

# Bulletin

Institute for  
Employment  
Research

## LABOUR MARKET INFORMATION MANAGEMENT

### Introduction

This *Bulletin* differs from the usual format, in that it does not focus upon a specific research topic, but describes a set of related activities carried out within a programme of research at the IER, the Labour Market Information Management Programme (LMIM). Activities which fall within this programme are many and various, but they all have a common purpose - to provide better labour market information to interested parties in a timely and efficient manner.

The phase 'labour market information' is an all-embracing term for data which act as indicators of the processes through which individuals search for, obtain and leave paid employment; by which employers recruit, train, promote or make workers redundant; and by which people create and develop businesses using their own labour services. Well-known indicators include such familiar statistics as the monthly count of claimant unemployment and the index of average earnings. Of less general interest, but equally important, are the many and detailed tabulations on such topics as employment, hours worked and labour costs available in periodical publications such as the *Employment Gazette* and obtained from a variety of survey or administrative data sources.

Users of labour market information rapidly discover that the specific information needed to satisfy a particular requirement is often not available. This may arise because of a need for detailed geographical coverage, through interest in a narrowly-defined occupational group or industrial sector, or it may result from some basic 'inadequacy' in the data source itself. As examples of the

latter problem, coverage of low-paid workers in a particular survey may be less than complete, the data classification systems used in the information source may differ from the user's requirements and varying definitions of employment, unemployment, etc. may lead to difficulties with the interpretation of the information available. The aim of providing 'better' labour market information involves identifying these gaps between the production of labour market information and its usage and in seeking to rectify these problems.

The LMIM programme of research at the IER spans a wide range of topics, from the identification of new statistical measures for monitoring labour markets, the introduction of new data classification systems to meet the demands of users and their harmonisation across and between different producers of labour market information, the trialing and testing of new methods for data translation, editing and statistical processing, to schemes for improving access to large, complex datasets or reviewing the methods and procedures used by official statisticians to collect labour market information. Headed by Peter Elias, the programme is funded mainly by the sponsors of projects which fall within it. However, some projects are funded directly by the IER given that the outputs generated may be marketed to cover the development costs.

As new labour market information management techniques evolve, the techniques and the information they generate may be incorporated into other IER projects. This arrangement not only provides a live 'test-bed' facility for new developments, but also ensures that all sponsors of IER research projects can be assured of access to the most appropriate labour market information for their needs.



### The LMIM Programme

LMIM activities can be grouped under the following headings:

- developing new labour market indicators;
- reviewing existing labour market information requirements and the procedures through which information is made available;
- developing improved data collection methods;
- improving/harmonising data definition and classification systems;
- developing improved data access facilities.

The remainder of this *Bulletin* describes recent or current research activities under each of these headings.

#### Developing new labour market indicators

A key labour market indicator is the monthly count of claimant unemployment. While there is a significant interest in the overall level of claimant unemployment as a measure of involuntary joblessness, many users are more interested in the direction of change from one month to the next. The rise or fall in the level of unemployment is often used as an indicator of progress into or out of recession. There are problems with the use of the unemployment count for this purpose in that it measures a labour market outcome, not a labour market development. At a particular point in time it indicates the outcome of the process through which people attempt to enter the labour market, their attempted movement between labour market states or the outcome of firm closures, 'restructuring' and recession. For this reason, it lags significantly behind the upturns and downturns in the economic cycle. Unemployment can continue to rise even though a recovery in the economy is well underway. Correspondingly, unemployment may still be falling as the economy moves into recession.

To determine the current state of the labour market cycle a more appropriate indicator derives from the existence of job vacancies. These can be regarded as indicative of employers' recruitment intentions and their desired labour force. However, problems arise in operationalising such an indicator. To date, the monthly count of job vacancies notified to the public employment service has been the only consistent information available on a regular basis which relates specifically to labour recruitment. Such notifications constitute only a fraction of all job vacancies and are biased

specifically towards low-skilled manual jobs and clerical occupations. Although some analysts regard vacancies notified to Jobcentres as good proxy measures for changes in the demand for labour, the low level of such vacancies generally and their low skill content makes such information less useful as a consistent and timely indicator of change in the labour market.

