UNITED NATIONS **E**

Economic and Social Council

Distr.
GENERAL

CES/2001/14
7 May 2001

Original: ENGLISH

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Forty-ninth plenary session
(Geneva, 11-13 June 2001)

Characteristics of Sample Surveys in Ireland

Invited paper prepared by the Central Statistical Office of Ireland [1]

I. Introduction

1. The dominant characteristic of sample surveys in small countries is that they cost substantially more in relative terms than they do in large countries. This of course reflects the known statistical fact that for samples with small sampling fractions the accuracy of the estimates is largely determined by the absolute sample size. This is a particularly important issue in surveys where the data are collected by personal interview; e.g. most household surveys, since the field expenditure is the dominant cost component of the survey.

2. The field costs associated with household sample surveys can in fact be very significant items on the overall budget of National Statistical Institutes (NSIs) in small countries. As a result, the range of surveys undertaken tends to be narrower, the frequency of mandatory surveys lower, and the survey methodologies employed often tend to be sub-optimal in order to contain costs. Countries with well developed administrative registers of the population and/or households have however the potential to achieve substantial savings through linking administrative and survey data to positive effect.

3. The pressure on statisticians in small countries is therefore to find the most cost-effective means to design and conduct their household surveys. The responses have included efficient sample designs, development of omnibus surveys (to reduce the number of individual surveys), linkages

[1] Paper prepared by Mr Gerry O'Hanlon, Central Statistical Office, Ireland

between surveys and exploitation to the extent possible of cheaper survey methods such as postal and telephone surveys rather than personal interviewing.

4. The problem is essentially the same for business surveys i.e. surveys cost more in small countries in relative terms. However, the collection costs are not the main concern as these surveys are usually conducted by post, telephone or, in an increasing number of cases, electronically. Instead, the problem manifests itself in a different manner because the structure of enterprises and the survey objectives are markedly different. Specifically, the structure of business enterprises is usually dominated by a relatively small number of large concerns and there is a demand to provide quite detailed results across the whole spectrum of economic activity. For a small country the combination of these circumstances means that the sampling fraction for medium to large enterprises often needs to be one (i.e. a complete census of firms above moderate size thresholds is required) in order to meet survey objectives. This frequently gives rise to response burden issues (particularly when the same firm is automatically included in all relevant surveys) and problems with preserving statistical confidentiality at the dissemination stage.

5. In this paper the experience of Ireland as a small country is examined in relation to the above issues. The situation of Ireland is particularly interesting since our statistical programme over the past thirty years has been driven by the need to meet the demands of the EU statistical system. The "one size fits all" focus of the latter means that small countries such as Ireland must to all intents and purposes match the statistical programme of larger countries. This of course causes particular problems in the case of sample surveys where the economy of scale does not favour the small country.

II. Household surveys in Ireland

The absolute sample size problem

6. We know from our statistics theory that the standard error of estimates based on samples that involve small sampling fractions are approximately proportional to the inverse of the square root of the absolute sample size. What this means in practice is that if we wish to halve the standard error associated with an estimate we must, all other things being equal, quadruple the sample size. This is bad news for small countries since it means that, in order to achieve a given level of precision; their samples must effectively match those taken in much larger countries. There are many valid reasons why larger countries might need to use larger samples – for example in order to identify sub-populations, regions etc. However, in practice the economy of scale works very much in favour of the larger country. Table 1 indicates the target national samples proposed for a new survey on income and living conditions in the EU. The target samples took into account the need to have larger samples in the larger countries in order to provide a better level of precision at EU level.

Table 1: Proposed Target Sample Sizes for new EU-SILC(1)

Country	Population (million)	Target Household Sample
Luxembourg	0.4	3,500
Ireland	3.6	4,500
Finland	5.1	5,000
Denmark	5.3	5,000
Austria	7.9	5,500
Sweden	8.5	5,500
Portugal	9.9	5,500
Belgium	10.1	5,500
Greece	10.2	5,500
Netherlands	15.3	6,500
Spain	38.9	8,500
Italy	56.9	9,500
France	57.2	9,500
United Kingdom	57.7	9,500
Germany	81.0	11,000
EU Total	**368.0**	**100,000**

(1) EU –Survey on Income and Living Conditions

7. The table speaks for itself. Germany, with a population 200 times that of Luxembourg, will need a sample only slightly more than three times that required in the latter.

