

**From UKHLS PI (Professor Nick Buck) and team**

## **UK Household Longitudinal Study (UKHLS): overview and background**

### **Summary**

This paper provides an overview of the UKHLS including the background to the study, its key aims, and features of the design which may influence the choice of topics for inclusion in the study.

The UKHLS is a major new household panel study commissioned by the Economic and Social Research Council. It will provide valuable new evidence about the people of the UK, their lives, experiences, behaviours and beliefs, and will enable an unprecedented understanding of diversity within the population. Key features proposed include:

- a total target sample size of 40,000 households/100,000 individuals
- an ethnic minority booster sample of over 3,000 households
- incorporation of the existing British Household Panel Survey (BHPS) sample
- interviews from all household members aged 10 and above
- topic coverage relevant to a wide range of disciplines and policy fields
- linkage with geo-coded and administrative record data
- collection of health indicators and biomarkers
- a platform for the collection of qualitative data
- an Innovation Panel for methodological research.

For further information about the UKHLS, please see <http://www.iser.essex.ac.uk/ukhls/>. Note also the Consultative Conference, Essex, Thursday, January 10th 2008. Further information and registration forms are at <http://www.iser.essex.ac.uk/ukhls/conference/>.

Further information concerning the establishment of the UKHLS may be found on the ESRC website (use this short-cut link: <http://tinyurl.com/2ns4vs>).

### **Background**

The UKHLS is a study which will follow individuals over time, regularly collecting data about each sample member and his or her household. Such panels have provided a major resource for understanding key issues which face societies around the world. They provide unique information on the persistence of such states as child poverty or disability, on factors which influence key life transitions, such as marriage and divorce, and they provide information on the effects of earlier life circumstances on later outcomes. They also support research relevant to the formation and evaluation of policy and enable improved and more reliable analytical techniques which cross-sectional data, based on only a single observation of each individual, cannot support. In the UK, the British Household Panel Survey has been particularly successful, has already been accessed by more than 2000 users and generates more than 150 publications per year; it is heavily used by government departments and by researchers outside the UK. It has an extremely high international reputation.

The UKHLS will be the largest and most ambitious household panel survey in the world. It will be a major advance on the BHPS. It has seven key features that reflect its scientific rationale, and which can be exploited to generate major innovations in scientific research:

- (i) *Sample size*. The target sample of 40,000 households will give a unique opportunity to explore issues for which other longitudinal surveys are too small to support effective research. It will permit analysis of small subgroups, such as teenage parents or disabled people. Examples include analysis at regional and sub-regional levels, allowing examination of the effects of geographical variation in policy (notably differences between the countries of the UK). A large sample size also allows high-resolution analysis of events in time, for example focussing on single-year age cohorts.
- (ii) *Household focus*. Data will be collected on all members of sampled households and their

interactions within the household. This has major advantages for important research areas such as consumption and income, where within-household sharing of resources is important, or demographic change, where the household itself is often the object of study. Compared to individual-based birth cohorts, it will give better and more continuous information on the family and household environment within which development takes place. Observing multiple generations and all siblings allows examination of long-term transmission processes and isolates the effects of commonly shared family background characteristics. UKHLS will also provide opportunities to explore linkages outside the household.

- (iii) *A full age range.* The UKHLS will complement existing age-focused studies sampling elderly people (e.g. ELSA) or young people (e.g. birth cohort studies), and provide a unique look at behaviours and transitions in mid-life. Moreover the large sample size means that all cohorts can be analysed at a common point in time.
- (iv) *Innovative data collection methods.* UKHLS will involve continuous development in data collection methods benefiting from experience from other longitudinal surveys, and the introduction of new technologies. This entails additional modes of interviewing, external record linkage, possible collection of qualitative and visual data, and an Innovation Panel to allow experimentation and methodological development.
- (v) *Broad, interdisciplinary topic coverage.* The UKHLS is a multi-topic survey for the study of a range of life domains. While meeting the needs of ‘traditional’ quantitative social science disciplines such as economics, sociology and social policy, it will also serve other disciplines (both in the social sciences and the biomedical sciences) and make possible a wider set of methodological approaches (e.g. via linked qualitative studies).
- (vi) *Ethnic minority research.* The incorporation of an ethnicity research agenda within the UKHLS recognises the increasing prominence of research into ethnic difference for our understanding of the make-up of British society and issues of diversity and commonality.
- (vii) *Biomedical research.* The UKHLS will support collection of a wider range of biomarkers and health indicators than any previous social-science focused survey. This opens up exciting prospects for advances at the interface between social science and biomedical research. It will provide the opportunity to assess exposure and antecedent factors of health status, understanding disease mechanisms (e.g. gene-environment interaction, gene-to-function links), household and socioeconomic effects and analysis of outcomes using direct assessments or data linkage.