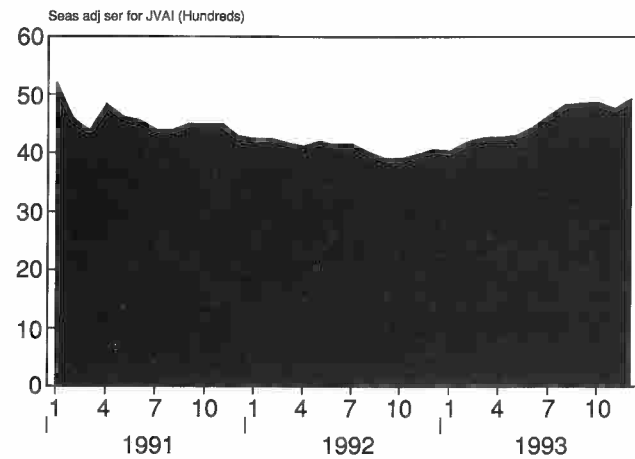
Over the last two years, the IER has been assisting the Employment Department with the development of a new job vacancy indicator based upon newspaper and journal advertising. Such measures have been used with some success in the USA and Canada, helping to pinpoint labour market upturns and downturns and providing a better indicator of the state of the labour market. The advantage of using such information is that it covers a wider spectrum of skills than is the case with Jobcentre vacancies. The main disadvantage is simply the cost of collating and counting advertised job vacancies. Which newspapers and journals should be used? How frequently should counting take place within each period? How should the information from different newspapers and journals be combined to create an index? And how can an historical series be created so that, when it is first introduced, some evidence exists to demonstrate its usefulness?

Work conducted by the IER during 1993 and 1994 has sought to answer these questions. The development has moved from the definitional phase to consider the operational aspects of production of the index. Important considerations are the different costs of collecting information by different methods and the reliability of various counting procedures. Training and Enterprise Councils in England and Wales and Local Enterprise Companies in Scotland helped to identify the most appropriate local newspapers. The IER, together with a firm specialising in media monitoring, has compiled and analysed a vast number of job vacancy advertisements in various national and regional newspaper and specialist journals. From the enormous volume of information these procedures generated it was possible to identify key sources for the index. Many of these sources, particularly the regional newspapers, were able to provide historical information from their management records relating to the numbers of job vacancy advertisements placed in their publications on a weekly or monthly basis.

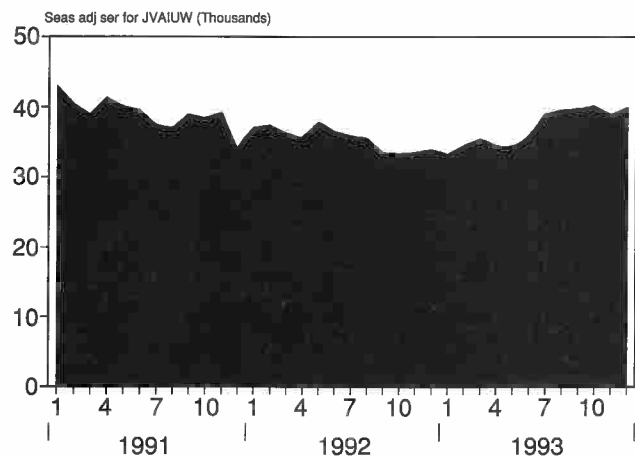
These data series were seasonally adjusted and combined in various ways to determine the best indicator of changes in the volume of job vacancy advertising. Figure 1 shows a trial series, covering the period 1991-94. There is clear evidence here of the recovery from recession during 1992, a process which appears to be continuing, judging from the latest information supplied by newspapers.

Figure 1

### Job Vacancy Advertising Index, 1991-93 (seasonally adjusted series, weighted)



### Job Vacancy Advertising Index, 1991-93 (seasonally adjusted series, unweighted)



The Employment Department has now taken the project out of the research phase and plans to make the index operational during 1995. After further testing, and once the Department is assured of the validity and reliability of the techniques used to create and maintain the index, publication of the series will commence. The programme of development work to produce the index will be described in detail in a forthcoming Employment Department Research Series publication.

### Reviewing labour market information requirements

A good example of work in this area is the review of labour information collected and processed by the agricultural departments of the United Kingdom.

