

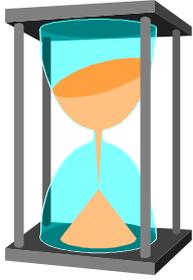


UPDATE - News from the LS User Group

Issue no. 11
August 1995

| Contents: | Page |
|-----------------------------------------------------------------------------------------|------|
| 1 Diary | |
| - LS Workshops | 2 |
| - One-day conference: "The enlargement of a complex dataset: some first analyses" | 2 |
| 2 The new LS Undertaking Form | 2 |
| 3 Technical issues: | |
| - 1991 Travel-to-work Areas (TTWAs), Brian Dodgeon | 3 |
| - Information on the LS Support Programme on the Internet and | 4 |
| - Statistical software developments on the OPCS computer, Simon Gleave | 4 |
| 4 New LS related publications | 5 |
| 5 Facilitating comparability in international comparative research, Rosemary Creeser | 6 |

This newsletter is designed to provide rapid dissemination of information on the OPCS Longitudinal Study (LS) and a forum for the exchange of users' views and comments. It is produced by the LS User Support Programme at the Social Statistics Research Unit (SSRU) at City University. All comments and contributions for the newsletter should be sent to Rosemary Creeser, LS Support Programme, SSRU, City University, Northampton Square, London EC1V 0AR tel. 0171 477 8487 Email: rc@uk.ac.city.ssrु. Contributions on IBM -formatted floppy disk are always welcome and should be sent, clearly documented (file name, wordprocessing package and version used) along with a hard copy of the text.



1 Diary

This section highlights forthcoming events of interest to LS Users.

If you are arranging an event and wish to publicise it in future issues of *Update* you should send details to Dina Maher, the LS Administrative Secretary at SSRU.

LS Workshops

SSRU hold regular 2-day workshops to provide detailed information on the study and to enable researchers to gain practical hands-on experience of accessing the data. The workshops also provide an ideal opportunity to meet members of the LS Support Team and to discuss the suitability of the LS for investigating specific research questions. The next LS Workshop will be held on 12th/13th December. A programme and booking form are enclosed with this newsletter.

As part of the hands-on element of the workshop participants are able to specify a statistical analysis of their choice using a small sub-set of variables extracted using the new LS FLEXTRACT facility, described in issue 8 of *Update*, and a test data-set based on 1% of the LS data-base. As usual, the number of places is limited to ensure that participants get sufficient individual attention and hands-on experience. A fee of £20 will be charged to cover documentation and administrative costs. Researchers who plan to start work shortly on projects using LS data, are advised to contact Dina Maher immediately on 0171 477 8486 to reserve a place.

One-day conference on "The enlargement of a complex dataset: some first analyses"

On Wednesday October 18th, SSRU and the ESRC ALCD (Analysis of Large and Complex Datasets) programme will be holding a one-day conference entitled "The enlargement of a complex dataset: some first analyses". The event, which will be held at the Senate Suite, City University will incorporate presentations on technical, analytic and substantive issues relating to the LS. A non-returnable fee of £7 will be charged to cover administration and refreshments. More details may be found on the enclosed programme and booking form. Individuals wishing to reserve a place for this event should return the completed booking form with their remittance to Dina Maher, LS Support Programme, SSRU, City University, Northampton Square, London EC1V 0HB.

2 The new LS Undertaking Form

The *LS Undertaking Form*, which new users are required to complete before they are allowed to proceed with projects based on LS data, has recently been re-designed. In addition to details of the background and aims of the study and the proposed dissemination of research results, the new form now records detailed information on the population of interest, the proposed methods of analysis and the individual variables which will be required to carry out the research.

The main reason for expanding the scope of the form is to ensure that the LS Research Board, the committee that evaluates all potential new LS projects, has sufficient information to decide whether a project should be approved. The board is particularly concerned that research projects maintain the confidentiality of the data which are included in the LS and that there are no significant conflicts with existing, ongoing work. Both of these issues are now explicitly addressed by the new form.

