Longitudinal Data for Policy Analysis

Michael White, Joan Payne and Jane Lakey
Policy Studies Institute
ABOUT THE GSS METHODOLOGY SERIES

The aim of this series is to publish monographs with a substantial methodological interest written by people across the Government Statistical Service. Findings can be included provided that they illustrate methodology, but the series is not for findings per se.

Publication process
Contributors should send their submissions to the series editor at the address below. Before doing so, however, they should clear the work with their line manager. The submissions will then be subjected to peer review, external or internal, or both, as the series editor see fit. Authors can submit their work for peer reviewed journals as well.

The series is aimed at getting out results quickly and easily. It is the intention that authors will prepare the final documents themselves, on their own PCs. The style has therefore been kept simple. Style guidelines specifying the fonts, type size, margin settings, etc. are available from the series editor (at the address below). The series editor should need to do minimal editing.

Dissemination
Members of the GSS can obtain copies free of charge from the National Statistics Library, Room DG/18, tel 0171 533 6266 or GTN 3042 6266 at the address below. Copies are also available for sale to customers outside the GSS. These can be obtained from ONS Direct, Room DI40, Government Buildings, Cardiff Road, Newport NP10 8XG, tel 01633 812078 or fax 01633 812762.

Series editor
John Charlton
Office for National Statistics
Room DG2/08
1 Drummond Gate
London SW1V 2QQ
Tel 020 7533 6239
email john.charlton@ons.gov.uk

Assistant editor
Sanchia Bailey
Office for National Statistics
Room DG2/07
1 Drummond Gate
London SW1V 2QQ
Tel 020 7533 6188
email sanchia.bailey@ons.gov.uk

GSS MISSION STATEMENT

“To provide parliament, government and the wider community with the statistical information, analysis and advice needed to improve decision making, stimulate research and inform debate”
<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive summary</td>
<td>5</td>
</tr>
<tr>
<td>1  Introduction</td>
<td>8</td>
</tr>
<tr>
<td>2  Some distinctions and definitions</td>
<td>9</td>
</tr>
<tr>
<td>3  The potential need for longitudinal data</td>
<td>11</td>
</tr>
<tr>
<td>3.1 Descriptive information</td>
<td>11</td>
</tr>
<tr>
<td>3.2 Explaining long-term social change</td>
<td>12</td>
</tr>
<tr>
<td>3.3 Evaluating policies or programmes</td>
<td>13</td>
</tr>
<tr>
<td>4  Types of longitudinal data, and issues of method</td>
<td>15</td>
</tr>
<tr>
<td>4.1 General purpose longitudinal surveys</td>
<td>15</td>
</tr>
<tr>
<td>4.2 Long-term cohort studies</td>
<td>18</td>
</tr>
<tr>
<td>4.3 Single-purpose longitudinal studies including evaluation studies</td>
<td>20</td>
</tr>
<tr>
<td>4.4 Administrative databases with a longitudinal element</td>
<td>22</td>
</tr>
<tr>
<td>4.5 Combined or linked uses of administrative data</td>
<td>25</td>
</tr>
<tr>
<td>5  Illustrative applications for longitudinal data</td>
<td>27</td>
</tr>
<tr>
<td>5.1 Incomes and low-income households</td>
<td>27</td>
</tr>
<tr>
<td>5.2 Longitudinal data on youth transitions</td>
<td>28</td>
</tr>
<tr>
<td>5.3 Longitudinal data on ethnic minorities</td>
<td>31</td>
</tr>
<tr>
<td>6  Extension and improvement of longitudinal data</td>
<td>35</td>
</tr>
<tr>
<td>6.1 General purpose longitudinal surveys</td>
<td>35</td>
</tr>
<tr>
<td>6.2 Long-term cohorts</td>
<td>36</td>
</tr>
<tr>
<td>6.3 Special-purpose longitudinal studies</td>
<td>36</td>
</tr>
<tr>
<td>6.4 Administrative databases</td>
<td>37</td>
</tr>
<tr>
<td>6.5 Some additional points</td>
<td>37</td>
</tr>
<tr>
<td>7  Summary and conclusions</td>
<td>39</td>
</tr>
<tr>
<td>References</td>
<td>41</td>
</tr>
</tbody>
</table>
This is a blank page
Executive summary

Aim

The aim of this paper is to discuss the requirements for longitudinal data for purposes of government policy, and to suggest steps involved in meeting those requirements.

Scope of review

The paper is largely based on the experience of the authors in creating and using longitudinal data in a variety of policy-related research studies. It does not claim to be a comprehensive review.

Background leading to increased policy requirement for longitudinal data

In the past two or three decades, a greater variation or fluidity has become characteristic of people's working and family lives. In addition, the linked and complex nature of social disadvantage has become more apparent. There is a particular need for government to know how disadvantages are distributed within individuals' lives as well as across individuals at particular times. Recent policy responses have involved greater use of integrative concepts such as social exclusion, employability and joined-up government. These are some of the underlying pressures which create new requirements for data that will track individual change over time.

Distinctive requirements for longitudinal data

The majority of government's requirements for monitoring social change continue to be met by repeated cross-sectional surveys. However, we have identified four types of need for longitudinal data that are distinct from cross-sectional data.

a) Descriptive information relating to life-patterns (such as the persistence or recurrence of unemployment over long periods, or individuals' marital and partnership histories)

b) Descriptive information relating to life-transitions (such as the proportions moving between various employment statuses, or in and out of benefits)

c) Information required to explain and interpret social change or social problems

d) Information required to evaluate public policies or programmes.

The paper argues that the distinction between description and explanation is an important one. Only longitudinal data will provide the requisite descriptive information about the events and stages in individuals' lives and how these are linked together. But longitudinal data do not necessarily make the task of explanation easier. Such data are inherently complex, and rigorous yet transparent interpretation is often elusive.

Longitudinal data sources

Four main types of longitudinal data sources have been distinguished:

a) General purpose longitudinal sample surveys (particularly represented by the British Household Panel Survey, but also by the Quarterly Labour Force Survey)

b) Long-term cohort studies (notable examples are the birth cohorts of 1946, 1958 and 1970)
c) Special-purpose longitudinal sample surveys (including shorter-term cohort studies, 'one-off' longitudinal surveys, and evaluation studies)

d) Administrative databases (such as those for claimants of various benefits, or for National Insurance contributions).