Most of the information on employees in the agricultural sector of the economy is collected not by the Employment Department, but through the agricultural departments. This division arises for historical reasons, and continues because it is the responsibility of the agricultural departments to hold and maintain lists of all agricultural 'holdings' - defined simply as places where agricultural activities occur.

It is, therefore, the responsibility of the agricultural departments to produce information on farm labour both for national accounting purposes and for the purpose of reviewing farm incomes. This they manage successfully through an annual census of agricultural holdings and via various surveys of farm business activity, earnings and hours worked. Complicating this situation is a requirement to prepare different estimates of farm labour, expressed as Annual Work Units, for the European Union. This requirement is fulfilled by conducting a separate survey across 20 per cent of all agricultural holdings once every two or three years.

These requirements clearly place a heavy burden on form-fillers in the agricultural sector. A project being conducted by the IER, in collaboration with the agricultural consultants Laurence Gould Partnership Ltd, is seeking form-fillers' views on this matter and is examining the variety of methods and procedures which could be adopted to make the whole process less burdensome.

### Developing improved data collection methods

The advent of low-cost high-powered personal computing has changed the nature of data collection processes, particularly those which stem from personal interviews. Increasingly, interviewers use computer-assisted personal interviewing (CAPI) techniques to conduct interviews based upon highly structured questionnaires. Interview questions are prompted on the screen of a lap-top computer. Interviewers ask the questions and type the responses into the computer. The *Labour Force Survey* introduced CAPI techniques in 1992. Other well-known surveys which are likely to benefit from this technology are the *Family Expenditure Survey* and the *General Household Survey*.

Despite the obvious benefits of using lap-top computers for personal interviewing (eg. the exact routing it provides to questions which are only asked in response to other

questions and the scope it allows for immediate validation of data), many data collection methods still rely upon postal questionnaires. Once returned, the information contained on such forms usually has to be entered into a computer. Additionally, some of the responses to questions need to be coded and classified to large and complex coding frames. It is in these areas (data entry and coding procedures) that the LMIM programme has been addressing over the last two years.

#### *CADE (Computer-assisted Data Entry)*

In collaboration with the Department of Computing Services at the University of Warwick, staff at the IER have assisted in trialing and developing a technique for computer-assisted data entry. CADE is a technique specifically designed to assist in the creation of data files from survey questionnaires, facilitating the entry of information from multiple choice questions, tick boxes, text, numbers or scale positions marked on the paper. Using a digitising tablet and pen connected to a personal computer, the system requires the user to define the questionnaire, to mark active areas on the paper (areas which may contain data), and to define the nature and allowable scope of each data item. Once defined, data entry proceeds rapidly as the operator points to data on each questionnaire as it is placed in turn on the digitising tablet. The screen mirrors the data entry process, allowing for visual checking with the questionnaire. The program prompts the operator to enter numeric data or text from the keyboard and checks such data for validity as it is entered. On completion, data are immediately available for processing via one of the major statistical packages.

The technique has been tested thoroughly over the past five years and is now used for data entry connected with all survey work carried out by the IER. While it was originally envisaged to be used for internal data management, the successful adaptation of the software for a wide variety of 'paper-to-computer' data entry uses has prompted IER and Computing Services to consider its commercial development.

#### *CASOC (Computer-assisted Standard Occupational Coding)*

CASOC was developed as a result of the work undertaken by the IER to develop the Standard Occupational Classification. Prompted by the success of the Australian Bureau of Statistics in using microcomputer systems for the coding of text to large and complex coding frames, a project commenced in 1989 to package the Standard Occupational Classification (SOC) and its 25,000 word coding index as an integrated coding tool, facilitating the coding of textual descriptions of occupations to the SOC.

User requirements quickly escalated to change what was originally envisaged as a small-scale programming project into a major software development. Included among these requirements was a specified need for fully automated coding, a facility to code text to the Social Classifications (eg Socio Economic Groups and Social Class based on occupation) and to earlier classifications, to facilitate default coding in situations where low quality descriptions were available and to provide managerial control over the coding operations. All of this was eventually realised in a software package which required no hard disk storage and could operate within the conventional memory limitations of the earliest PCs. A major independent evaluation of the software was conducted by the Centre for Survey Methods. These tests showed that CASOC compared very favourably with professional occupation coding using manual index look-up techniques. Most surprising was the evaluation of fully-automated coding, which gave results which were generally as good as, and sometimes better than, expert coding by manual methods.

