THE VIEWS EXPRESSED IN THIS PAPER ARE THOSE OF THE AUTHOR. PLEASE DO NOT QUOTE WITHOUT PERMISSION FROM THE AUTHOR.

EXAMINATION OF AVAILABLE DATA ON INDIVIDUALS RELATING TO EMPLOYMENT AND SKILLS ISSUES IN THE UK

Assessing the Evidence Base on Skills and Employment in the UK

Paper prepared for a UKCES LMI Workshop on 22nd January 2009

Anne E. Green

IER, University of Warwick, Coventry CV4 7AL

Email: Anne.Green@warwick.ac.uk Tel: 024 765 24113

December 2008

(slightly revised January 2009)

Contents

			Page
1.		Introduction	1
2.		What information and intelligence is required and desired?	2
3.		Generic issues	4
4.		Assessment of selected LMI sources on individuals	10
	4.1	The Labour Force Survey (LFS) and Annual Population Survey (APS)	10
	4.2	The Census of Population (CoP)	13
	4.3	The Annual Survey of Hours and Earnings (ASHE)	16
	4.4	Skills at Work Surveys	17
	4.5	The Workplace Employment Relations Survey (WERS)	18
	4.6	The British Household Panel Survey (BHPS) and Understanding Society	20
5.		Overview – information gaps and priorities	23

1. Introduction

The *purpose* of this paper is to:

- examine whether available data on **individuals** specifically relating to employment and skills issues in the UK is fit for purpose; as a source of consistent and comparable LMI to inform policy deliberations;
- identify any key gaps in the data currently available;
- indicate how current data might be improved.

Scope and structure: Note that this paper is **selective**, rather than *comprehensive*. *It does not aim or claim to cover all possible data sources.*¹ Rather it aims to provide:

- a review of key features of employment and skills on which information and intelligence is required and desired (section 2);
- an overview of *generic issues* relating to the assessment of data sources providing data on individuals (section 3);
- an assessment of selected key data sources (section 4):
 - the Labour Force Survey (LFS) and Annual Population Survey (APS) as the single most important source of information on economic position, employment and other important features of the labour market in the UK; and
 - the Census of Population (CoP) providing the greatest coverage of individuals of any census/survey and providing greatest potential for geographical disaggregation of any of the data sources reviewed.

Foremost attention is placed on these two sources as they are of particular importance in generating estimates of employment. Other sources reviewed in less detail include:

- ➤ the Annual Survey of Hours and Earnings (ASHE) as a key source of information on hours worked and earnings and an example of a source of data on individuals provided by employers rather than individuals themselves;
- Work Skills in Britain providing information specifically on skills;
- the Workplace Employment Relations Survey (WERS) a source where individuals provide information alongside their managers and workplace representatives on various aspects of the nature and practice of employment relations at the workplace between managers and workers including training, organisation of work and employee attitudes to work;
- ➤ the British Household Panel Survey (BHPS) and Understanding Society which provide insights into experiences of a panel of individuals and households over time.
- a **summary of some of the** *key gaps* in available data sources, the potential of other data sources to fill them and associated priorities (section 5).

There are other useful sources on individuals – including various data sources with an EU focus. However, several of these sources provide information on specific topics and the UK sample size is relatively small, so precluding geographical disaggregation.

2. What information and intelligence is required and desired?

Potentially there is a long 'shopping list' of LMI required on employment and skills at individual level. At a minimum and a very broad level this includes trends over time in:

- the economic position of individuals numbers in employment (as an employee or selfemployed), unemployed or economically inactive;
- the *composition* of employment by sector and occupation;
- the skills levels of the population (for both those in employment and those not in employment); and
- the *utilisation of skills* of the population.

Thereafter, there is an interest in:

- other features of employment including earnings, hours worked, employment arrangement (i.e. temporary, fixed-term, permanent position, agency worker, etc), training undertaken both formal and informal, work experience, etc;
- *job mobility* what skills are needed to move between occupations and sectors, and geographically (within the UK and elsewhere);
- how and when skills are acquired through formal training (either on the job or, onthe-job training);
- **skills gaps** skills that individuals need to undertake their jobs better, or to move into a particular job; and issues such as
- *job search* how individuals search for work and how the methods used correspond with recruitment methods used by employers.

There are requirements for *disaggregation* on all of these dimensions; with the degree of disaggregation varying according to the specific research question being asked. Key disaggregations include:

- geography the UKCES has a requirement for data on the four countries of the UK (i.e. England, Scotland, Wales and Northern Ireland). Beyond this, in many instances there is an interest in regional disaggregation within England; (here is it is salient to note that in population and employment terms most English regions are at least as large as the other countries of the UK). At a finer level, there are interests in disaggregation for a range of geographies including 'city regions' (usually defined as aggregations of local authority areas), local authorities, policy areas (defined in many different ways), deprived areas, etc. This highlights the need for geo-referenced data and the requirement for access to such data in order to aggregate it to different geographies.
- sub-group while aggregate data provides a benchmark, there are varying needs there
 is a desire to disaggregate LMI on employment and skills for sub-groups of the
 population notably by:
 - sector of particular relevance for UKCES;
 - occupation;
 - age; and
 - gender.

Other dimensions of interest include:

- ethnicity;
- when arrived in the UK;

- nationality;
- > firm size.

Often there is a need for *multiple disaggregation* on several of these dimensions at any one time.

While there is a need for consistent information across the UK, the potential to make *international comparisons* is also required. This means attention has to be paid to harmonisation issues in terms of how data are collected and classifications used.

There is a requirement for *timely* LMI – providing a 'snapshot' of the *current position* (or as close to the current position as is feasible). Thus current position needs to be set in the context of *historical trends* and *likely future trends*.

3. Generic issues

Before considering the strengths and weaknesses of specific data sources, it is appropriate to consider a number of generic issues relating to the use of these data sources and their fitness for purpose. These include:

Who is providing the information? – Typically, information on individuals is collected
via household or individual surveys. This is the case for the CoP, LFS and panel
studies. However, in the case of ASHE information is provided by employers and where
information is available from administrative records information may be provided by the
individual or a mediator.