The recognition of administrative databases as a source of longitudinal data is particularly important. In some other countries, including Denmark, Sweden and the USA, administrative data play a major part in social and economic research.

**Methodological issues**

All sources of longitudinal data are confronted with a range of methodological issues. The main issues of this type discussed in the paper are as follows:

a) Topic coverage. Both general-purpose longitudinal surveys and long-term cohort studies can cover a wide range of topics and thus provide flexibility for a variety of *ad hoc* policy inquiries. This raises important issues about the best strategies for topic selection in such surveys. Any administrative database typically covers only a narrow range of information, but can become more useful when linked to other sources, including surveys.

b) Sample size and representativeness. Policy inquiries often focus on sub-groups with particular disadvantages, who constitute a small proportion of the population. General-purpose surveys need large sample sizes to cope with these needs. In addition, nationally representative samples are not necessarily representative of sub-groups. Administrative databases have large sample sizes, but may be unrepresentative of groups such as non-claimants of benefits.

c) Response rate and sample attrition are of crucial importance to those surveys which continue over many years. US longitudinal surveys have achieved higher overall response than in Britain. There is currently growing concern about survey response rates.

There are other methodological issues that are of particular importance in specific applications. For example, 'missing data' is a special issue in studies of income.

**Application areas for longitudinal study**

The review also briefly considered three applications for longitudinal data: income dynamics, youth transitions, and ethnic minority disadvantage. These were intended purely as illustrative examples.

‘Income dynamics’ is an important subject for policy since it provides a relatively new means for studying individual and family deprivation over time. However, there are many technical problems in this area and further methodological work would be of value.

‘Youth transitions’ constitutes a topic where a variety of sources and methods could be developed to provide improved longitudinal information for policy. Better quality data may be as important as more data.

‘Ethnic minorities’ involves a particularly complex requirement, arising from the diversity of the ethnic groups and the varied nature of their disadvantages. There are also special sampling issues arising from waves of migration that produce peaks and troughs in the age distributions. It appears virtually impossible to meet the information requirements from existing sources. The requirement is for special-purpose datasets, especially cohort studies.
Suggestions for development and improvement of longitudinal data

The paper has put forward a range of suggestions for filling gaps and improving the effectiveness of longitudinal data. The following appear the most important of these suggestions (not arranged in any order of priority).

a) To enlarge the sample size of the British Household Panel Survey (or a successor sample) in future years. This would increase the value of the BHPS as a flexible resource for answering policy questions.

b) To mount an initiative to increase awareness of general purpose surveys and other existing longitudinal data sources among potential users. This might be coupled with assistance in mapping their requirement into existing datasets.

c) To foster the use of the Quarterly Labour Force Survey as a general purpose longitudinal data source. Specifically, to include questioning about wages and non-wage income at all five waves of the survey, instead of just at the initial and final waves as at present.

d) To conduct a further review of issues concerning response rate and sample attrition in longitudinal surveys. This review should take account of experience and practice in the USA.

e) To facilitate the development of longitudinal datasets from administrative sources, and their linking to survey data, inasmuch as this is compatible with legal and ethical requirements. Ancillary to this, to initiate a review of the methodological issues connected with this type of use of administrative data.

f) To encourage and support improved documentation of, and access to, existing longitudinal datasets, especially by the exploitation of advancing information and communications technology.
1. Introduction

The Office for National Statistics (ONS) has invited us to prepare a paper discussing the requirements for longitudinal data for purposes of government policy, and the steps which would be involved in meeting those requirements.

The emphasis upon policy is important, since there are practical issues in meeting these needs that do not arise, at least not to the same extent, in the case of academic research. These include the timeliness of information in response to new issues, and the clarity or transparency of conclusions that may be drawn.

In more detail, the aims of the paper are:

a) To outline in non-technical terms the main methodological issues involved in the use of longitudinal data

b) To discuss existing sources of data that could contribute to policy needs

c) To consider alternative sources of data or alternative methods for meeting policy needs

d) To consider aspects where longitudinal data might be strengthened, for example, by identifying significant gaps in information, by seeking greater comparability, representativeness, or data quality, by increasing accessibility of data sources

e) Where appropriate, to make proposals for improvement.

We start by considering some definitional issues and the background to the increased policy interest in longitudinal data. Then items a) and b) are taken together since it seems better to discuss methodological issues in terms of actual data sources, rather than in the abstract. Similarly, c) and d) are taken together since they are highly complementary. Item (e) concludes the paper.

The limitations of our discussion should be stated. Within the time available, we have not been able to conduct any systematic cross-disciplinary review of longitudinal datasets in Britain or the literature on these. Our examples are mainly drawn from the longitudinal datasets we have used, primarily in studies of the labour market and social disadvantage. We lack familiarity with other areas where longitudinal data are important, such as child development and health. Accordingly, both the issues we discuss and our proposals are illustrative rather than exhaustive. In addition, we make only passing reference to the various methodological and theoretical debates in the analysis of longitudinal data, and exclude any consideration of the impact of new technology on future methods of data collection (though there is little doubt that this will be considerable). Our discussion is confined to relatively practical issues of method and application, and will hold no surprises for those who provide or use longitudinal data. Despite these limitations, we hope that our views on these topics will aid further constructive debate.

\footnote{We have however been able to draw upon a recent ONS reference work on UK longitudinal data sources (Office for National Statistics, 1999).}
2. Some distinctions and definitions

Longitudinal data can arise when at least two observations are made on the same units at different times. This distinguishes longitudinal data from single cross-sectional surveys, where observations are made on a sample at just one time, and from repeat cross-sectional data, where observations are made on separate though similarly constituted samples at different times. It also distinguishes longitudinal data from aggregate time-series data, where observations are based on a collection of units at each time. The term ‘panel data’ is often used for repeated observations on a sample, especially where techniques of ‘panel data analysis’ are to be applied (Hsiao, 1986).

Cohort data are a particular class of longitudinal data, in which the sample consists of units that pass through a defining event at about the same time, and are then followed over time. Examples are birth cohorts, age cohorts, or school-leaver cohorts.

The most usual units for longitudinal data are individuals, since these maintain their identity over time. Note that households and families do not maintain their identity in the same way that individuals do, since their members can change over time. For this reason, longitudinal household studies are, in general, studies of individuals who were members of households at some initial time of sampling, together with other individuals who subsequently join their households.