Impact on the statistical budget

8. Ireland does not have comprehensive or integrated registers of persons or households and thus the provision of statistical data on the population is met, almost exclusively, from traditional censuses or sample household surveys.

9. Migration, particularly external migration, has been a major influence on Irish demography over the past 150 years or so. In the absence of administrative registers, and with freedom of movement between Ireland and its nearest neighbour the UK, regular censuses of population are the only means available for measuring migration accurately. Over the past 50 years traditional censuses have been conducted every five years in years ending in a '1' or '6'[2]. These censuses have involved the use the Enumerators for both the delivery and collection of forms and, as a result, are extremely expensive undertakings in the context of the statistical budget.

10. Since 1975 the Central Statistics Office (CSO) has been conducting labour force surveys as part of the mandatory EU programme. Up to 1983 the surveys were conducted every two years, from 1983 to 1997 they were held annually and a continuous quarterly survey was introduced in

[2] The one exception was 1976 when the census was cancelled for budgetary reasons but the lack of accurate data prompted the Government to undertake a limited headcount census in 1979 – only two years ahead of the next scheduled full-scale census in 1981.

September 1997. The principal purpose of the LFS (or Quarterly National Household Survey as it is now referred to in Ireland) is to track developments in the labour force with a high degree of precision. To achieve this objective a sample of around 40,000 households i.e. between 3 and 4% of all households is required on each occasion.

11. The third major household survey conducted by the CSO is the Household Budget Survey (HBS). The main purpose of this survey is to update the weighting basis of the Consumer Price Index. Surveys were held every seven years between 1966 and 1994 when the frequency was increased to a survey every five years. This is a very extensive income and expenditure survey that requires interviewers to work very closely with the respondent households. A target sample of 7,500 –8,000 household has been achieved for each survey.

12. Aside from the above-mentioned surveys, the only other regular national household survey is the Living in Ireland Survey undertaken by the Economic and Social Research Institute with the support of Eurostat. This is an annual sample of around 3,000 households with a focus on income and living conditions and is part of the European Community Household Panel project.

13. Table 2 summarises the impact of the CSO surveys on the overall statistical budget of the Office and also identifies the proportion of total costs taken up by field costs. The table shows that, on an annualised basis over a five-year period, the census accounts for a fifth of the total expenditure on statistics by the CSO. The continuous QNHS takes up a further 12 per cent and the HBS approximately 2 per cent. In summary, therefore, the three main household surveys/census account for over a third of the total. While not measured, it can be safely assumed that their statistical output as a proportion of total statistical output is significantly less.

14. Adoption of proposals to introduce new surveys such as the EU Survey on Income and Living Conditions would add a further 5 to10 per cent to annual expenditure.

Table 2: Impact of expenditure on household surveys on overall statistical budget

Survey /Census	**Share of annualised expenditure on statistics * %**	**Field costs as a percentage of total survey costs %**
Census of Population	20	60
QNHS	12	70
HBS	2	75

* Calculation based on expenditure over a five-year period.

15. The table also shows the dominant influence of field costs on survey expenditure. In all cases field costs accounted for between 60 and 80 per cent of overall survey costs.

Obtaining value for money

16. Labour costs are, of course, the major component of field expenditure. In an era of high unemployment it was possible to engage good quality interviewers and enumerators at relatively low cost. However, this situation has changed dramatically in Ireland over the past five years with the

unemployment rate falling from 12% in 1996 to less than 4% in 2001. The result has been a significant tightening in the available labour supply. As a result it has been necessary to increase, for example, the fees payable to Census Enumerators by around a third in real terms between 1996 and 2001. Obtaining value for money must therefore be accorded a higher priority in all aspects of the surveys. In Ireland attention has been focused on sample design, development of omnibus surveys, linkages between surveys and the introduction of cheaper collection methods. Some examples are given below:

- For the introduction of the QNHS the CSO engaged an international statistical consultant to advise on a cost-effective sample design for the new survey. The result of this was that it was possible to reduce the target quarterly sample from 47,000 to 39,000 without any negative impact on the precision of the results. The solution involved the development of a less clustered sampling scheme with interviewing organised so as to minimise any additional survey costs arising from increased travel etc. It is beyond the scope of this short paper to elaborate on the technical issues involved. The essential point, however, is that savings of this magnitude are possible through careful analysis of the options.