## **Research agenda**

The UKHLS research agenda is extremely broad. It is based on the assumption that it is being developed as a long term study, lasting several decades, and will be a key resource for future progress in social science research. Some examples of the main areas where the UKHLS will open up new opportunities are given below, but the potential research agendas are broader and will grow with the consultation process. It is also important that UKHLS retains flexibility to meet emerging new issues, while maintaining continuity of measurement in areas retaining importance.

### Households, families and individuals

- How household and family composition are changing, and the ways in which their boundaries are becoming more fluid.
- The impact of a range of personal characteristics and expectations including cognitive abilities, personal plans, and expectations of the future of public provision and the welfare state.

### Socio-economic well-being

- The development of well-being over the life-cycle interpreted broadly to include health, lifetime income, consumption and expenditure, savings behaviour, asset accumulation, time use and individuals’ scope to plan and manage risks both in the short term and over the life course.
- Employment issues, including wage determination and job matching, job quality and how these relate to changing forms and valuations of human capital

- Patterns of participation in market and non-market work, and the dynamics of the work-life balance.
- Changing patterns of economic opportunity for individuals and households, and their influence on the prevalence of poverty and social exclusion.
- The implications of education, training, and life-long learning for labour market success, citizenship, consumption, and socio-economic well-being.

#### Diversity and disadvantage

- The causes and long-term consequences of exclusionary states, conditions or events such as disability, mental illness, drug taking, criminality, truancy, teenage pregnancy and unemployment at individual and household level.
- Issues of diversity, cultures, practices and identities, defined in terms of class, ethnicity, religion, gender, sexual orientation, citizenship, national identity, age group, disability status, consumption and lifestyle – all increasingly important for public policy and an understanding of UK society.
- Patterns of international and internal migration.

#### Health issues

- The formation of individual health-related behaviours (diet, tobacco and alcohol consumption, exercise) and their consequences
- Disability and its management by individuals, families, and private and state institutions.

#### Individuals in society

- Environmental impacts of households' energy consumption and travel, and their responses to changing prices and congestion. Also crucial is the change in attitudes to environmental threats as the scientific and political scene develops.
- Changing patterns of political engagement and civic participation, and the extent of 'individualisation' and disengagement from public institutions.

#### Strategic methodological objectives

- Developments in experimental research, achievable through links to longitudinal survey information on subjects' characteristics and family circumstances.
- Understanding of the behaviour underlying survey non-response, measurement error, and other barriers to reliable quantitative research.

### **The study design**

The basic principles of the UKHLS design will be common with those of most other household panel studies, i.e.:

- It will start with a sample of addresses, and where these consist of private households, all members of those households will be defined as sample members.
- At the first wave and each subsequent wave all sample members above a threshold age will be eligible for interview. Data will also be collected about all sample members, irrespective of age.
- In addition other individuals who form households with sample members after wave one will be eligible for interview.
- Thus the UKHLS will be a longitudinal sample of individuals representing the whole UK population, and interviewed within a household context.
- Individuals will be followed as they move and form new households. The following rules mean that the UKHLS will remain representative of the UK population as it changes, subject to weighting and to developing mechanisms for incorporating new immigrants to the UK who are not naturally represented by these following rules.

The UKHLS sample of 40,000 households will include:

- a) An Innovation Panel of 1500 households to enable methodological research, including experiments for developing and assessing mixed modes of data collection. The fieldwork for the Innovation Panel will commence in January 2008.

- b) A new equal probability main panel achieved sample of around 27,000 households. The fieldwork for this sample will commence in January 2009.
- c) A boost ethnic minority sample, to provide 1,000 adult individuals in each of five minority ethnic groups: Caribbean, African, Indian, Pakistani, and Bangladeshi, in addition to ethnic minority respondents in the main sample. The boost sample will include members of other ethnic minorities in the areas covered by the boost sample. The expected total number of completed households will be around 4,200, a total in the region of 7 – 8,000 interviewed adults. The fieldwork for this sample will commence in January 2009.
- d) The British Household Panel Survey (BHPS) sample of approximately 8,400 households. BHPS sample data collection as part of the UKHLS will start with wave 2 in October 2009. Wave 18 of BHPS will be collected at the same time as wave 1 of UKHLS, using the current BHPS questionnaire.