Longitudinal data can vary in their overall period, the frequency of observation, and the regularity of observation. Major panel surveys in the USA, Britain and some other countries tend to have annual interviews which continue for periods of 10 years or more. But much longitudinal data consist only of an initial observation and a single follow-up. In some cases, the members of a sample are followed up over long periods of their lives and with long gaps between interviews: this is characteristic of some birth cohort studies. In yet other cases, the longitudinal record is built up cumulatively at irregular times – for example, an individual’s health record may be up-dated with each visit to their doctor or to hospital. These differences greatly affect the practical uses to which longitudinal data can be put and the forms of analysis which can be applied. For example, frequent interviews or cumulative administrative records are more likely to capture transitions, or brief episodes, than are widely-spaced interviews which rely on the respondent’s memory.

Social surveys with repeated interviews can generate longitudinal data to specification. But longitudinal surveys are not the only sources of longitudinal data. Information is accumulated over time in a longitudinal form by many administrative systems, such as health records, tax and National insurance records, and benefit payment records.

Both longitudinal surveys and administrative databases require subsequent processing to construct a usable longitudinal dataset, and this may involve a lot of work. Many potential longitudinal data sources have not yet been used to create longitudinal datasets. Examples would include local authority records of Housing Benefit claims, or the New Earnings Survey as a potential source of longitudinal data concerning employers.

Longitudinal datasets can, in principle, be constructed from separate single or multiple observations in different sources, if these different sources contain common identifying information that permits a link to be created. The existence of standard identifiers, such as National Insurance number, potentially enlarges the scope for longitudinal data but there are often practical and legal barriers to be faced.

Longitudinal data can be simulated in several ways. The most common method is to ask retrospective questions within survey interviews, to fill gaps in information about the past. Virtually all surveys - including longitudinal studies - ask retrospective questions about both
distant and recent events, such as age of leaving school, or the number of hours worked during the previous week. In the past two decades this has been extended to include more extensive 'work and life history' sequences.

Another method of simulating longitudinal data, is to construct 'pseudo-panels' within large repeated cross-sectional surveys, by defining cohorts which reappear in each cross-section. For example, one might identify 16 year olds in year 1, 17 year olds in year 2, 18 year olds in year 3, and so on. This method depends upon the ability to define a cohort unambiguously across separate samples.

In summary:

- Longitudinal data exist where there are observations on the same units over time
- The overall time-period covered, the frequency of observation, and the regularity of observation all affect the uses to which longitudinal data can be put
- Longitudinal data sources include both surveys and administrative records
- Longitudinal datasets are constructed from longitudinal data sources, and this is often a substantial task in its own right
- Longitudinal datasets can be constructed from multiple sources when a common identifier is available
- Longitudinal data can to a limited extent be simulated, within both cross-sectional data sources and longitudinal surveys.
3 The potential need for longitudinal data

Why are longitudinal data needed? Are they needed more now, or in the immediate future, than they have been in the past? Without some assumptions about the underlying need, discussion about the provision of longitudinal data may lack direction.

There are four types of need for longitudinal data which we identify and distinguish:

- Descriptive information of two kinds:
  relating to life-patterns
  relating to life-transitions
- Information required to explain and interpret social change or social problems
- Information required to evaluate policies or programmes.

3.1 Descriptive information

Government has a general need for descriptive, arithmetical information about society and about social change, without which it cannot make sound plans. However, most of this need can be met by cross-sectional surveys that are repeated at regular intervals, because social changes can be inferred from changes in the proportions or averages observed across years. We have assumed, in all the following discussion, that longitudinal data are not perceived as a substitute for repeat cross-sectional data. Indeed, longitudinal surveys are less appropriate instruments for estimating most types of social change between periods than are repeat cross-sectional surveys. We must therefore identify the particular kinds of descriptive information that cannot be provided by repeat cross-sectional methods.

Characterising individual life-patterns

Cross-sectional information reveals how many people are in particular states at a given time. Some characteristics defined in this way are likely to be continuing, persistent, perhaps even permanent, for example, whether a person has a qualification or a vaccination. However, many other states can be transitory, persistent, or recurrent, and which of these they are may make a considerable difference, not only to the individuals involved, but also to their policy significance. Unemployment provides an example. To the individual, it matters a great deal whether unemployment is a brief and isolated episode, or whether it is long-term and/or recurs frequently in their working career. At the aggregate level, it also makes a great difference to policy if unemployment is widely spread in the population, with most people having only a slight experience of it, or if it is concentrated in a narrow group of individuals, most of whom have long and/or recurrent periods out of work.

Individuals experiencing one type of problem or disadvantage tend to have others as well. Each of these problems may come to the forefront at different times, and this makes it more appropriate to assess their difficulties over a substantial period in their lives rather than at one single point. At the point of service, this type of thinking is represented by the “case management” model where clients have a one-to-one relationship with an adviser, who acts as the gateway, advocate and hub for the whole range of helping and facilitating services needed to address the range of problems. At the policy level, as ‘joined-up government’ advances, descriptive information that shows the extent, persistence, recurrence and interconnectedness of problems will increasingly be required. The solutions offered to problems will need to be assessed at the individual level over more extended periods, as these

---

2 Longitudinal surveys are appropriate instruments for measuring social change which takes the form of within-person change over time, such as change in the hazard of marital dissolution. For this type of purpose, however, longitudinal surveys must themselves be repeated.
solutions seek to become more fundamental and less palliative. Concepts such as 'pathways to employment' and 'employability' have been adopted in European Union employment policies, specifically to underline the need for continuing, long-term improvements for disadvantaged workers.

Another reason why there is likely to be an increasing need for longitudinal data is the emphasis on 'social exclusion' as a central theme of European social policy. This complex concept refers among other things to the role which intensive or cumulative disadvantages have in preventing individuals from taking a full place in social activity and citizenship. This concept goes beyond multiple disadvantage, and points towards individuals' long-term position in society.

Describing transitions

Much social change comes about through a change in the rate of transition between states or behaviours. Often, too, the outcome results from two-way traffic, with some people moving from A to B and others from B to A. Repeat cross-sections will not fully reveal these movements, and whenever one is interested in transitions, longitudinal data are likely to be needed.