In the case of household surveys, it is important to note that responses might be provided by one household member on behalf of others – i.e. there may be **proxy** responses. When another person provides information on behalf of a specific individual, there is a danger that the answer provided may be inaccurate/different from that which the individual himself/herself would have provided. Considerable work has been undertaken on proxy responses in the LFS, where proxy responses are accepted from related adults (in the household in question). Key conclusions from this work² includes:

- proxy responses are highest for individuals in the youngest age groups (those aged 16-19 years) and lowest for females and the inactive;
- agreement levels between information given by proxy informants and the same information given by the subjects themselves were above 80% for many key variables:
- key demographic variables (such as age and sex) are no problem for proxy responses;
- the more detailed the information being collected (i.e. detailed occupation or industry), the less reliable the proxy, although this is not so much of a problem at a less disaggregated level;
- lower qualifications are more vulnerable to misreporting than higher level qualifications (e.g. degrees);
- misreporting in relation to training relates mainly to minor levels of training;
- where detailed numerical information is required (e.g. hours worked, income, etc) proxy responses are less satisfactory.

Overall, this suggests that proxy responses are most reliable for relatively simple information, but are less reliable for more detailed information. This needs to be borne in mind if consideration is being given to new surveys.

• Subjective versus objective measurement — A further key issue relating to the question 'who provides the information', is whether measurement (e.g. of skill) is subjective or objective. In the case of the question: 'what skills are needed to do your job', two individuals undertaking the same job may provide two different answers (and those answers may be different from the answer that an employer might give). Most large scale surveys of individuals rely on subjective measurement, since administering

4

Taken from ONS (2008) Labour Force Survey Performance and Quality Monitoring Report July-September 2008.

objective tests is difficult and costly.³ The advantage of objective measurement is that standards are being measured on a consistent basis in relation to a given threshold, while with subjective measurement this consistency in relation to an external standard is lost because of the reliance on self-classification. On the other hand, it should be borne in mind that people act in accordance with their perceptions, and from this perspective a subjective viewpoint is important.

- What is the sampling frame? As noted above, most information on individuals is provided by means of household or individuals surveys. How the sampling frame has implications for the coverage of a survey. Individuals who are not included in the sampling frame because of where they live, how long they have been there or because they have certain individual characteristics. Here it is important to note that sub-groups who are least well covered in sampling frames tend to include:
 - recent movers (i.e. internal movers within the UK and recent movers to the UK);
 - those in communal establishments (as opposed to private residences).

From a sectoral perspective it is important to note that these limitations have implications for coverage of workers who are more likely than average to share these characteristics – e.g. those in agriculture and construction.

- How is the sample drawn? Ideally, a sample would be the population with characteristics of interest. However, money and time constraints mean that in most instances this is not possible. In probability sampling (incorporating simple random, systematic and cluster sampling) all elements have an equal chance of being selected. Non-probability sampling (including convenience, purposive and quota sampling) rests on availability or judgement that individuals selected are representative. For generating estimates, a probability sampling approach is preferable, while non-probability sampling may be used where the population of interest is difficult to identify or contact, or where insights into processes are required.
- Response rates Although the CoP aspires to complete coverage of the population, surveys of households and individuals are voluntary. Hence, it is important to consider overall response rates and also variability in response rates, and implications of these for the quality and robustness of information collected. The CoP fails to achieve complete coverage of the population (for example, in 2001 it is estimated that 3 million persons [6%] of the population was missed) and response rates to other surveys are lower. In general, there has been a trend towards decreasing response rates and this needs to be borne in mind when considering the merits of conducting new surveys. Moreover, non-response is an important issue because it can cause bias in survey results.

For instance, the LFS Performance and Quality Monitoring Report from the LFS for July-September 2008 shows an overall response rate of just fewer than 59%; (note that the response rate is higher in the 'first wave' than in subsequent waves). Of those refusing to take part, the single largest category is 'outright refusals' (accounting for nearly half of all non-respondents). Geographically, response rates were highest in West Yorkshire

Objective measurement is more likely to be used for measurement of basic skills and basic literacy.

(nearly 72% overall) and lowest in Inner London (just under 52%), Outer London (just under 54%) and the West Midlands Metropolitan County (nearly 59%). Similarly, data on attrition indicates that those who fall out of the LFS between the first and fifth waves are over-represented in London, in the younger age groups and amongst employees, whereas attrition rates are lowest amongst the economically inactive.

Likewise earlier work conducted on non-responding households on the LFS using linked data from the CoP showed that:

- households that are difficult to contact tended to be located in London, contain one person, have one adult in employment, have no dependent children, contain people who did not form a family, live in a purpose-built flat or in a converted/shared house, be private renters, have lived at their sample address for less than a year; and have a household reference person born outside the UK, from an ethnic minority group; aged between 16 and 34 years, single, an employee or self-employed;
- households that refused to take part in the LFS were most likely to be located in London or the South East and have no dependent children; with a household reference person who has no academic qualifications or have qualifications other than a degree.

This information on response rates underlines:

- the importance of weighting to adjust for non-response;
- confirms the difficulty of covering certain individuals in certain places in surveys notably young people (especially males), those from outside the UK, those with limited English and recent movers; and
- ➤ the trend towards lower response rates⁴ and the challenges of tackling apathy and mistrust in collecting information from individuals.
- Sample size The size of a sample has implications for the possibility for, and robustness of, subsequent disaggregation especially when there is more than one dimension of interest (e.g. sector and region, sector and age, sector and ethnicity, sector and occupation, etc). Hence it may be necessary to 'trade off' one dimension of interest against other dimensions. The source with the largest sample size is the CoP and there is potential here for use of individual data as well as aggregate data. An alternative approach is to 'pool' data across consecutive periods in order to achieve a larger sample so as to enable greater disaggregation, but obviously this means that the temporal dimension is lost. Note that this approach required consistency in survey methodology, classifications and coding over time.
- Timeliness This is a key issue for UKCES: timely data are needed to measure trends.
 This is particularly important for those aspects of employment and skills that are changing quickly, but may be less of an issue where change occurs more slowly. There

_

For a recent assessment relating to the LFS see: Barnes W., Bright G. and Hewat C. (2008) 'Making sense of Labour Force Survey response rates', *Economic & Labour Market Review 2* (12), 32-41. Public attitudes, media representation, data collection methodology, the weather, sporting events, staff training and political issues are amongst the factors identified as influencing response rates.

tends to be considerable concern about timeliness at times of economic change (e.g. at the time of writing there is considerable interest in short-term changes in employment levels and flows into non-employment and in the sectoral, occupational and geographical dimensions of change), and while this is important, it is also crucial that a medium- and longer-term perspective on employment and skills is maintained too.