An additional reason for re-designing the *LS Undertaking Form* is to facilitate the work of the LS Support Team at SSRU - for example, to assist the scheduling of work and the extracting of sub-sets of

data for specific research projects. We also hope that the new form will prove helpful to researchers planning to use the LS for their work.

Copies of the new *LS Undertaking Form* are available in either paper or machine-readable form from Dina Maher (tel: 0171 477 8486, Email: dm@ssru.city.ac.uk).

It is important to note that the new form is designed to be completed in consultation - either with a member of the LS Support Team at SSRU or the LS Unit at OPCS.

3 Technical issues

3.1 1991 Travel-To-Work Areas, Brian Dodgeon

Travel-to-work areas (TTWAs) are available in the LS for both the 1971 and 1981 Censuses and are an attempt to approximate self-contained geographical labour markets.

Based on ward-level information, the original variables were constructed by comparing home and workplace locations to identify commuting patterns. Cluster analysis, a multi-variate statistical technique, was used to aggregate wards into areas which were "substantially self-contained".

For 1981 TTWAs had to have a working population of 3,500 or more. The total of those who lived and worked in the same area had to be at least 75% of both the number of people who worked in the area and the total number of workers who lived in the area. The one exception to this rule was that areas with a working population in excess of 20,000 could be still be allowable as a TTWA at a "self-containment" rate of 70%. Further details may be found in the *Employment Gazette Occasional Supplement, No.3*, September 1984.

The 1991 TTWAs have been developed by researchers at Newcastle University and were submitted for approval last year to the Department of Employment (DE). Unfortunately, since then no further news has been received from the Department. It is likely that the new areas may be contentious because they are used in connection with the publication of regional unemployment data and are substantially larger than the 1981 areas.

Following the July Cabinet reshuffle which abolished the DE, responsibility for most statistical functions has passed to the Central Statistical Office (CSO). OPCS hope to hear from the CSO soon in order that they can proceed with incorporating the look-up table for the 1991 TTWAs into the LS Model 204 database.

3.2 Information on the LS Support Programme on the Internet, Simon Gleave

The LS Support Programme at SSRU have recently established a link with the World Wide Web (WWW). The site currently contains summary information on the OPCS Longitudinal Study and the LS Support Programme. At some point in the future, we also hope to be able to use the link to make most of the free LS documentation available for users to "pull down" onto their own personal computers using file transfer software such as FTP.

While the LS pages are still under development they should be complete by the end of August. To link up with the LS site simply point your web browser at the following address:

<http://ssru.city.ac.uk/Ls/lshomepage.html>

Readers in academic establishments who do not currently have access to the Internet should be able to obtain free copies of the most widely used browser software, *Mosaic*, *Netsape* etc from their computing support staff.

3.3 Statistical software developments on the OPCS computer, Simon Gleave

In recent years the LS Support Team at SSRU has supported users wishing to use both SPSS and SAS on the OPCS computer. In order to broaden the types of analysis on offer OPCS have now added two further statistical software packages - GLIM and ML3.

The first of these packages, GLIM, will allow members of the LS Support Team to analyse categorical data from the LS at the individual level. Until now this type of analysis could only be carried out by requesting a large machine-readable table which was then analysed at the user's own site. The rules governing the release of this type of data from the LS have meant that users have previously been limited to a maximum of 7 variables and 250,000 cells. It is obvious that this may be particularly limiting when analysing longitudinal data based on three time points. The availability of GLIM will help to minimise these limitations.

ML3, the second package, has been developed by researchers at the Institute of Education for analysing hierarchical or "multi-level" data. Taking an example from the LS (female labour force participation), the software may be used to model both the characteristics of individuals and those of the areas in which they live and work. As outlined above for GLIM, the LS Support Team will be able to use ML3 to analyse individual-level LS data, subject to the usual confidentiality constraints and the competing demands of other LS work. Users who wish to find out more about GLIM and ML3 should note the *LS User Guide* series includes two relevant titles, which are referenced below. There will also be a presentation on ML3 at the LS seminar "The enlargement of a complex dataset: some first analyses" on October 18th.