The need for increased information about transitions results from increasing social fluidity and a less formalised or regimented society. Consider the case of marital status. In some earlier generations, most individuals made only few partnership transitions and these were nearly always legalised. Today, this picture is complicated by more frequent transitions and by legal distinctions between married and non-married partnerships and, consequently, between types of partnership break-up. Our usual classifications of marital status are becoming increasingly inadequate to describe the complexity of partnership histories. Similarly, in the labour market, there has been a substantial growth in temporary or contract employment, self-employment and second jobs, and perhaps a growth of casual employment and the 'grey economy' though the evidence for this is largely anecdotal.

Capturing and linking individual transitions over a period of time is not easy. If policy needs demand a fairly complete picture of individuals’ statuses and activities over time, then new and increasingly sophisticated methods of data capture and analysis will probably be required.

3.2 Explaining long-term social change

Increases in descriptive information about people over time stimulate a desire for explanations about long-term causes and consequences. This is of primary interest to the social science academic community, but government may also wish or need to have more penetrating explanations of the social problems which present themselves, particularly if preventative policies become favoured, since these have to be directed at supposed causes.

Some commentators feel that longitudinal data make causal explanation much easier, even suggesting that causal relationships can be directly inferred from longitudinal descriptions. We caution against this view. Longitudinal data may uncover surprising associations between individuals’ circumstances or behaviour at different times, and these findings often suggest new causal hypotheses. These insights should not be confused with rigorous tests of causal hypotheses, or with the identification of causal mechanisms. Superficial interpretations can influence policy in socially harmful ways. For example, in the USA’s 'underclass' debate recurring problems for the same individuals have sometimes been attributed to their personality, motivation or upbringing. These untested hypotheses have then been directly translated into wide-reaching policy prescriptions.
Longitudinal data are inherently more complex than cross-sectional data. The proper analysis of longitudinal data is correspondingly complex. The simple and transparent findings, which are often desired for policy purposes, are hard to reconcile with the specialised and technically advanced nature of longitudinal data analysis. Sponsors who are motivated by policy concerns must be prepared to support rigorous analysis of longitudinal data, if they are to obtain reliable findings from such studies. But they should also be prepared for these findings to be complex, and need to allocate adequate time and effort to the interpretation, 'unpacking' and presentation of findings in a policy-relevant form.

There is also some danger in supposing that, because longitudinal data stimulate interest in causal explanations of social problems, only longitudinal data can be used to investigate causal relationships. Longitudinal data are often relevant and effective for this purpose, but in many cases other types of data can be used. In one of the most thorough reviews of the topic to date, Heckman and Robb (1985) concluded that almost all the results (concerning causal inference) obtainable with longitudinal data could also be obtained with repeat cross-section data or in some instances with single cross-section data. The choice between longitudinal and other approaches is often a matter of what is practical and cost-effective, rather than a matter of what is absolutely necessary.

One of the most powerful methods of advancing understanding of the underlying processes of social disadvantage is by piloting new policies and programmes to combat them, and then evaluating their impact. This is discussed next.

### 3.3 Evaluating policies or programmes

There has been an increased emphasis in recent years on the need for policies and programmes to be tested by the results of evaluation studies. This is sometimes referred to as 'evidence-based policy'. Many types of evaluation study demand longitudinal data, and it may be useful to provide a brief outline.

The term 'evaluation' is used in several ways. Sometimes it refers to what might more properly be called 'monitoring': collecting basic information on how the programme or policy is implemented, its costs, the number of people affected, their experiences both during and after their involvement, and so on. Many of these information needs can be met by cross-sectional data. Here, however, we use the term in its more formal sense of assessing the net impact of the programme or policy on the outcomes that are its object. This means measuring the difference between the extent to which these outcomes are achieved when the programme or policy is in place, and the extent to which they would have occurred in its absence. This type of question almost always requires information on individual transitions rather than aggregate changes, as most programmes are relatively small elements within a complex set of agents for change.

Sometimes programme outcomes can be measured immediately after participants have completed the programme, as, for example, with the effects of an intensive tuition programme on reading ability. In other cases, a certain length of time must elapse before outcomes can be measured at all, as with recidivism rates for offenders or with rates of job entry for unemployed people. Often there is interest in both short-term and long-term outcomes: do unemployed people who have received skills training keep the jobs that they may get, or do children who have received intensive reading tuition retain their advantage in years to come?

The net impact of a programme cannot usually be assessed simply by comparing participants with non-participants, because in almost all cases these groups differ systematically in ways that affect outcomes. For example, the better-motivated offender will volunteer for the rehabilitation programme, or the more 'employable' among the unemployed will be offered work placements. The counterfactual question, 'What would have happened if these people
had not been on the programme? can be addressed in a number of ways, with varying data needs.

Social experiments make perhaps the smallest demands for longitudinal data. If participant and non-participant groups are big enough and random allocation procedures meticulously followed, then it can be assumed that all prior differences between the two groups are eliminated, so that only post-programme outcomes need to be measured\(^3\). However non-experimental methods, such as matched comparison groups or statistical controls for selection effects, often require information on past history, not least because past behaviour is generally the best predictor of future behaviour. Another approach to assessing net impacts involves comparing trends in the participant group in the outcome of interest over a period time with trends in the population from which participants are drawn. Under certain assumptions, any closing or widening of the initial gap between the participant and non-participant groups can be attributed to the impact of the programme. Clearly this methodology requires data covering a long enough period for random fluctuations to be evened out.

\(^3\) Even so, there may be a need to assess outcomes at various points after participation in the programme, giving rise to a longitudinal data requirement.
4 Types of longitudinal data, and issues of method

This section discusses the main types of longitudinal data that are or could be available, and the difficulties or limitations that they involve. There are five main sources of longitudinal data:

- General purpose longitudinal sample surveys
- Long-term cohort studies
- Single-purpose longitudinal studies including some evaluation studies
- Administrative databases with a longitudinal element.
- Combined or linked uses of administrative data.

4.1 General purpose longitudinal surveys

Large continuing longitudinal surveys, with wide-ranging subject matter, potentially constitute a strategic resource for policy-relevant research. A new study might take many years to complete, whereas an existing longitudinal survey database might already have collected the information needed, producing considerable savings in both time and costs. In addition, large longitudinal surveys covering many different facets of economic and social conditions provide a long-term research resource.