Key features of LMI sources with regard to timeliness are:

- frequency of data collection here there are obvious differences between the decennial CoP and continuous surveys; and
- > elapsed time between data collection and publication/availability for analysis.

 With respect to the latter, there may be a trade-off between release of selected data sooner, subject to subsequent revision, or waiting longer for more complete data.

A concern for timely data does not mean that information that is dated is unimportant. For example, the decennial CoP provides a useful benchmark and enables detailed crosstabulations on multiple dimensions.

■ Cross-sectional and longitudinal data — Cross-sectional views are valuable in providing detailed 'snapshots' at particular points in time. Consecutive cross-sectional views can provide information on trends over time. Note that this is different from a longitudinal perspective measuring changes over time for particular individuals. The CoP provides the classic cross-sectional perspective, with a detailed 'snapshot' every ten years. Note, however, that linked Census records (the Longitudinal Study) provides links between Census records (between 1971, 1981, 1991 and 2001 in the case of England and Wales, and shorter periods in Scotland and Northern Ireland) enabling examination of comparative individual changes over the long-term, linked to certain life event data. The LFS has a five-wave structure, enabling tracing of longitudinal trends over five consecutive quarters (e.g. providing useful information on flows between employment, unemployment and economic inactivity).

Longitudinal sources are more expensive to set up and maintain than data sources providing a cross-sectional view. However, the ESRC and others have made considerable investment in longitudinal and panel surveys. In general, the possibilities for more detailed disaggregation are limited – especially when producing estimates, as opposed to using the data for modelling purposes. Considerable investment is needed in 'keeping in touch' with respondents – especially for those groups who are particularly mobile (e.g. young people).

Linked administrative records also provide potential for longitudinal analyses. For example the Work and Pensions Longitudinal Study (WPLS) links spells on different DWP benefits and so is useful for measuring moves in and out of employment. For some population sub-groups there may be potential for greater use of such sources in future, but there are important issues of confidentiality and access to be resolved.

• **Coding**: Classification schemas (for sectors and occupations) inevitably lag behind changes in the real world. While official sources use *standard classifications* (e.g.

Standard Industrial Classification [SIC] and Standard Occupational Classification [SOC]), which in turn enhance international comparability of sources. When classification schema change there are issues of discontinuities in data sets and so it is useful to 'dual code' surveys in order to measure how different categories fit together and to provide 'conversion factors'. In one-off surveys – focusing on a particular sector, area, topic – the temptation to use non-standard coding may be greater, and the advantages of non-standard schema need to be traded off against the ability to make comparisons with other sources.

When analysing data at a sub-regional scale (and also at a regional level in some instances, depending on the functional geographies concerned), there are also issues relating to whether data are coded on a *residence- or a workplace-basis* (or both). Note that most data relating to individuals is coded on a residence basis – although workplace data are provided in addition to residence-based data in the LFS and CoP, and other sources (e.g. ASHE) relate to the employer address. Generally, LMI from employers relates to a workplace (or employer headquarter address) basis.

- Data security, confidentiality and disclosure: Issues of data security, confidentiality and disclosure have a high profile at the current time. In summary, the key issue is that it may be possible to identify individual data. Concerns about these issues have implications for response rates and may mean that information sources that exist cannot be used to their full potential especially because of actual or potential limitations regarding the flexibility of users to make use of the maximum level of disaggregation available on particular variables in order to aggregate data in user-defined ways.
- Measures of skills: There is a considerable literature on the concept of skill and how to
 measure skills. In the largest general-purpose and labour market oriented surveys of
 individuals, the measures of skills available are:
 - Qualification as a measure of achievement of competence. Although data may be collected on a range of qualifications obtained, often only the 'highest qualification' is coded. Although qualifications vary in different parts of the UK, it is possible to code them onto a common structure relatively easily. Foreign qualifications pose much greater difficulties in terms of provision of information from individuals and subsequent translation into a UK-based framework either by the individual concerned or by LMI analysts;
 - Length of education number of years in schooling;
 - Occupation as a measure of how competence is used; (note that skill level is an important element in the Standard Occupational Classification [SOC]).

Of these, qualifications and occupations have been most widely used and are the measures for which the greatest time series of data are available. Qualification is a key measure because they are available for those individuals not in employment as well as for those in employment.

The LFS and some other individual surveys also contain information on:

Training – related to specific jobs.

Other approaches used are:

- Tests which measure individuals' performance to a specific standard. However, these are very expensive and tend to measure only a narrow range (of usually basic) skills.
- ➤ Job requirements as in Skills Surveys of individuals and/or employers. Note that these are subjective, but can cover a wide range of skills.
- *Linking data sources*: There is potential for more work here. Obvious examples include:
 - > linking administrative data and survey data; and
 - adding ecological data to survey data.
- Quantitative and qualitative methods: The emphasis of this selective review is on large-scale quantitative surveys. Qualitative methods are useful in providing insights into processes, motivations and attitudes. Hence, they can provide useful complementary information to large scale quantitative surveys. Biographical approaches can also be valuable here.

4. Assessment of selected LMI sources on individuals

This section provides an overview of some key LMI sources.⁵ Foremost emphasis is placed on the Labour Force Survey and the Census of Population because of their importance in generating estimates of employment.

4.1 The Labour Force Survey (LFS) and Annual Population Survey (APS)

Why the LFS is important. The LFS is a large household survey designed with the primary purpose of providing information on the UK labour market that can then be used to develop, manage, evaluate and report on labour market policies. According to a 'Review of the Labour Force Survey' by the ONS in 1992 the primary purpose of the LFS is "the prompt publication of key aggregate, whole economy indicators, for the integrated assessment of labour market conditions". It is the main source of timely information on economic position, the number of people in employment, the sector and the characteristics of that employment (including sector and occupation). It also detailed information on individuals' qualification levels and on a wide range of other LMI – covering all aspects of work, wanting to work and not working, income from work and benefits. As such, it is the most comprehensive single source of LMI in the UK.

Coverage and responsibility for conducting the LFS: The ONS is responsible for conducting the LFS in Great Britain, while in Northern Ireland the Department of Trade and Investment (DETINI) has a similar responsibility. The ONS publishes full UK LFS results – hence there is complete coverage across the UK. The LFS is carried out under a European Union Directive and uses internationally agreed concepts and definitions. Hence, the LFS is a key source of LMI for international comparative work. It is the source of the internationally comparable ILO unemployment measure.

