and EFS. For example, there will be improved precision for measures such as average household income, % reporting a limiting longstanding illness, % reporting health is not good, % living in flats or maisonettes, by ethnic group using information from the CPS *whole sample*. Estimates of precision for these types of output will be calculated as part of the development work for the CPS.

As well as more regular reporting, there is increasing demand for small area statistics which cannot be met by current survey arrangements. The existing surveys are designed to provide regional/country estimates and only the LFS is large enough to estimate for smaller areas. Together with local LFS boost samples in England (LLFS) and Wales (WLFS), the LFS sample provides annual estimates for a range of indicators (mainly concerned with the labour market and adult training targets), down to Local Education Authority (LEA) level (similar in size to county). To provide a wider range of information at (the smaller) Local Authority District level would be a big step forward. The CPS would also provide a significantly larger sample to improve small area estimates as part of the Neighbourhood Statistics Programme.

It will be possible to use the CPS *whole sample* and *part sample* output separately or in combination to paint accessible and comprehensive pictures of society in general and to explore specific topics in detail. There will be scope for short, topical outputs from the CPS on a regular (quarterly) basis and for multi-topic analyses and articles with the Web at the centre of providing the new look outputs.

4. CONSULTATION AND DEVELOPMENT WORK

ONS believes the arguments for developing in this direction are compelling. Informal discussions suggest some other government departments will be attracted to join in at later stages. For the initial stage, ONS is consulting with departments who pay into the four ONS-led surveys.

Resources will be needed to develop the fieldwork, methodological and computing infrastructure for the CPS, and to test it thoroughly. We will need to test the new style of interviewer work package and, in particular, to check the potential impact on response. Much of the methodological work and testing has already been done during previous exercises, but important new work remains to be done on sample design and estimation.

Thorough testing will be required to ensure that there are no insurmountable problems for the delivery of high quality data as a result of the transition from the current survey designs to the CPS. In particular, ONS will need to maintain the continuity of key time series, e.g. macro-economic time series produced from the LFS. Care will be taken at each appropriate stage of development to measure and to minimise the impact of survey changes on the continuity of key time series. If, in spite of these precautions, any unavoidable discontinuities arise, ONS will consult fully with users about ways in which continuous historical time series for the key variables can be estimated.

5. TIMING

The CPS is at an early stage of development. Staff for development are being recruited. Key milestones are (1) a dress-rehearsal in autumn 2005 of all systems and of sufficient size and design to test for effects on key estimates; (2) a decision point in June 2006, following investigation of the dress rehearsal results; (3) if the decision is to go ahead,

start of data collection in January 2007; (4) first annual results in the first half of 2008; requirements for other outputs for other periods, particularly for labour market statistics, will be to current timetables or better.

ONS is developing a range of major improvements at present. Some of the office-wide initiatives (such as the Statistical Infrastructure Development Programme, Information Management Programme and Labour Market Re-engineering Project) will deliver new methods, tools and infrastructure that ONS will be able to take full advantage of in developing an CPS to provide world-class survey taking. Given these developments, and the interaction between them, it is likely that January 2007 will be the earliest possible start date for data collection on the CPS.

Appendix A Illustrative diagram of the CPS' Modular Structure

The diagram is simplified for illustration. It shows the basic (*whole-sample*) module and 5 *interview combinations* of the whole sample module with *part-sample* modules. Some of the *part-sample* modules are covered on more than one *interview combination* to achieve the required sample size. Modules with particular requirements, such as the detailed labour market modules, are in *interview combinations* which meet their design needs. The detailed labour market *interview combination* has a 5-wave rotating panel design. In the simplest conception of what it will contribute to the CPS, it might contribute only data from a single wave for any one sampled household. Some key labour market concepts, such as the ILO definition, will be measured in the *whole-sample* module, i.e. on the largest possible sample.