The model for this type of general-purpose social survey comes from the USA. A large proportion of all social and economic research there has been based on two major longitudinal surveys, the Panel Survey of Income Dynamics (PSID) which began in 1968, and the National Longitudinal Surveys of Labor Market Experience (NLS)\(^4\) which began in the late 1960s, with additional samples in the early and late 1970s. These surveys have each accumulated data on more than 30,000 individuals, across several thousand items of information, and each has been the basis for hundreds of published studies and monographs, many government-funded. Numerous other longitudinal surveys, with narrower samples and subject matter, have also been conducted in the USA.

In Britain, the main example of this type of survey is the British Household Panel Survey (BHPS). The BHPS is based on a national sample of households constructed in 1991. Interviews of each household member aged 16 or over were conducted for the first time in that year. All the individuals have subsequently been followed-up annually, and members of new households to which they move, and new individuals joining existing households, are also included. Further waves are planned, with firm funding, to 2003.

The BHPS survey interviews cover a wide range of standard information. In particular, BHPS has detailed information on income of all types and on housing costs. Since this is a household survey, the information from individuals can be combined to produce household income variables and, from these, equivalised income can be calculated for each individual. Other topics covered each year by the BHPS include employment status and labour market transitions, family composition, housing, savings and assets, indicators of consumption, self-reported health status and health behaviours, childcare arrangements, and a range of attitudinal or perceptual variables. The usual range of demographic or classificatory variables

\(^4\) The NLS consisted of four separate samples: men aged 45-59, women aged 30-44, young men aged 14-24 and young women aged 14-24.

\(^3\) Council Tax Benefit was not included initially but the BHPS contains the necessary information for it to be imputed with reasonable accuracy.
is available. All told, this is a rich data source with many possibilities for longitudinal analysis.

**Methodological issues in general purpose longitudinal surveys**

The methodological issues that general-purpose longitudinal surveys must address, though often shared by other types of survey, must be dealt with over long periods and numerous interviews.

*Topic selection, breadth and depth*

One obvious limitation of a general-purpose longitudinal survey is that interviews cannot cover many topics in great detail. Over-long interviews increase costs and may lead to increased refusals at subsequent waves; in addition it is impossible to brief interviewers adequately on more than a limited number of specialist topics. The survey can either cover a few selected topics in depth or many topics with just a few selected questions for each. These choices have to be made in a relatively fixed way for years ahead, because much of the strength of this type of survey lies in having the same information at each wave. There may be some scope to add new questions in response to new concerns, but if these are numerous they will undermine the overall project.

In practice, general purpose surveys, including the BHPS, tend to focus upon the main, long-established topics (such as employment, income, family, and health) which are of common importance to many areas of policy and have a cross-disciplinary interest. Standard questions in these areas have proved their value over many years of research. Within such a framework, it may be possible to retain some flexibility for a few new questions, reflecting emerging policy concerns, to be added to the current sweep of the survey. This ad hoc questioning will benefit from having background data from previous sweeps to link with it.

Thus a general-purpose survey is more likely to meet future needs if it concentrates on “depth” rather than “breadth”. Given that future needs cannot be foreseen, it is better to provide excellent coverage of core variables that are likely to have a wide application. On balance, the issue of topic coverage may not be as problematic as it first seems. Knowledge of what are good and crucial questions has developed greatly over the past two decades. There may equally be a need for users to understand the potential value of the standard questioning, and to develop resourcefulness in applying it to their needs.

*Sample size*

The high cost of maintaining a general-purpose longitudinal survey places considerable pressure on sample numbers. Yet, to be truly general purpose, the survey must be able to provide information on many different sub-groups, as policy issues frequently focus upon quite small sections of the population.

The British Household Panel Survey (BHPS) has achieved around 10,000 interviews in around 5,000 households at each wave. This is a good size for inquiries on topics which concern the entire sample or large sub-groups: for example, income (whole sample), employment (whole sample), health (whole sample), earnings (employed sub-sample), transfer payments (recipient sub-sample). Sample size is more problematic for minority sub-samples: people with disabilities, those in further or higher education, women with school-age children or using childcare services, and so on. Also, the proportions experiencing particular kinds of transition (such as losing a job or getting divorced) in any one year may be small. Combining data from several years can sometimes increase numbers.

On balance, we feel that sample size in the BHPS is too small for its role as a general-purpose longitudinal survey. It compares unfavourably with the General Household Survey and the
Family Resources Survey, and the main longitudinal surveys in the USA have substantially larger sample numbers.

Sample attrition in longitudinal surveys

A general problem of longitudinal surveys is the cumulative effect of non-response. The longer a survey is continued, the more serious this becomes. A response rate of 70 per cent is regarded as normal in a single cross-sectional survey, but if this were the response at each successive wave of a four-wave study, the final response would be only 24 per cent of the initial sample. To combat this problem, multi-wave longitudinal studies have to put additional resources into tracking and re-interviewing the original sample. Yet, even with small sample losses at each stage, cumulative attrition remains problematic, especially if response at the initial wave is not well above the normal level.

In the USA, longitudinal surveys have usually achieved high initial response rates, and panel attrition, a serious problem at first, has been much improved. For example, the initial samples of the NLS at the end of the 1960s achieved a 98 per cent screening success rate, and initial interview rates (from the designated post-screening sample) of 91-94 per cent, depending on sub-group. Ten years on, 70-71 per cent of men, and 76-78 per cent of women (depending on age group), remained from the originally interviewed samples, leading to effective cumulative response rates of 65-70 per cent. The 1979 National Longitudinal Survey of Youth (NLSY) achieved an initial response of 90 per cent and also retained 90 per cent of these over the first 10 years, leading to an effective cumulative response rate of 81 per cent.

(Details of the procedures and response rates obtained in these surveys can be obtained from the Data Archive at the University of Essex.)

In Britain, sample non-response has been a troubling problem for many types of surveys. The BHPS obtained a positive response from 74 per cent of the households approached initially, and was then able to interview 95 per cent of the eligible individuals at the households to which access was obtained. Subsequently, 73 per cent of the sample interviewed at Wave 1 were also interviewed at Waves 5 and 6, and 70 per cent at Wave 7.

Random sample attrition does not matter, provided that the total sample size is still adequate. Comparison with the 1991 Census of Population indicated little bias in the initial BHPS sample, but subsequent attrition means that the BHPS sample now under-represents married people and those aged over 54 (Nathan, 1999).