The study is still in process of being commissioned, and some aspects of the design remain to be finally determined. Some of these depend on the scale of the available budget in relation to data collection costs.

### **Data collection**

Data collection is conceived to involve an innovative mixture of

- interview-based field work with ‘traditional’ questionnaire-based collection, plus
- physical measurements of bio-medical data,
- linkages to geo-coded information and administrative records on e.g. health and health care usage, social security benefits and contribution histories, and education.

### **Fieldwork**

- It is expected that the UKHLS will be based on twelve month intervals between interviews for each respondent.
- Fieldwork capacity constraints will mean that interviewing for a single wave will be spread over a 24 month period. In this design, it is expected that second wave fieldwork would overlap with that for the first wave so as to preserve the twelve month interval between individual interviews.
- Fieldwork will be continuous through the period (in contrast to most other panels, including BHPS where fieldwork is concentrated over a core period of around three months). This has implications for reference periods for retrospective questions. The sample will be divided into sub-samples by month or quarter of interview with a separate twelve month reference period for each sub-sample.
- The first wave fieldwork will take place by face to face interview methods. It is likely that the second wave would use a mixed-mode strategy, in which the first approach is by telephone and if an interview cannot be achieved in this way there is a face-to-face follow-up. This will have implications for the design of individual questions.
- The average individual questionnaire length at wave 1 will not exceed 40 minutes, on grounds of respondent burden, though it also appears likely that available budget will mean that the individual interview will be restricted to 30 minutes at both wave 1 and wave 2.
- It is assumed that all sample members aged 16 and over will be eligible for interview. It is hoped also to collect some data from children below this age. It appears likely that it will be possible to collect data by self-completion from children aged 10-15 at wave 1.

### **The UKHLS questionnaire**

The length constraints cited above are likely to be acute given the broad scope of the UKHLS and wide range of demands that are likely to be placed on it. In order for the UKHLS to exploit its large sample size, and to enrich the range of data collected, it needs to move away from a uniform structure in which, in principle, all respondents answer all of the same questions. Hence the UKHLS cannot repeat the great majority of questions at each wave. Instead, interview time will be saved by asking a much higher proportion of questions less frequently, and some questions asked only of

subsets of respondents. Topic can be classified into different groups, i.e. topics for which data is collected from:

- A. The whole sample annually.
- B. The whole sample on a regular basis, but the frequency is less than annual.
- C. The whole sample once only – e.g. initial conditions and retrospective life history data.
- D. The whole sample after key life events: e.g. questions associated with the birth of a child or other key stages in the child's life (e.g. school choice); in adult life, partnership formation and dissolution, or the onset of disability or retirement.
- E. Parts of the overall sample, defined by respondent characteristics, e.g. characteristics such as ethnic group membership, or life stage.
- F. A random sub-sample, in the case of topics for which effective research can be undertaken with fewer than 40,000 households. Given the pressure on questionnaire space, use of this category increases the range of issues that can be addressed. This may be especially relevant for the introduction of new questions in response to new research agendas.

The design process for the UKHLS aims to allocate topics so as to maximise the scientific value of the UKHLS by ensuring that as wide a range of issues as possible can be addressed in sufficient depth to support innovative research.

There are a number of cross-cutting issues which are likely to be relevant for most topics proposed for inclusion in the UKHLS:

- What is the optimal data collection frequency for measures from a research perspective?
- What level of detail is really required?
- For retrospective and flow measures what is the most appropriate reference period?
- To what extent is it necessary to collect information about each individual within the household?
- To what extent can data be reliably collected by one respondent on behalf of all others in the same household?
- How important is continuity of measurement relative to the existing BHPS, and comparability with other UK national surveys?
- To what extent is cross-national comparability an important consideration when choosing a measure?
- To what extent can linkage with administrative and other data sources provide data that can substitute or complement collection of that data within the UKHLS?

The conclusions of the consultation process will be brought together in order to convert this material into draft and final questionnaires by the principal investigators of the UKHLS, working with the Scientific Advisory Committee and the Governing Board.