Differences across the UK: There are relatively minor differences across the UK – these being the result of a small number of different questions (e.g. on religion) in Northern Ireland compared to Great Britain.

Methodology: The LFS is a quarterly sample survey of households living at private addresses. The Great Britain sample is taken from the Postcode Address File, except in that part of Scotland north of the Caledonian Canal where a random sample is drawn from the published telephone directory. In Northern Ireland the sample is taken from a rating and valuation list.

The sample design currently consists of around 60 thousand responding households every quarter. Each quarter's sample is made up of five 'waves', each of approximately 11 thousand private households, and each wave is interviewed in five successive quarters. Generally, the first interview is undertaken face-to-face and subsequent interviews are

As noted above, the review is selective rather than comprehensive.

The ONS is looking to undertake a Communal Establishment Survey to help fill part of the gap.

undertaken by telephone. Answers are self-reported. Note that some information is collected in one specific wave only.

From 1979 the LFS in Great Britain was biennial, from 1984 it was annual and since 1992 the LFS has been undertaken continuously with the results published quarterly.⁷ The quarterly survey in Northern Ireland commenced in winter 1994/5.

The Annual Population Survey (APS): From 1994-5 the annual local area database was developed to provide LFS data for local areas by combining data across quarters. Sample boosts (not included in the LFS) were undertaken to provide a minimum number of economically active people in each local authority area in England and in each unitary authority area in Wales and Scotland. This boosted Annual Local Area Labour Force Survey (ALALFS) provided greater potential for examining local areas or small sub-group populations. In 2004 the APS was introduced, including all of the ALALFS, providing a further sample boost, but with quarterly (as opposed to annual) publication. This boost was subsequently withdrawn in 2006.

Topic coverage: Key topics covered by the LFS include:

- household and respondent characteristics age, sex, marital status, housing / accommodation, nationality, citizenship, ethnicity, religion, length of time at current address, previous whereabouts;8
- employment details
 - o type of work including *industry*, public/private sector, *occupation*;
 - employment characteristics and conditions employed / self-employed, government training scheme, full-time / part-time, permanent / temporary, working from home, travel-to-work, working times and patterns, when started current job and how it was found;
 - o redundancy;
 - o under-employed looking for more work;
- ➢ looking for work ⁹
 - o whether looking for paid work;
 - o reasons for not seeking work;¹⁰
 - type of work sought (employed / self-employed, full-time / part-time);
 - o characteristics of job search;¹¹
- benefit entitlement;
- > previous employment (12 months before)
 - o type of work (industry / occupation);¹²
 - o characteristics (employed / self-employed, full-time / part-time);
- education and training
 - o current qualifications (type, level, number);

In 2006 there was a change to calendar quarters for reporting purposes.

These individual and household variables enable disaggregation of key employment and skills variables on a number of different dimensions.

The LFS is the key source of information on this topic.

¹⁰ The LFS is unique in the detail reported here.

¹¹ The LFS is the key source of information on this topic.

Thus enabling analyses of sectoral and occupational mobility.

- where highest qualification acquired;
- o current study (qualifications studied, type of course, etc);
- apprenticeships;
- o job-related training (in last 13 weeks, last 4 weeks and last week);
- characteristics of job-related training (on / off job, site, fees, reasons for training, pay whilst training, time spent training, qualifications from training);
- o adult learning (taught or non-taught learning undertaken in the last 3 years);
- health long-term health problems / disabilities (effects, characteristics and intensity), past health problems;
- accidents at work
- income gross income and net income (in main job and second job).

Hence, the LFS is a key source of information on economic position, the characteristics of employment, job mobility, labour market flexibility, equal opportunities, earnings, training, qualifications, health and safety at work and union membership.

Accessing data and interpretation: A limited amount of aggregated LFS variables are made available via Nomis (for countries, regions and local authority areas), with associated confidence intervals. Individual records are also available electronically.

As a sample survey, the LFS is subject to sampling variability. This is especially important when considering change over time – i.e. Is a small change real? Or is it a consequence of sampling variability? Clearly, estimates are more accurate for larger samples or groups than for smaller ones. Estimates below 10,000 are not released. This underlines the importance of the large sample size of the LFS.

The smallest geographical area for which LFS data are released is the local / unitary authority district.

Future plans: The LFS is continuing. At the current time a *Review of the LFS Questionnaire* is being undertaken.¹³ This includes questions on the degree to which the variables meet user needs for both coverage and content.¹⁴ There is also a range of 'open' questions relating to use made of the information, whether information is made available sufficiently frequently, relevance of information to needs, whether the information meets requirements, whether requirements could be met by other sources, and whether it would cause a problem if information was no longer available or available less frequently.

The five Government household surveys on which ONS leads (i.e. the LFS [and its boosts], the General Household Survey; the Expenditure and Food Survey, the National Statistics Omnibus Survey and the English Housing Survey) are to be merged into an *Integrated Household Survey (IHS)*. An integrated field force of interviewers will administer a

12

³ See https://www.nomisweb.co.uk/articles/news/files/LFS%20User%20questionnaire.dot

By broad module – i.e. individual demographics, household characteristics, family characteristics, economic activity, main job, second job, seeking / not seeking work, ILO unemployment, benefit entitlement, employment twelve months ago, transition from work to retirement, education and training, health, childcare and income.

See http://www.statistics.gov.uk/articles/nojournal/IHS web update May08.pdf

common modular questionnaire which will include all the topics covered by the separate surveys. A short core module, collecting information on socio-demographic variables, will be administered to all respondents, while different topic modules will be administered to parts of the sample. The sampling frame for the survey will encompass those for all the individual surveys. The IHS is being introduced in a phased manner. Addition of the core IHS questions to the LFS/APS has been delayed until at least 2009.

Assessment: The LFS is the key sources of information on employment and skills in the UK. Key strengths are:

- large sample size;
- the fact that it meets international standards (hence permitting international comparability as well as consistency across the UK);
- it provides quarterly estimates (of employment);
- each respondent is categorised as either in employment, unemployed or economically inactive:
- it provides a range of data about the details of individuals' jobs and whether they are looking for work, etc so giving a picture of how people fit into the labour market
- it provides estimates of earnings and provides a particularly valuable source of data about the low paid.