Re-weighting to correct for sample attrition is readily available to users of the BHPS dataset. However, professional opinion differs on the issue of corrective weighting. Re-weighting evidently can only be applied to observable characteristics of the sample, but attrition may also be influenced by unobserved factors. For example, one reason for attrition is geographical mobility, and mobile individuals may differ from non-mobile individuals in terms of unmeasured motivational factors. Weighting on the basis of observable characteristics may even increase the unrepresentativeness of the sample in terms of unobserved characteristics.

The Quarterly Labour Force Survey as a general purpose longitudinal survey

Although the BHPS is the only general-purpose longitudinal survey in Britain comparable to those in the USA, it is also worth considering the Quarterly Labour Force Survey (QLFS) under this heading. Until recently, the QLFS (which replaced the annual Labour Force

---

6 These figures are not directly comparable with initial response rates because they are affected by mortality, emigration and other factors which reduce the Availability of original sample members.

7 Weights to correct for attrition are made available with the BHPS dataset.
Survey in 1992) has been used as if it were a cross-section survey, but its underlying design is longitudinal. The QLFS is based on a nationally representative sample of households, and it seeks interviews with all household members aged 16 or over. Excepting non-response, individuals are interviewed for five successive quarters, and one-fifth of the sample is replaced at each interviewing wave. The QLFS is suitable for the analysis of short-term transitions although less so for longer-term life-patterns.

In terms of its topic coverage, the QLFS is designed solely to provide information about individuals in the labour market. However, there is substantial questioning on health and disability, on receipt of out-of-work benefits, and (in the first and the last of the five interviews) on earnings and non-wage income. There is the usual range of classificatory questions, and it is possible to derive household information. There is also useful information on post-compulsory education and training. While narrower than the BHPS, the QLFS provides sufficient information to be regarded as a general-purpose survey. Use as a longitudinal survey has recently been facilitated by developing weights to correct for cross-wave attrition.

The QLFS has a much bigger sample size than the BHPS, with 110,000-120,000 interviews in each wave. This makes it possible to study minor sub-groups and unusual transitions. In addition, the QLFS is unclustered within sampling area, which increases its representativeness. Weighting is used to gross the cross-sectional samples to national population projections.

Sample non-response and attrition are likely to be less problematic in a short-period longitudinal study such as QLFS than in a panel survey spanning many years. For the 1996-97 period, the response rate at initial interviews was around 80 per cent and at the final interview, around 73 per cent. A further issue with the QLFS, however, is the use of proxy interviews. Whereas in the BHPS, these interviews are restricted to less than 4 per cent, in the QLFS the proportion is typically around 30 per cent. Certain questions are excluded from proxy interviews, and some analysts prefer to exclude proxy respondents altogether.\(^8\) If this is done, further re-weighting or another method of adjustment should be considered.

4.2 Long-term cohort studies

The most prominent type of long-term cohort study in Britain has been the birth cohort. There have been three post-war national birth cohorts, those of 1946 (the National Survey of Health and Development, NSHD), 1958 (the National Child Development Study, NCDS) and 1970 (the British Cohort Study, BCS). All three began with a focus on health issues at birth and in early childhood. While this focus has been retained in the case of NSHD, the two subsequent birth cohorts have developed into much wider-ranging studies which cover similar ground to a general purpose longitudinal survey. The following discussion of methodological issues focuses on NCDS.

Topic selection, breadth and depth

As a birth cohort study develops from a study of child development into a more general inquiry it faces the same problems of topic selection as general-purpose longitudinal surveys. These problems are aggravated by the longer periods over which studies such as NCDS or BCS continue, as the possibilities for analysis at later stages are constrained by the information collected earlier.

None the less, NCDS or other long-term cohort studies can, by a process of accumulation, offer a wider range of information about the sample members than any other study. For example, NCDS contains a mass of information about cognitive and social development.

\(^8\) A methodological study on the effects of proxy response is reported in the QLFS Methodology Manual.
(psychometric test scores, school assessments, etc.), health, physical measurements at various ages, and educational attainments. These can be linked to subsequent details of employment, occupation, earnings, family formation, or adult health. It is thus a unique resource for examining the links between parental home, childhood, adolescence and adult life. Information has also been gathered on the children of the original sample, so that it is possible to link three generations.

However, there are difficulties in using information that spans long periods. These difficulties include the complexity of analysing so many layers of information, the multi-disciplinary knowledge required to bring them together, and the dearth of concepts and theories about long periods in individual lives. Much social and economic analysis is carried out under the general assumption that causal processes can be broken down into largely separate stages. An analysis which links events widely separated in time has to show that this explains more than the conventional approach. Our tentative impression is that few researchers have ventured into this little-explored and possibly risky territory. For policy purposes, also, there may be doubts that conditions several decades back can still be relevant. Childhood conditions may have been important for people now in mid-life, but in what way can such a link (if demonstrable) be related to current concerns about (say) national childcare policy? So, although the range of information accumulated by long-term cohort studies is remarkable, more attention may need to be paid to how this information can be used. A relevant step would be to promote more discussion between researchers involved in long-term analysis, and policy makers who want to apply findings to current issues.

Sample size and representativeness

The NCDS has a much bigger sample than the BHPS – initially, more than 17,000 respondents (with a parental response rate of 98 per cent). The sample is unclustered and representative of all sub-groups within the birth-week. It generates substantial sub-groups at each life-cycle stage – for instance, unemployed young people, recent parents, or university graduates – and this has been one of its chief attractions to researchers.

On the other hand, a birth-cohort is at any time representative only of one age group, whose experiences are influenced by the historical times at which they reach critical transitions. For example, most of the NCDS sample entered the labour market between 1974 and1980, a period of deteriorating prospects for young people, but far less adverse than the first half of the 1980s. Thus, their employment histories differ from the histories of those who entered the labour market earlier or later.

Because of these effects of historical time, the value of a cohort study is increased if there are other not-too-distant cohorts available for comparison. A single cohort study is for many purposes not suitable for studying the impact of policy changes except by comparison with other cohorts. A number of comparisons have been made across the existing birth cohort studies, but the future potential for this type of comparison has been impaired by the lack of any further birth cohort study since 1970. Similarly, the year 2000 birth cohort study which is now projected will lack any longitudinal comparison study from the previous two decades. This problem has been recognised in the draft (September 1999) policy statement of ESRC and ONS on the future development of longitudinal data.