Marketed by HMSO, the software has been successfully adopted by leading market research agencies, by government departments involved in data collection, Training and Enterprise Councils and academic researchers. Plans are now being formulated to bring the software up-to-date in terms of integrating it within a Windows environment and including the new index to the SOC planned for 1996.

#### *CASIC (Computer-assisted Standard Industrial Coding)*

CASIC has been developed as Windows compatible software which performs essentially the same function as CASOC, but codes economic activity descriptions to the new (1992) and the 1980 Standard Industrial Classifications. Trialing of the software is currently taking place within the IER. External trials will commence shortly. The final product will be introduced during 1996.

#### **Improving and harmonising data definition and classification**

One of the principal ways in which we record change in labour markets through time or measure differences between groups is by making comparisons between indicators. Yet the necessary comparability between indicators which lies at the foundation of this process is often seriously lacking. For true comparability, measurement must involve similar concepts and definitions and great care must be taken with data processing to ensure that comparability is preserved. Frequently the users of labour market information fail to heed the warnings of data

procedures about the non-standard nature of the data they produce, leading to incorrect inferences about the process of change or the nature of differences between groups in the labour market.

Occupational classification is a prime example of an area where comparability is difficult to achieve and where misuse of information can arise through a lack of understanding about the nature of the information. Occupational classification arises from the application of large and complex coding frames to the responses to open-ended questions. The raw material for occupational information arises from questions such as 'What is your job title?' or 'What is your job?' sometimes with a request for a brief description of the main tasks or duties. Respondents may be the job-holder, or may be answering on behalf of the job-holder (eg an employer may complete this information on behalf of an employee, a census form-filler may reply on behalf of others in the household). These job descriptions are then compared with lists of job titles (coding index) to obtain a classification category. Data which have been classified may be aggregated to a number of broad categories.

Classifications may differ in terms of the coding index used and the level of detail may vary according to the final use which is required of the data. Worse still, the classifications may differ in some fundamental respect at a detailed level, yet apparently obscure these differences at the aggregate level.

Prior to 1990, occupational information in Great Britain derived from official surveys and sources lacked comparability. Even at the most aggregate level of the classifications which were used it was often difficult to compare occupational structures and to draw meaningful conclusions about differences in occupational structure. The *New Earnings Survey* employed a version of classification which had been developed in 1972. The *Labour Force Survey* utilised a version of the classification used in the 1981 *Census of Population* whereas the *Family Expenditure Survey* had evolved its own classification of occupations. The Employment Service used a variety of different classifications for analysis of the occupational structure of unemployment or job vacancy statistics. Different classifications were also employed to describe the occupations of persons involved in accidents at work.

Out of this disorder, a project which was originally commissioned by the (then) Manpower Services Commission and undertaken by an IER team of researchers, developed the Standard Occupational Classification (SOC). The purpose of the SOC was to facilitate the comparison of occupational information obtained from different sources.

While this would not eliminate the need for care in making such comparisons, it would at least remove or reduce the complications associated with differences in the underlying classification process.

Four years after its publication in 1990, introduction of the Standard Occupational Classification into all major official sources of occupational information is virtually complete. One major application remains, the survey of the first destinations of leavers from higher education institutions. This is now the subject of a major project conducted within the LMIM programme.

#### *SOC (First Destination Supplement)*

SOC (FDS) is the name given to a version of the Standard Occupational Classification which has been specially prepared for the coding of occupational information describing the jobs which leavers from higher education institutions are recorded as holding approximately six months after completing their studies. Carried out by higher education careers services, the survey is the major annual enquiry into the destinations of students who have completed their studies. It provides important information on the short term outcomes of the higher education process.

Until recently it was felt that the 1990 SOC did not provide sufficient detail in those occupational areas in which higher education leavers were likely to obtain jobs. A project funded by the education departments of the UK, the higher education funding councils and the Higher Education Statistics Agency, with assistance from the Association of Graduate Careers Advisory Services, requires the IER to develop a version of the SOC which addresses this concern, whilst retaining comparability with the previous coding structure and the SOC itself. A draft version of this classification has been prepared and plans for its implementation during 1995 are being made.