4.2 The Census of Population (CoP)

Why the CoP is important: The decennial CoP is the most comprehensive source of data on the characteristics of the population. It has been conducted every ten years since 1801, with the exception of 1941 (due to World War II). The uniquely valuable feature of the Census is that it strives to achieve complete coverage of the UK population. In order to provide a comprehensive picture of the population, the Census is adjusted for non-response, using the responses to a Census Coverage Survey (CCS) conducted soon after Census day. Unlike all other data sources on individuals it can yield detailed socio-economic information for small geographical areas.

Coverage and responsibility for conducting the CoP: The CoP covers all parts of the UK. The ONS is responsible for the CoP in England and Wales. In Scotland responsibility rests with the General Registry Office for Scotland (GROS) and in Northern Ireland with the Northern Ireland Statistics and Research Agency (NISRA). In recent years censuses have been conducted simultaneously throughout the UK.

Differences across the UK: There are small differences between questionnaires used in England and Wales (relating to Welsh language), and some further differences in questionnaires used in Scotland and Northern Ireland. Of particular relevance to employment and skills issues are that:

- ➤ in Scotland only in 2001 information was collected about 'place of study' and means of travel to place of study;
- ➤ England was the only part of the UK to include a question in 2001 on professional qualifications.

Methodology: The CoP adopts a *self-completion methodology*. This means that the onus is on relatively simple questions and there is limited scope for extending the amount of information collected. To date, this has involved delivery of questionnaire to households.

Topic coverage: The *main topics* covered by the Census are:

- household accommodation;
- household relations;
- demographics;
- cultural characteristics;
- > state of general health
- > qualifications
- employment / economic activity
- workplace and journey-to-work
- migration

While the amount of information on employment and skills is limited, the key dimensions of economic position, sector and occupation of employment, hours of work and formal qualifications are covered, as is journey-to-work (allowing a link between place of residence and workplace).

The key *weakness* of LMI from the CoP is that the information is *dated* and is not suitable for measuring short-term trends. The most recent data from the CoP is for *April 2001*. Nevertheless, the CoP remains very important for benchmarking purposes and for studying long-term trends. It is also the key source of information on travel-to-work patterns (i.e. the link between where people live and where they are employed).

Accessing CoP data: Most CoP data are published in the form of standard pre-specified tables for a hierarchy of geographical areas from the smallest geographical units (Output Areas) through electoral wards to local authority districts, to Government Office Regions and nations. In addition to the standard tabular outputs from the CoP, a 3 per cent sample of CoP returns (the Sample of Anonymised Records [SAR]) is made available for further analysis in a completely anonymous form. This can be used to produce bespoke tables from the CoP. A more detailed version (the Controlled Access Microdata Sample [CAMS]) can be accessed under conditions of strict confidentiality via the ONS Virtual Microdata Laboratory. Another method of obtaining more detailed areas for bespoke areas is to request commissioned tables from the CoP.

Planning for the 2011 Census is now at an advanced stage. Questions from previous censuses likely to be included (albeit with changes in classifications in some instances) cover the following (those of particular relevance to issues concerning employment and skills are italicised):

- basic demographics: sex, date of birth, marital/civil partnership status;
- student status (including term-time address);

- ➤ migration: country of birth (asked since 1841) and address one year ago (asked since 1961);¹⁶
- ethnic group (asked since 1991);
- Welsh language proficiency (Wales only asked since 1891);
- religion (a voluntary question first included in 2001);
- health status (asked since 2001);
- long-term illness / disability (asked since 1991);
- > occupation (asked since 1831);
- self-employed or employee (asked since 1851);
- ever worked (asked since 2001);
- > economic activity (asked since 1851);¹⁷
- address of workplace (asked since 1901);
- travel to work (asked since 1971);
- qualifications (asked since 1961 although in 2001 the number of qualifications covered was extended to include the full range of levels of qualifications [academic and vocational], whereas in 1991 questions were on the highly qualified);
- industry (asked since 1911);
- > name of employer (asked since 1911);¹⁸
- > carers (asked since 2001).

New topics proposed for 2011 include:

- rational identity (this is designed to complement the ethnic group question);¹⁹
- > date of most recent arrival to the UK for those who were born outside the UK;20
- > English language proficiency;²¹
- ➤ Second addresses.²²

There are proposals for questionnaires to be returned in one of three ways:

- completed and posted back by 'householders';
- completed and submitted online not used formerly for the CoP;
- completed via a telephone interview not used formerly for the CoP.

Looking further ahead: The future of the Census beyond 2011 is unclear. For 2011 a decision was made that the traditional census was the only viable option for collecting the

¹⁶ This provides information on the stock of migrants.

This provides details of employment and economic inactivity, as well as employment.

Asking for the name of employer helps improve the accuracy of industry coding.

¹⁹ It is likely that the precise format of this question will vary between countries of the UK.

Respondents will be asked to exclude holidays or short visits overseas. This question is designed to gather better information on recent to the UK.

The 2007 Census Test included the following categories: 'no ability', 'understand spoken', 'speak', 'read' or 'write' – but this categorisation and the exact format of the question may be changed.

This topic is designed to gather information on more fluid living arrangements which make the concept of 'usual residence' more difficult. The 2007 Census Test included questions on:

Do you stay at another address for part of the week or year? (with categories including 'a student's home address' and 'an address you use when you work away from home', as well as holiday homes, etc): with a request to write in the address and country.

Why do you stay at this other address? (including a category 'I stay there when I work away from home')

How long do you stay at this address?

information that has been included in former censuses. The replacement of the census with an Integrated Population Statistics System bringing together information from multiple sources has been debated.

Assessment: The value of the CoP for individual data on employment and skills rests on the fact that it aims to cover the entire population of the UK (with relatively small differences between the four countries) and enables detailed geographical disaggregation and crosstabulation of multiple dimensions of interest. As such it remains an important source of LMI.

4.3 The Annual Survey of Hours and Earnings (ASHE)

Overview and purpose of ASHE: As the title suggests, the ASHE provides information on the levels, distribution and make-up of earnings and hours for employees within industries, occupations and regions. ASHE was developed to replace the New Earnings Survey in 2004. Changes were made to improve the coverage of employees and derivation of earnings estimates.

Coverage and responsibility for conducting ASHE: Responsibility for ASHE rests with the ONS in Great Britain, while DETINI is responsible for the equivalent survey in Northern Ireland. Hence, coverage is UK-wide. Note that ASHE covers employees only – it does *not* cover the self-employed.²³ Note that there was a 20 per cent reduction in the size of the ASHE sample in 2007 – with 142,000 returns in 2007, down from 175,000 in 2006. Largest sampling reductions were made in industries where earnings were least variable.