Sample attrition

The British birth cohort surveys achieved impressive sample retention rates while cohort members were still in compulsory full-time education and could be traced through schools,
but subsequently have been more seriously affected by attrition\textsuperscript{10}. At the age of 11, 87 per cent of the initial NCDS sample were successfully followed-up, and at age 16 the proportion was still 83 per cent. At age 23 this fell to 76 per cent of the original birth cohort less those known to have died or emigrated, and at age 33 to 74 per cent, calculated similarly. The effects of sample attrition include under-representation of ethnic minorities, early low achievers, and those from disadvantaged family circumstances (Iyer, 1984, cited in Nathan, 1999). Adjustment weights for non-response and attrition have not been supplied with the NCDS or BCS.

4.3 Single-purpose longitudinal studies including evaluation studies

Chapter 3 outlined some of the reasons why evaluation studies often demand longitudinal data. Outside of evaluation studies, single purpose longitudinal studies designed to serve the needs of social policy makers are relatively rare, because of their long lead times and high costs. They are useful when the phenomenon under study is of considerable social importance and has a dynamic that must be understood before appropriate policy interventions can be designed. Examples include the Youth Cohort Study (YCS), which has surveyed a succession of cohorts of young people at the end of compulsory education, the Programme of Research into Low Income Families (PRILIF) with its longitudinal study of lone parents and the Survey of Living Standards During Unemployment. Panel studies measuring changes in public attitudes (for example, the panels used by polling organisations to measure changes in voting intentions in the run-up to an election, or the British Elections Study’s 1992-1997 panel) may also have an indirect influence on the social policy priorities adopted by politicians, and may be commissioned by political parties but not usually by government. In this section we discuss some of the methodological issues that arise in single purpose longitudinal studies of all of these kinds.

**Topic coverage and sample size**

Because they have one main focus, well-designed single-purpose longitudinal studies should avoid most of the problems of topic coverage and sample size that arise when we try to use general purpose surveys to study particular sub-groups or to evaluate specific programmes. Indeed, it is in principle unlikely that any general-purpose study would be able to supply comparably detailed data on particular topics (see the earlier discussion of general purpose surveys).

Even with a single-purpose study, analysis will of course sometimes throw up further unplanned questions to which sample numbers in sub-groups do not permit an answer.

**Sample response, attrition and representativeness**

Like most longitudinal studies, single purpose longitudinal studies tend to suffer from biased sample attrition. This is particularly problematic if attrition is correlated with outcomes of especial interest. For example, in the YCS, young people who are unsuccessful on the courses on which they have embarked are less likely to respond than those who do well. A major cause of sample attrition in longitudinal surveys is geographical mobility, and unfortunately this is often associated with the phenomenon under study. In the evaluation of a job-training programme, for example, people may move house in order to take up a new job, or lone parents may move house if they find a new partner. Though statistical techniques can be used to correct estimates of programme effects for biased sample attrition, it is desirable to

\textsuperscript{10} The 1946 NSHD appears to have retained larger proportions of its sample in the long-term than the NCDS or BCS (see Nathan, 1999) but has only attempted to follow a sub-set of the original birth cohort. The study is not in the public domain and we do not have full methodological details.
combine these with serious efforts to maintain contact with sample members. In the case of certain geographically mobile populations, telephone interviewing can be useful, as collecting data from those who have moved house entails only small extra costs. This was used, for example, in the evaluation of High Technology National Training.

The problem of non-response in single-purpose longitudinal studies conducted for evaluation purposes is often aggravated if they rely on administrative sources to supply names and addresses. In a number of studies, problems with the addresses extracted from administrative records have represented the biggest source of non-response. In the evaluation of Training for Work, for example, nearly one in five addresses supplied for TTW leavers was untraceable or referred to an empty or derelict property (Sproston 1999). In the evaluation of the first Project Work pilots, one quarter of the original issued sample appeared to have moved from the address on Employment Service records without leaving a forwarding address (Stratford 1998). It is hard to understand why such problems are so common, especially when the addresses supplied are those used for the payment of benefits. Some of the biases consequent on this source of non-response can be guessed at; others are wholly unknown. An investigation into the causes of problems with addresses on administrative records could be of long-term value to a number of studies.

In evaluation studies using a matched comparison group design, the problem of sample attrition is somewhat different - here the comparison sample becomes in effect a quota sample, with fieldwork continuing until sufficient responses have been obtained to match the required profile. The problem here is that this technique could make it impossible to generalise from the comparison group to the population of all who are eligible for the programme. The solution is to calculate selection probabilities for each member of the comparison sample, so that estimates of programme effects can be weighted back to apply to the full eligible population. This of course complicates further the selection procedure for the matched comparison sample.

In some circumstances, it may be better to settle for a smaller achieved sample size in order to reduce attrition bias. The regular 100 per cent follow-up surveys conducted by the Department for Education and Employment (DfEE) of all leavers from special government programmes can have response rates as low as 35- 40 per cent, and as far as the present authors are aware, no corrections for response bias are made to the estimates based on these surveys. In this case, more useful information could perhaps be obtained if the follow-up was directed towards tracing a smaller random sample of leavers.

Most single purpose longitudinal studies cover a limited time span, but where the time span is long, the original sample will tend to become less and less representative of the present day population and thus less useful for understanding current social problems. For example, many sample members in the PRILIF longitudinal survey of lone parents who formed a nationally representative sample of lone parents when they were first interviewed in 1991 have now re-partnered, and given the increase in births to teenagers during the 1990s, the age structure and partnership histories of the sample have become less representative of lone parents at the end of the millennium. Thus to gain a full picture, longitudinal data need to be supplemented by repeat cross-section data - as was the case with PRILIF in 1993 and 1994, and with the new sample of lone parents included in the 1999 Survey of Low Income Families. Alternatively, representativeness can be maintained if the original panel is continually replenished with new entrants. An example of this is the British Youth Panel of 11-15 year olds (an 'add-on' study to the BHPS), which each year takes in a new cohort of 11 year olds and loses a cohort of 16 year olds. However, panels that are neither supplemented by repeat cross-sections nor continually replenished can be valuable for other purposes, notably for studying various aspects of development, patterns of activity and behaviour over the life course, and the long-term consequences of earlier events.
Problems of timing and timeliness

One of the main problems with single-purpose longitudinal studies is the time needed to complete them. They are often commissioned to meet an immediate information need, and by the time the study has been completed, this need may have passed. This problem is particularly evident in the case of evaluation studies, where the timing is rarely ideal. If evaluation is begun before the programme to be evaluated has reached a settled and workable format, it may be difficult to identify or measure interventions. If evaluation is left too late, the programme may be closed down before the study is complete - a problem that arises from the mismatch between political and research timetables. This may be a topic that deserves further strategic review.