If you are interested in analysing LS data using GLIM or ML3, you should contact Simon Gleave (Tel: 0171 477 8000 X 4129 or EMAIL: sg@ssru.city.ac.uk).

References:

Ward C and Dale A (1991) *Modelling categorical data from the OPCS Longitudinal Study using GLIM*, *LS User Guide no. 5*, London: LS Support Programme, SSRU, City University

Creeser R (1994) *Analysis issues in the LS*, *LS User Guide no. 12*, London: LS Support Programme, SSRU, City University

4 New LS related publications

The Longitudinal Study, 1971 to 1991: History, organisation and quality of data, Lin Hatterseley and Rosemary Creeser, ISBN 0 11 691637 0, £27.30

On July 27th HMSO published a detailed 374-page companion volume to the OPCS Longitudinal Study. This volume, which is a collaborative venture between colleagues at OPCS and SSRU, covers the development of the OPCS Longitudinal Study from 1971 up until the 1991 Census. Separate chapters are devoted to the history of the study, the scope for analysis (including international comparative work), its organisation and an overview of the data sources used. It also examines the quality of the linkage of key socio-demographic variables from the 1971, 1981 and 1991 Censuses, together with that of vital event data spanning this 20-year period. The breadth of material covered and the large number of illustrative examples from previous research make the volume equally suitable for experienced LS researchers and

those with little or no previous knowledge of the study, working in public health, academic, local or central government .

LS User Guide no. 13: Using the OPCS mainframe computer with the LS, Brian Dodgeon

A new LS guide will be available in August for those users with special authorisation to work directly on the LS at OPCS St. Catherine's House, London. The guide follows a similar step-by-step format to that of *LS User Guide no. 4* which explained how to use the previous ICL computing environment.

Separate sections are devoted to the MVS Operating System, how to create, maintain and archive files, alternative methods of editing, input/output, printing, uploading and downloading files, virus-checking, extracting data from the Model 204 Database using the FLEXTRACT utility, and how to use the SPSS and SAS statistical packages.

LS Working Paper 74: Patterns and determinants of birthweight in consecutive live births: results from the OPCS Longitudinal Study, 1980-88, Susan Macran and David Leon

This working paper summarises findings on the patterns and determinants of birthweight among consecutive live singleton births which occurred between 1980-88 to a sample of 10,000 women born in 1950 or later and enumerated at the 1971 Census. The work reported in the paper was funded as part of an ESRC research project *Health inequalities among women and their families*, undertaken at the Centre for Population Studies, London School of Hygiene and Tropical Medicine. A strong tendency to repeat birthweight was observed for women whose first birth weighed less than 2500 grammes (the WHO definition of a 'low' birthweight), compared to those whose first birth weighed 4000 grammes or more. Another significant finding is the "regression to the mean" effect. Both the tendency to repeat birthweight and the regression to the mean effect were observed in sets of first and second births of the same sex and where the social class of the father recorded at birth registration was the same. While these effects are consistent with previous international research, it is the first time that they have been studied using a large representative sample of the population of England and Wales.

Copies of LS User Guides and LS Working Papers are available free of charge from Dina Maher (tel: 071 477 8486 or EMAIL: dm@ssru.city.ac.uk.)

LSE Working Paper no. 1 (May 1995) : Labour mobility and employment stability: Is there a continuing sex differential in labour market behaviour? Catherine Hakim

In this LSE Working Paper Catherine Hakim analyses data on labour mobility and employment stability from the LS and other OPCS datasets.

Analyses of certain topics, such as occupational change, can only be carried out using a sub-group of people who are in the labour market at two consecutive censuses. The paper compares this sub-group with the total workforce in employment at any single time-point. Using LS census data for 1970, 1971 and 1981, it shows that there are large sex differentials in movement in and out of the workforce over the periods 1970-71 and 1971-81. As a result there are large differences between men and women in the sub-group in continuous employment - eg those in work at two consecutive censuses, at the start and end of a decade.