- The principle purpose of the QNHS is to collect data on the labour force (it does in fact meet Ireland's requirement under EU legislation to conduct a continuous labour force survey). However from the outset, provision has been made to cover other social issues on a periodic basis through the development of special self-contained modules. This aspect of the survey is still at a developmental stage but to-date special modules have been undertaken to meet national and EU requirements. The latter are those specified as part of the EU labour force survey programme while the national modules in the first three years of the survey's existence have covered topics as diverse as: *Housing; Crime and Victimisation; Recycling and Energy Conservation; Travel to Work; and Home Computing*. There has been a very positive public response to this approach and the list of topics proposed by users is lengthening all the time! There is of course a constraint in that the topic and type of question must be compatible with the core labour force survey. In addition, the response burden on the interviewee must also be taken into account. Nevertheless, our experience has been that a considerable amount of valuable additional information can be collected on a range of topics at relatively low marginal cost.

- For a country with a significant agricultural population, it is essential for the purposes of the periodic Household Budget Survey to have an accurate measurement of income from farming for farm households in the sample. In earlier surveys the CSO conducted a special farm accountancy survey in parallel with the HBS in farm households. This was a huge overhead not only in terms of costs but also in terms of technical complexity. For the more recent surveys, however, the CSO reached an agreement with Teagasc (the national Farm Advisory Service) to canvass farm households in their annual farm accountancy survey to participate in the HBS. A good response has been achieved and this co-operation has not only reduced costs but has also meant that this important income component is measured more comprehensively, and indeed more accurately, than under the old approach. With the agreement of the participating households, Teagasc receives the individual HBS data on household expenditure and other household income from the CSO and is thus in a position to add considerable value to its standard analyses of the farming sector.

- As already stated, the major cost associated with most household surveys are the field costs involved in interviewers conducting face to face interviews. Processing costs are of course also significant not only in money terms but also in terms of time. To mitigate the overall costs, CSO has introduced CAPI for the QNHS and is also committed to introducing CATI, where appropriate, when technical resources permit. There are, of course, substantial overhead costs associated with the introduction of these new technologies which makes them more appropriate for frequent surveys such as the QNHS but less attractive for periodic and complex surveys such as the HBS.

Future directions

17. Compared with most developed countries the range and frequency of household surveys in Ireland is still relatively poor despite major improvements in recent years. The cost factor has undoubtedly been the main reason for this. The demand for household survey based information is, however, increasing all the time and it is clear that more surveys will be required. The conduct of a Time Use Survey (for which a successful pilot survey has been undertaken), the establishment of an annual cross sectional survey on income and living conditions, and more frequent surveys on household expenditure are likely to be the priority areas. To meet these new demands the CSO will seek to build on the field-force structure now in place for the continuous QNHS and to maximise the synergies between the various surveys.

18. The five-year frequency for the Census of Population is likely to continue for the foreseeable future. To reduce field costs the option of mail-back will have to be thoroughly examined. A small pilot test in 1999 was encouraging but much more needs to be done to assess the potential impact on the quality and completeness of the census results.

III. Business Surveys

Censuses rather than samples

19. The focus in business surveys is usually on providing key statistical information (e.g. employment, output, earnings, and stocks) for a wide range of economic sectors or products. In small countries the number of firms in a particular sector is often so small that samples will not give sufficiently reliable results. Accordingly, all firms in the sector above a relatively low threshold must be surveyed. In Ireland, for example, there are approximately 5,000 industrial firms employing three or more persons. Of these approximately 2,000 have 20 or more employees and each of these firms must be included in all regular industrial inquiries conducted by the CSO. Since the absolute numbers involved are relatively small, the burden on the statistical budget is not a major issue. Instead the focus has to be on the burden on respondents, particularly those in the small to medium size category (i.e. employing between 20 and 50 persons). Table 3 summarises the increase in the burden of response entailed for a firm going over the 20 person employed threshold.