Methodology: ASHE is based on a sample of employee jobs taken from HM Revenue & customs PAYE records. Information on earnings and hours is obtained in confidence from employers; ASHE is unusual in providing data on individuals from employers. Hence there are no issues of inconsistency in interpretation between different individuals respondents. Earnings information collected relates to gross pay before tax, National Insurance or other reductions, and excludes payments in kind.

A panel dataset (selected by the last two digits of the National Insurance number) enables earnings for individuals to be linked over time.

Information is collected annually.

Topic and output coverage: ASHE classifies earnings by features such as public / private sector, gender, age, whether on adult rates of pay, full-time / part-time status, size of establishment, occupation, and permanent / temporary employment status. Information is collected on total gross earnings, overtime earnings, shift and other premium payments, normal basic hours and overtime hours.

-

Hence the importance of the LFS in providing information for this group.

Most *published* ASHE analyses relate to full-time employees²⁴ on adult rates whose earnings for the survey pay period were not affected by absence; however, in non-published data a full-time / part-time breakdown is available. At UK, country and regional level, information is disaggregated by occupation and by industry. Disaggregations by sex are available also. At local authority district level, median earnings are available (without industrial and occupational disaggregation). Information is available on both a residence and a workplace basis; (previously having been provided on a workplace basis only).

Assessment: The ASHE is a key source of information on employee earnings. Key strengths of the dataset include coverage and the fact that information is provided on a consistent basis across individuals by their employers. It provides information on a regular and reasonably timely basis, but does not cover all workers.

4.4 Skills at Work Surveys

Overview and purpose of Skills at Work Surveys: The purpose of Skills at Work Surveys is to gather information on the skills used at work via questions directed at workers themselves²⁵ - so complementing the use of qualification and occupational as measures of skills. The 2006 Skills Survey is the latest in a series of such surveys on work skills in Britain dating back twenty years. The 2006 Survey (funded by ESRC and a consortium of government agencies – including from Wales and Scotland) sought to update the 2001 Skills Survey funded by the Department for Education and Skills.

The objectives of the 2006 Survey were to:

- provide an analysis of the level and distribution of skills both broad and generic (including computing) skills – being utilised in British workplaces in 2006;
- > to provide a picture of recent trends in broad and generic skills;
- ➤ to update knowledge on the valuation of skills, and the association of skills usage with other worker rewards and indicators of well-being, and of how skills are related to the evolution of inequality;
- ➤ to provide a description of the work preferences and work motivation of those in employment and how preferences and motivation relate to skill development that people experience in their jobs;
- to provide an analysis of skills levels and distributions within and between regions of the UK.

Methodology: The target population in 2006 was those people in employment aged between 20 and 65 in 2006. A clustered random sampling method was used and respondents were interviewed in their homes using computer-assisted questionnaires. In addition to the achieved core sample of 4,800 respondents from England, Wales and Scotland (south of the Caledonian Canal), additional samples were drawn from the Highlands and Islands, Wales and Northern Ireland (so providing UK coverage) and the East

17

_

Defined as those working more than 30 paid hours per week or those in teaching professions working 25 hours or more per week.

See Felstead A., Gallie D., Green F. and Zhou Y. (2007) Skills at Work, 1986 to 2006.

Midlands.²⁶ The resulting total sample size (including these supplementary samples) was 7.787 cases.

Topic coverage: The 2006 Skills Survey questionnaire gathered material on:

- job characteristics;
- detailed job analysis;
- > computing skills and qualifications;
- work attitudes;
- the work organisation;
- pay;
- > the job five years previously;
- > recent skill changes and future perspectives; and
- basic demographic data.²⁷

Together these topics enable analyses on themes such as:

- qualifications required for entry to jobs (measured subjectively by the individual) and qualifications supplied;²⁸
- > task discretion; and
- characteristics of those looking for opportunities for skills development.

Assessment: The value of the Skills Survey rests on the wealth of information it provides about *self-assessed skills utilisation of those in employment* and how this relates to attitudes, values and outcomes. Hence, a key contribution is to measuring medium-term change in (subjective) skills utilisation and their association with labour market contexts, processes and attitudes, with associated policy implications.

4.5 The Workplace Employment Relations Survey (WERS)

Overview and purpose of WERS: The WERS is a survey mapping the state of employment relations in workplaces. The objectives are:

- to provide a mapping of employment relations practices in workplaces;
- > to monitor changes in those practices over time;
- to inform policy development and permit an informed assessment of the effects of public policy;
- > to enhance understanding of employment relations and the labour market.

Coverage and responsibility for conducting the WERS: The WERS covers Great Britain; (*N.B.* information is *not* available for Northern Ireland). Surveys have been conducted in 1980, 1984, 1990, 1998 and 2004. The next WERS is scheduled for 2010. The 2004

Illustrating the potential for more detailed disaggregation and analysis at the level of an individual region in England. Similarly there would be potential to boost the sample in other regions, in particular sectors, etc.

Thus enabling disaggregation on selected individual characteristics – e.g. men and women.

So enabling insights into (mis)matches between skills required and supplied (i.e. 'overqualification' and 'under-qualification'.

WERS was sponsored by the Department of Trade and Industry, the Advisory Conciliation and Arbitration Service, the ESRC and the Policy Studies Institute.

Methodology: The sampling frame for the WERS is the IDBR. Within scope are all workplaces with five or more employees engaged in all sectors with the exception of primary industries and private households with domestic staff. The 2004 WERS (and its predecessors²⁹) contained both a *cross-section survey*, based on a random sample of establishments in existence in Great Britain in 2004. This comprised a, comprising:

- ➤ a four-page self-completion questionnaire for the main management respondent about the composition of the workforce;
- ➤ a four-page self-completion questionnaire for the financial manager about the financial performance of the establishment;
- ➤ a face-to-face interview with the senior person at the workplace with day-to-day responsibility for industrial relations, employee relations – covering issues such as recruitment and training, consultation and communication, employee representation, payment systems, equal opportunities, work-life balance, flexibility, etc;
- ➤ interviews with employee representatives covering issues such as the structure of representation at the workplace, negotiation and consultation over pay and other matters; relations with management, etc;

and of particular relevance for the current focus on individuals:

➤ an eight-page self-completion questionnaire distributed to a random sample of up to 25 employees in each workplace (with responses from 22,500 employees in total).³⁰

The WERS also has a *panel element*, which in 2004 involved returning to a random selection of over two thousand workplaces that had participated in the 1998 cross-section, with the purpose of investigating change in those workplaces over the six year from 1998 to 2004. (The next WERS is scheduled for 2010.)