Women in continuous employment are an unrepresentative minority. Over a decade, one-third of the workforce will be new entrants and another third will leave it. Women dominate both groups, even though they are a minority of the whole labour force. During a period of ten years, only two-thirds of the workforce remains continuously employed - representing three-quarters of men and fifty per cent of women in work at the start or end of the decade. Though women constitute just under half of the total workforce, they

compose less than a third of the continuously employed. Furthermore, those in continuous full-time work are exceptionally rare. On the other hand, continuously employed men are a representative subgroup of all men in employment.

Although the LS offers scope for longitudinal analyses of women's employment, it is important to remember that women in continuous employment are an unrepresentative minority. To ensure that a representative sample is covered, the conceptual framework has to be recast to take account of women in intermittent employment and full-time housewives who leave the workforce permanently, early in adult life, either at marriage or childbirth. The paper presents estimates of the distribution of the three employment profiles (women in continuous employment, women in intermittent employment, full-time housewives) with a consideration of the implications for research and the potential for bias.

Copies of this Working Paper may be obtained free of charge from Amor Vieira in the LSE Sociology Department (tel: 0171 955 7309, fax: 0171 955 7405).

5 Facilitating comparability in international comparative research, Rosemary Creeser

Previous issues of *Update* have carried articles on some of the longitudinal studies maintained by countries such as Finland, France and Israel (see Creeser, 1993; Roualt, 1994 and Peritz 1994). The fact that there are many similarities between these studies and the LS, both in their coverage and the types of data which are recorded, means that there is considerable scope for international comparative research. For example, LS mortality data has been used for an international study comparing the size of socio-economic mortality differences among men in the 1970s (see Kunst and Mackenbach, 1992) and more recently to study occupational differences in mortality and hospitalisation in three EC countries (England and Wales, Denmark and Italy). The first of these studies is currently being extended using more recent data to ascertain to what extent international variation in health inequalities changed in the 1980s. In addition, linked census data from the LS has made an important contribution to work comparing migration patterns. An interesting example is the research carried out by Cribier and Kych (1993) to compare retirement migration from Paris and London.

Although there are many similarities - for example between the EDP (the French equivalent to the LS) and the OPCS Longitudinal Study covering England and Wales - there are also many "obstacles" to international comparison. The most obvious of these is that studies are primarily established for national analyses. One of the consequences of this is that the data which are recorded will tend to reflect the classifications used in the country of origin, which may not necessarily facilitate international comparisons.

This article explores several of the issues that researchers planning to use the LS for international comparative research may need to address. It is based upon the comparatively small number of research projects which have drawn upon LS data for this type of work. As new projects come online we hope to include further related contributions in forthcoming issues of the newsletter.

Figure 1: A comparison of three definitions of the family

According to the three definitions a family is defined as:

1. "British" definition used at the 1971 and 1981 Censuses

- i. A married couple with or without their never married child(ren) **or**
- ii. A father or mother together with his or her never married child(ren) **or**
- iii. Grandparent(s) with grandchild(ren) if there are no apparent parents of the grandchild(ren) usually resident in the household.

2. "French" definition

- i. A married or unmarried couple with or without their never married child(ren) aged under 25 years **or**
- ii. A father or mother together with his or her never married children aged under 25 years.

3. "Census" definition

- i. A married or unmarried couple with or without their never married child(ren) **or**
- ii. A father or mother together with his or her never married child(ren).

Comparability of standard definitions

Differences in the definition of what constitutes key concepts such as the "family", "household", "an elderly person" etc may present one of the main obstacles to international comparative research. For example, preparatory work for a project comparing the "English and French family in historical perspective", based upon census data from the LS and the French censuses of 1975 and 1982, highlighted the need to harmonise key concepts such as the definition of a family. In international terms, the definitions used by both of the countries' census offices contain several idiosyncratic features (see figure 1). For example, the "British" definition includes families composed of grandparent(s) and grandchild(ren) while the French impose an arbitrary age limit of 25 years on children. Furthermore, the "British" definition used for the 1971 and 1981 Censuses also excluded unmarried couples. This last feature both fails to reflect developments within British society and recommendations made by the United Nations for the most recent round of European censuses (see Priest, 1987).