20. The table shows that for the five regular inquiries undertaken by the CSO a firm with 20 employees receives 22 forms each year whereas its counterpart with 19 engaged receives only two. This is a sizeable disparity, which occurs at a relatively low threshold, and it results in the CSO being regularly criticised for the form filling burden it places on small firms. While the Office can

point to the fact that amount of information collected in the more frequent surveys is not great, and is usually relatively straightforward, it nevertheless recognises that for small firms with poor (or non-existent) administrative systems the surveys can have a very high nuisance value.

Table 3: No. of forms issued annually by CSO to industrial firms by size of firm and type of inquiry

Inquiry	3-19 employed	20+ employed
Monthly Production	-	12
Quarterly Earnings	-	4
Quarterly Stocks and Assets	-	4
Annual Prodcom	1	1
Census of Production	1	1

21. The CSO's stated policy is to keep the response burden as low as possible. Initiatives to date have included: willingness to accept best estimates; preparedness to use a firm's management accounts to extract statistical data; and visits by field personnel to help the respondents. The Office is also embarked on a major re-structuring of its Business Division so that a consolidated approach is taken in seeking data from a company. The objective will be to ensure that specific information is requested once and only once from an enterprise. In addition, an integrated business register is being developed which should make it possible to make greater use of sampling in targeted cases.

Confidentiality concerns

22. Irish industry is dominated by a relatively small number of large companies as can be seen from Table 4.

Table 4: Manufacturing enterprises classified by size – 1998

Size group (persons engaged)	No. of enterprises	Share of turnover(%)	Share of employment (%)
Under 20	2,358	3.2	8.7
20-49	1,055	6.0	13.5
50-249	828	26.7	35.4
250-499	118	31.0	16.6
500 and over	65	33.1	25.8
Total	4,424	100	100

23. The table shows that the number of small enterprises with less than 20 persons engaged represents over half the total but these firms account for only three per cent of turnover and less than nine per cent of employment. At the other extreme, less than 200 firms employed 250 or more persons but they accounted for almost two thirds of turnover and over 40 per cent of employment. The dominance of the latter firms poses significant confidentiality disclosure problems for the CSO. The following are some examples of the impact on published results of invoking standard disclosure protection procedures:

- In the Census of Industrial Production the main aggregate variables can only be published for 69% of the 227 populated 4-digit NACE cells;
- In the case of cross sectional analyses for 3-digit NACE by size class over two thirds of the potential cells have to be suppressed;
- In the PRODCOM survey sales were identified under 2,100 headings but over 1,500 value and almost 1,400 volume figures cannot be shown.

24. Under the Statistics Act, 1993 access may be given to anonymised microdata for the purposes of research. However, the structure and size of business enterprises in Ireland makes it virtually impossible to grant such access while guaranteeing the confidentiality of individual returns. In practice the CSO will grant access to business data in exceptional circumstances only and imposes additional restrictions on the researcher. The latter include the requirement that access must take place on CSO premises and that the researcher agree to be bound by the same legal restrictions regarding confidentiality as CSO staff.

IV. Concluding remarks

25. This brief review of the characteristics of sample surveys in Ireland clearly demonstrates the relatively greater burden placed on smaller countries in implementing a national statistical system. This burden covers not only the budgetary costs but also the burden on respondents and the restrictions on the use of the data collected due to confidentiality considerations. In short, the cost benefit assessment of new statistical projects must of necessity be more searching in smaller countries. In the past, when countries tended to develop their national statistical systems on an individual basis, such assessments tended to have a much sharper focus – often, admittedly, with an outcome which was to the detriment of the statistical user! Nowadays, with the emphasis on developing internationally comparable statistical systems, smaller countries face a greater challenge in responding to the “one size fits all” paradigm which is the norm and which tends to be based on the needs, resources and experiences of larger countries. Ireland, as a small country, has struggled in particular to meet this challenge since joining the EU in 1973 and it is an experience that many more similar sized countries will face as the Union is enlarged. This is an issue that is not limited to the EU as it applies also in other international statistical contexts and is thus worthy of a wider debate.

- - - - -