Topic coverage:

Topics covered in the *self-completion employee survey* include:

- about the job length of time at current workplace, whether permanent / temporary / fixed-term, hours worked, attitudes towards the job (in relation to hours of work, autonomy, flexibility, etc;
- ➤ about the workplace flexible working arrangements; amount of training during last 12 months, how well the work skills you have match the skills you need in your current job,
- ➤ views on working at the workplace relating to organisational values, manageremployee relations, etc;
- representation at work trade union membership;
- individual characteristics gender, age, ethnic group, dependent children, caring, academic qualifications, professional and vocational qualifications, use of a computer as part of work tasks, job title (for occupational coding), pay.

²⁹ In 1980, 1984, 1990 and 1998.

The fieldwork response rate in 2004 was 60 per cent.

Linkage with other information sources: A particularly valuable feature of the WERS is that each of the datasets may be linked to another by means of a unique workplace identifier. Moreover, through this workplace element the WERS data can be linked other data sources - such as:

- the Annual Business Inquiry, FAME [Financial Analysis Made Easy] and the Annual Survey of Hours and Earnings (ASHE)³¹ (at a Virtual Microdata Laboratory);
- ➤ the EU KLEMS industry-level database providing measures of economic growth, productivity, employment creation, capital formation and technological change at industry level; and
- > via postcode of the workplace a range of spatially-referenced ecological data.

Assessment: The value of the WERS rests on measuring medium-term changes in the workplace. Of particular relevance in relation to data on individuals are questions on qualifications and training in a wider context workplace context (derived from other elements of the WERS or linked data sources), and self-assessment of the (mis)match between work skills and the skills the individual feels he/she needs for the current job.

4.6 The British Household Panel Survey (BHPS) and Understanding Society

Overview and purpose of BHPS: While the LFS, ASHE and WERS have longitudinal elements, and it is possible to link some individual Census records over time, the main emphasis of the information sources reviewed above is on a cross-sectional view. By contrast, a panel survey provides a longitudinal focus. The BHPS is a social survey of households and families, which commenced in 1991, designed to provide information about individuals' lives, experiences and attitudes in a household and family context.

Responsibility for the BHPS: The BHPS has been funded by the ESRC and is based with a team at the Institute for Social and Economic Research (ISER) at the University of Essex.

Methodology: The BHPS is an annual survey consisting of a nationally representative sample of about 5,500 households recruited in 1991, containing a total of approximately 10,000 interviewed individuals. The sample is a stratified clustered design drawn from the Postcode Address File and all residents present at the first wave of the survey were designated as panel members. The same individuals are interviewed each successive year, and if they split from original households to form new households, all adult members of these households are also interviewed. New members joining sample households are eligible for interview and children are interviewed when they reach the age of 16.³² Extension samples of 1,500 households in each of Scotland and Wales were added to the main BHPS sample in 1999 to enable independent analysis in each of these countries. In 2001, a sample of 2,000 households was added in Northern Ireland, (so achieving UK coverage). Hence the total size for the BHPS (including these extension samples) grew to around 10,000 households across the UK.

-

WERS provides contextual information from the employee's workplace that can be used to help understand variations in earnings in the ASHE data.

Since 1994 children aged 11-15 have also completed a short interview.

Topic coverage: In relation to employment and skills the BHPS includes measures of:

- job satisfaction aspects of pay, job security, work itself, hours worked;
- preferences and expectations (e.g. would like better job with the same employer, new job with new employer, start own business);
- reason for leaving job and attraction of new job;
- for non-employed whether would like a job and likelihood of getting a paid job in the next 12 months; and
- > attitudes to gender roles within home and employment.

Looking ahead: The BHPS is to be incorporated into 'Understanding Society'³³ – a longitudinal study based on a household panel design, due to commence in 2009. This study is designed to provide valuable new evidence about the people of the UK, their lives, experiences, behaviours and beliefs. With a sample of 40,000 households across the UK (approximately 100,000 individuals), this is a very large study. The intention is to interview everyone in those households over the age of ten years. *Understanding Society* both replaces and incorporates the BHPS.

Understanding Society is designed to be *multi-purpose and multi-topic*. As such, it is designed to facilitate ethnicity research, biomedical research and qualitative research, and to provide comparability with other surveys. It is also intended to add information from other sources (i.e. geo-coded data and administrative records). Overall, it is designed to enhance the quality of analysis on outcomes, preferences, personal endowments and constraints, the wider social and spatial environment and behaviours,

Topics of particular relevance (or of related interest) to employment and skills include:

- education, human capital and work including education participation and qualifications, training and skill acquisition, activity and employment status, activity history over the previous year, current job characteristics, hours of paid work, employment conditions, self-employed characteristics, second jobs and voluntary work:
- > standard of living measures (income, consumption, material deprivation, expenditure, financial well-being);
- family, social networks and interactions, local contexts, social support, technology and social contacts
- > lifestyle, social, political, religious and other participation, identity and related practices, dimensions of life satisfaction/happiness
- > psychological attributes, cognitive abilities and behaviour
- preferences, beliefs, attitudes and expectations;

but not all information will be collected at every round of interviews.

-

³³ See http://www.understandingsociety.org.uk/

Assessment: The BHPS contains a wide range of variables on employment and other life domains. It enables examination of individuals within a household context and the availability of longitudinal data means that research questions that cannot be addressed using cross-sectional data may be examined. With Understanding Society there will be potential for further longitudinal analysis – especially of social aspect of employment – within a broader context. With a larger sample size it will be possible to derive more precise estimates for smaller sub-groups and at sub-national scale.

5. Overview - information gaps and priorities

LMI sources are dynamic: information currently available may not be available in the future and new information sources are likely to emerge.34

Official data from censuses and surveys and from government administrative sources (have the advantage of a standard methodology, providing a *consistent overview* across sectors / population sub-groups / areas. This enables comparisons to be made between such sectors, sub-groups and areas.