In order to facilitate comparisons with data from the French censuses the research team derived a set of variables which correspond to the French definition of a family. An additional set, analogous to the "census" definition shown in figure 2, were also defined. These variables which were originally derived in the previous OPCS computing environment will enable researchers to make comparisons with other countries and with the results of the 1991 Census. Further information on these variables may be found in Penhale (1990a), while some of the findings of the research project are summarised in Penhale (1990b), Wall (1990a and 1990b).

Defining comparable spatial areas

Researchers interested in using the LS and equivalent longitudinal datasets for making comparisons at a sub-national level also need to consider how comparable spatial areas are to be defined. The following paragraphs outline some of the issues that Cribier and Kych (1993) had to consider in order to define the exact geographic areas they were interested in for their work comparing retirement migration from Paris and London.

Previous research has highlighted the problems of comparing Paris and London (see Robert, 1989; Shepherd, 1989), due to the nature of the two cities, the boundaries used and the recording practices of the census offices in each of the respective countries. In a summary of their research methodology, Cribier and Kych note that compared to Greater Paris, which is fairly easy to describe, arriving at a definition of Greater London is complicated by the fact that, at no distance is there a distinct threshold in the population density which could serve as a boundary. According to whether the administrative region of Ile-de-France or the French census ZPIU definition ('conurbation Zone de Peuplement Industriel et Urbain') is used, the population of Greater Paris ranges from 9 to 10 million. Beyond this point population density drops suddenly and the city is replaced by a rural, agricultural region which is less densely populated than the rest of France.

Furthermore, the definitions of what constitutes a "greater London" which are in use encompass: Greater London, the London Metropolitan Area, the South East region and even the greater south east. (The last of this includes almost a third of England.) Depending on the choice of definition the population of a "greater London" ranges from 7 to 20 million inhabitants. In order to compare spatial areas with populations of a similar size, Cribier and Kych decided to compare the areas of Greater London (6.7 million inhabitants) and the London Metropolitan Area (12 million) with the Paris Conurbation (8.7 million).

They also had to consider whether to distinguish the inner/outer areas of the cities. Their final decision not to distinguish the areas any further was partly influenced by the confidentiality requirements of the data sources. (For example, in the LS ward-level data cannot be accessed because of the problem of small numbers.) However, for the purpose of their analysis they decided to create an "inner Paris" by combining the *communes* less than 10-11 km from the centre of the Ville de Paris (see figure 2), thereby creating an area comparable with the inner London boroughs. (Figure 2 is not available in the .PDF version of *Update*.)

Finally, to make comparisons at a sub-national level it may be necessary to identify other comparable spatial areas such as large cities. This component of Cribier and Kych's work was greatly simplified by the fact that they were able to use the urban units defined by INSEE (the French census office) recorded in the EDP. However, as there was nothing similar to this in the LS, the team chose to use the metropolitan counties and a sub-set of the more densely populated districts outside the London metropolitan area.

The (trans)portability of UK measures

Over the last decade there has been considerable discussion of the inadequacy of the Registrar General's measure of social class for investigating differences in mortality - both for women and for men. As a result of this, much of the LS research into health inequalities has also drawn upon "asset measures" which reflect responses to the housing tenure and car ownership questions recorded by the census (for example, see Goldblatt, 1990).

One of the issues that researchers wishing to using the LS to study international variations in health inequality will need to address is the relevance of these types of measure. It is obvious that household tenure will not differentiate as well in other European countries such as Germany and the Netherlands where rented housing is very much more common and where a large proportion of the population defer buying a home until middle-age. One solution explored by Kunst and Machenbach (1994) in their study of premature mortality in nine industrialised countries has been to use educational level as a proxy for social and economic well-being. While the use of education as a socio-economic indicator has been corroborated by other researchers such as Valkonen (1993) it is not completely trouble-free.