#### *ISCO 88 (COM)*

Parallel to these national developments the countries of the European Union (EU) have agreed to develop a common standard for the interchange of occupational statistics. With funding from Eurostat (the statistical office of the European Union), the IER has developed a version of the 1988 International Standard Classification of Occupations [known as ISCO 88 (COM)] onto which all countries can map their national classifications. In the same way that SOC becomes the common language of occupations across different statistical applications in the UK, ISCO 88 (COM) is the common yardstick defining occupations in a European-wide context.

Although it has taken four years to reach its current stage of development, ISCO 88 (COM) has not yet been finalised for all countries. In a paper presented at a recent EU Labour Force Survey meeting in Budapest, Elias and Birch show in detail where problems remain and for which countries these difficulties exist.

### Improving access to and use of large and complex datasets

Quite often the labour market information which can best address a particular issue lies tucked away in some large and complex survey dataset. Examples of such datasets are the *British Household Panel Survey*, an annual investigation into the changing circumstances of household members in a representative group of approximately 5,000 households across Great Britain; the *National Child Development Study*, an investigation into the development of approximately 12,000 people all born in the same week in 1958; the *New Earnings Survey Panel Dataset*, a sample of approximately 1 per cent of employees in the population of working age, covering nearly half a million persons and providing annual observations on each employee's earnings and hours worked from 1975-93; and the *Labour Force Survey*, a quarterly enquiry covering approximately 60,000 households each quarter. Although an enormous amount of information is contained within such datasets, it is seldom easy for an analyst to extract the exact data required. This situation arises because each survey gives rise to its own data design, and because of the complexities of selecting and manipulating particular cases within these datasets for statistical analysis. Faced with the complexities of secondary analysis of large datasets, labour market analysts usually turn to published tables or to related research based upon the data, hoping to find the evidence required from such sources and usually compromising between the available information and the appropriate requirement.

The problem, then, is not so much the availability of labour market information, but its accessibility. One approach to the improvement of access to data allows for the interrogation and distribution of data from a national database. In Britain we are fortunate to have NOMIS (National On-line Manpower Information System) for this purpose, but NOMIS is essentially targeted at users who wish to define information for specific geographical areas. For users who

wish to examine data for specific groups, or who wish to generate information about the labour market experience of individuals, they must access the datasets which provide such information. A large part of the LMIM is devoted to the development of software which eases the process of data analysis. One such package is described below.

### ACCNCDS5

The National Child Development Survey is a large and continuing study of all persons born in one week in March, 1958 and who are now living in the United Kingdom. At regular intervals, information about these individuals has been collected from their parents, teachers, doctors and from respondents themselves. Last contacted in 1991, respondents (and their parents) gave an historical account of their employment, housing, fertility and partnership from the age of 16 years. These data are invaluable for two reasons. First, they give a representative picture for the 'average' 33 year old. As such, the labour market information they contain can be used to contextualise similar information from other sources which may be based upon a specifically defined group of interest. Second, the work history information allows us to examine the interrelatedness of events by studying their timing. The difficulty with such analyses is the sheer complexity of the basic data. Organised as lists of event codes and dates, it is a major task to interrogate such data even with modern database management software.

ACCNCDS5 is a software package which offers an elegant and simple solution to this problem. The package operates in two parts. The first part establishes a dialogue with the user via question and answer. Users are prompted to choose the information they are interested in and to specify the nature of the analysis they wish to perform on these data. These instructions are fed to a second program which interrogates the NCDS data and creates the required data for analysis.

ACCNCDS5 was developed with funding from the Economic and Social Research Council and the Anglo-German Foundation. It is available free of charge from NCDS Support Group, Social Statistics Research Unit, City University London. Accompanying documentation gives some examples of the use of this software in collaboration with major statistical packages.

For further details of the LMIM Programme, or for information about the availability of software and documents detailed in this *Bulletin*, contact Peter Elias, Tel: 01203 523286 (or by fax 01203 524241) Email: P.Elias@warwick.ac.uk.