In the light of information shortcomings and/or gaps in official data, there is an understandable desire for 'one-off' or occasional surveys. These have provided useful information on specific issues (and combinations of issues) in the past (and are likely to continue to have potential to do so in future). At regional level such surveys have often relied on a number of partners to come together to provide funding; (and funding issues often mean that such surveys are of a 'one-off' nature). Such surveys have the advantage of addressing specific information and intelligence needs. However, comparisons between sources and over time are challenging because the focus, coverage, sampling selection, survey methodology and questions asked differ may differ from those used for other sources of LMI.

On balance, especially for generating estimates on employment and skills, there are considerable merits in making use of official data - on employment, unemployment, economic activity, occupations, qualifications, etc - and ensuring that existing sources continue to meet existing needs. In this respect the LFS is of foremost importance and the Census of Population is also significant in providing a comprehensive benchmark. Both of these sources provide coverage across the UK.

In considering collecting new information / incorporating new questions in existing surveys, it is important to remember that such changes take time to put into practice. In practical terms, it is often only possible to include further questions in a survey by dropping others – and this has implications for measuring trends over time. Expansion of the amount of information collected has cost implications and may also impact negatively on response Hence, while the needs of users are a factor in modifying questions, the acceptability of questions (to the general public) and of terms used is an important consideration.

Attempts have been made to provide guides to data sources - for example, see Green A.E., Davies R., Elias P., Hasluck C., Owen D. and Wilson R. (2002) Data Catalogue, Report for DTLR project on 'Regional and Sub-Regional Information and Intelligence Sources'. Coventry: IER, University of Warwick. http://www.local.communities.gov.uk/research/datacata2.pdf; and Green A.E. (2002) Regional and Sub-Regional Information / Intelligence: Needs, Uses, Gaps and Report DTLR. IER. Priorities. for Coventry: University http://www.local.communities.gov.uk/research/needs.pdf. The challenge is keeping such sources up-to-date. For official sources various information is provided online (e.g. via the Nomis website and the National Statistics website.)

However, in surveys where there is a multi-wave structure (such as the LFS) it is possible to include some questions at certain times only.

Moreover, with regard to data collection from individuals, it is important to bear in mind the following *trends* (highlighted in the 2011 CoP planning process) *which have heightened challenges in information collection*:

- increasing mobility and migration;
- complexity of family structures and living arrangements (i.e. this makes the concept of 'usual residence' which lies at the heart of sampling frames for many large surveys of individuals more difficult); and
- mistrust individuals tend to be less willing than formerly to comply.

There may be *potential to make greater use of administrative records* providing employment and skills-related information on individuals (and we have made extensive use of such sources in research on migrant workers),³⁶ although sectoral and occupational information is not always collected, or collected in a standardised form, on such records. Likewise, the Department for Work and Pensions Longitudinal Survey (WPLS) is potentially a valuable source for examining moves on and off work-related benefits.

From an economic perspective, the WERS demonstrates the value of *linking surveys of employers*, workplace representatives *and individuals* at a workplace level. WERS also provides a good example of how links may be made between surveys and other information sources.

Considerable recent investment has been put into *longitudinal studies* by the ESRC and over the coming years there is scope to mine these further. Such surveys are likely to be of particular value in providing information on social (as opposed to economic) issues relating to employment and skills. A challenge for longitudinal and tracking studies is to 'keep in touch' and maintain response rates. At the University of Warwick an ambitious study – *Futuretrack*³⁷ - of the relationship between higher education and employment is underway, investigating the impact of educational and community background on the information available to those applying to enter higher education, the choices that they make, and the implications of these choices as they progress through universities and colleges or make alternative choices. Futuretrack involves two cohorts who applied to study in higher education in 2005 and 2006. The surveys will comprise a six year tracking study, a sufficient period to follow most participants into employment or postgraduate training.

Regarding *gaps*, two that have been highlighted in my own work on employment and skills issues are:

A lack of information on the most mobile elements of the population – In some instances such individuals are outwith the scope of sampling frames and in other instances where they are within the scope of data collection for an information source they may be very difficult to contact. Mobile elements of the population include young people (particularly young males), those living in the private rented sector in large cities and migrant workers. There is increasing interest in use of diaries and blogs for recording experiences to derive intelligence on such groups – many of whom will be very

For further information see http://www2.warwick.ac.uk/fac/soc/ier/research/glmf/futuretrack

-

Green A.E., Owen D.W. and Adam D. (2008) *A resource guide on local migration statistics*. Coventry: IER, University of Warwick.

IT literate. In this way it may also be possible to build up emerging trends, experiences and factors influencing decision-making of such individuals.

• **Dealing with foreign qualifications** – There is a gap in terms of collecting information and dealing with foreign qualifications in analysis. This is a particular challenge given the complexity of different qualification systems and also the diversity of the workforce.

While it is appropriate to prioritise sources providing consistent information in the medium-term, the UKCES has also expressed a need to be relevant to current concerns and policies. Hence, at the current time there is potential to for *redundancy studies* to show how and whether individuals' skills are reutilised. Examples include:

- Ongoing work involving the University of Birmingham and the Work Foundation on the demise of MG Rover at Longbridge, West Midlands. This involves a longitudinal study of ex Rover workers, tracing their experiences of training, employment and non-employment after the closure of redundancy.³⁸
- ➤ A study of post-redundancy paths of workers from Harland and Wolff in Belfast, made redundant in 2000.³⁹

In the current climate there is scope for such studies in different sectors, local areas and institutional contexts.

Likewise, a topic of ongoing interest relates to *job search*.⁴⁰ Key issues here include:

- > What job search channels do individuals use?
- ➤ Which are the most productive in terms of securing employment for whom?
- ➤ How far is the Internet superseding more 'conventional' forms of job search (e.g. newspaper advertising)?
- ➤ How important are social networks / 'word of mouth' advertising?
- ➤ Is there a mismatch between job search channels used by individuals and employers?

See Bailey D., Chapain C., Mahdon M. and Fauth R. (2008) *Life after Longbridge: Three Years on. Pathways to re-employment in a restructuring economy*, Work Foundation, University of Birmingham and ESRC. http://www.theworkfoundation.com/Assets/Docs/MG Rover 2008.pdf

See Shuttleworth I., Tyler P. and McKinstry D. (2005) 'Redundancy, readjustment, and employability: what can we learn from the 2000 Harland & Wolff redundancy?' *Environment and Planning A* 37(9) 1651–68.

There is potential here for sector-specific studies.