For example, one of the problems highlighted by Kunst and Mackenbach is that the educational data for three of the nine countries included in their study (Sweden, Finland, England and Wales) assigned more than two-thirds of the population to a single category (the lowest one). Unfortunately, this is the way the England and Wales (LS) educational data are coded. The census question on which they are based is solely concerned with identifying those individuals with degrees, professional or vocational qualifications ('qualified manpower'), rather than recording the highest educational qualification, which in many cases would have been obtained at school. Another problem mentioned was that, in the same three countries, data on educational level was unavailable for more than 1% of the population. However, for England and Wales the research team concluded that as the mortality level of the "unknown" group was close to the national average, this had a negligible effect on the measurement of the association between mortality and educational level.

A final note

A distinctive feature of the LS is that due to its "safe setting" data are made available in their raw (uncoded) form and are not grouped or top-coded. One of the advantages of this is that variables may be recoded more easily to produce those which facilitate international comparisons, such as the EGP measure of social classification (see Erikson, Goldthorpe and Portocarero, 1983).

References:

Creaser R (1993) "The Finnish Longitudinal Study and its uses", *Update - News from the LS User Group*, **Issue no. 6 (October 1993)**, London: LS Support Programme, SSRU, City University

Cribier F and Kych A (1993) "A comparison of retirement migration from Paris and London", Special LS issue of *Environment and Planning A*, **Volume 25:10 (October 1993)**, pp 1399-1420

Erikson R, Goldthorpe J H and Portocarero L (1983) "Intergenerational mobility and the convergence thesis", *British Journal of Sociology*, **34**, pp 303-340

Goldblatt P O (1990) *1971-1981 Longitudinal Study: Mortality and social organisation, LS Series no. 6*, London: HMSO

Kunst A E and Mackenbach J P (1992) *An international comparison of socio-economic inequalities in mortality*, Rotterdam: Dept of Public Health and Social Medicine, Erasmus University

Kunst A E and Mackenbach J P (1994) "The size of mortality differences associated with educational level in nine industrial countries", *American Journal of Public Health*, **84:6**, pp 932-937

Penhale B (1990a) *Households, families and fertility, LS User Guide no. 1*, London: LS Support Programme, SSRU, City University

Penhale B (1990b) *Living arrangements of young adults in England and France, LS Working Paper no. 68*, London: LS Support Programme, SSRU, City University

Peritz E (1994) "Longitudinal Studies based on record linkage in Israel", *Update - News from the LS User Group*, **Issue no. 9 (November 1994)**, London: LS Support Programme, SSRU, City University

Priest G E (1987) "Considerations for the definitions and classifications of households and families and related variables for the 1990 round of censuses", *Statistical Journal of the United Nations Economic Commission for Europe*, **Vol. 4:3**, pp 271-303

Robert J (1989) "Comparer Paris et Londres: un exercice délicat", *Cahiers du CREPIF*, **26** (March), pp 17-19

Rouault D (1994) "The Echantillon Démographique Permanent: A French equivalent to the Longitudinal Study", *Update - News from the LS User Group*, **Issue no. 7 (January 1994)**, London: LS Support Programme, SSRU, City University

Shepherd J (1989) "Paris et Londres: deux dynamiques spatiales", *Cahiers du CREPIF*, **26** (March), pp 30-43

Valkonen T (1993) "Problems in the measurement and international comparisons of socio-economic differences in mortality", *Social Science and Medicine*, **36**, pp 409-418

Wall R (1990a) *English and French households in historical perspective. LS Working Paper no. 67*, London: LS Support Programme, SSRU, City University

Wall R (1990b) *Residence patterns of the elderly in England and France, LS Working Paper no. 68*, London: LS Support Programme, SSRU, City University

We will be delighted to receive your views and comments on any of the articles in this newsletter.