Some of these timing problems might be avoided if a clear distinction were made between piloting and evaluation. Piloting demands speedy and regular monitoring information, and the programme should be left free to evolve in response to such feedback. Formal evaluation, in contrast, requires the programme to have already reached a stable and viable format. Timing problems could also be much reduced and results produced much more quickly if studies could draw on existing general-purpose surveys and/or administrative databases, avoiding the need to collect new data.

Timing problems are generally less severe with single-purpose longitudinal studies designed to increase our general understanding of a phenomenon rather than for evaluation purposes, as results are less likely to be needed to inform immediate policy decisions. In addition, the first sweep of such studies often provides much valuable information about the phenomenon under study, in advance of further follow-up sweeps - as did the first sweep of the Living Standards During Unemployment study, which took place three months after signing on.

Other methodological issues

With some longitudinal studies, it is necessary to conduct an initial survey sweep, not because there is intrinsic interest in the data from that sweep, but because there is no other way to solve purely practical problems. For example, evaluations of government programmes for unemployed people are generally more concerned with medium- and long-term outcomes than with short-term outcomes. Nevertheless ex-participants usually have to be surveyed as soon as possible after they have left the programme, because if the first sweep is delayed, addresses become too out-of-date to yield a satisfactory response rate. If other ways of tracing individuals were available, this early sweep could become unnecessary.

4.4 Administrative databases with a longitudinal element

Administrative databases that accumulate information about individuals are longitudinal in nature. In some countries, these are the sources on which a large amount of social and economic research depends. For example, in Sweden and Denmark, the national agencies responsible for statistics construct combined databases, also called “registers”, which show, for each individual and each year, such items as earnings, taxable income, welfare payments (divided into various types), housing costs/imputed rents, pension contributions, number of weeks in employment, participation in labour market or training programmes, area of residence, sex, age, educational and vocational qualifications, marital status, number and ages of children if any, employment status of partner if any – and many other details. Researchers can then obtain extracts from this database, covering a selection of variables, types of individuals, and years, for the purpose of a particular project.

Another example concerns programme evaluation research in the USA. Those taking part in programmes or demonstration projects, and their non-participating comparators, are commonly tracked through subsequent years, using social security identifiers to pick up
details of welfare payments received, months in paid employment, and earnings. This is the stock information used to produce estimates of the programme impacts.

In Britain, medical records have long been used for clinical or epidemiological studies, subject to ethical approvals. A variety of administrative databases have been adapted for research within government departments, but are not accessible to independent researchers (these are described in ONS, 1999). The JUVOS\(^{11}\) database has been playing a growing role in evaluation studies of welfare-to-work programmes, and a special evaluation database drawn from administrative sources has been set up for the evaluation of New Deal (Daly and Bentley, 1999).

**Topic coverage**

As each public administrative database has a narrow purpose, such as control of benefit payments or of National Insurance contributions, we would expect each to cover only a small range of information. They are likely to become more useful for research purposes when several databases are combined or when administrative data are combined with survey data.

It does not seem likely that administrative data could end the use of surveys, since surveys ask many kinds of detailed questions that will never appear in administrative sources. Rather, administrative data are complementary to survey data. The largely factual and financial information available in administrative sources is hard to obtain from individuals, or tends to be reported inaccurately. Administrative data on these topics should generally be more accurate and more complete. Again, individual recollection of the dates at which various transitions took place tends to be quite poor, and may be more accurately obtained from administrative sources.

The range of classificatory information held on administrative databases is not too limiting, especially in the context of panel data analysis\(^{12}\). In JUVOS, for example, one can obtain information on sex, age, marital status, and disability. In addition, the history of prior unemployment can be considered as proxy information concerning the combined effect of various labour market disadvantages. The range of classificatory information entered on the Labour Market System (LMS) used for unemployed job seekers is much wider, and this is now being used to a greater extent in the New Deal Evaluation Database. For example, the various forms of assistance provided through New Deal are recorded so that one can obtain process information as well as outcome information.

**Sample size and representativeness**

A feature of the panel data research emerging from the Scandinavian countries is the large samples which they have drawn from the administrative registers, and the large number of years which they span. This permits the effects of some government policies to be estimated with a high degree of precision\(^{13}\).

In addition, since administrative databases are generally national in scope and do not involve multi-stage sampling or clustering, they should reduce problems of representativeness for sub-groups. However, there are two limitations to be noted. First, any one database is representative only of a specific section of the population. Benefits databases cover claimants, while National Insurance databases cover employed people. Second, and perhaps more seriously, some individuals and groups may avoid contact with public agencies and thus

\(^{11}\) Joint Unemployment and Vacancies Operating System.

\(^{12}\) In certain types of panel data analysis, it is possible to control for individually specific factors which are not measured in the study. This facility is of course no remedy for the absence of classificatory information which is of direct interest.

\(^{13}\) Jensen et al. (1993) is an example of a large panel sample from administrative data, which is used to estimate precisely the rather small employment effects of short training courses.
be under-represented on administrative databases. For example, they may not claim benefits that they are entitled to, or they may seek to conceal earnings. For some purposes, this source of under-representation may result in significant biases. These biases, however, are not necessarily easier to overcome in surveys, where similar groups may have higher refusal rates.

A related issue is that the categories used in administrative data may not correspond to the differences in condition experienced by people. Aggregate unemployment in the UK is measured by the ILO definition, which is behavioural, but some studies of unemployment have been based on samples drawn from claimants of unemployment benefit, an administrative category which excludes many unemployed people, and includes others who are economically inactive.

**Sample attrition**

Sample attrition does not take place for administrative data in the same way as in surveys, and this is one of the main attractions of using administrative data for longitudinal data analysis. For example, individuals who move from one area to another can still be tracked through their National Insurance number.

However, there will still be some degree of attrition over the long-term, resulting not only from death, emigration and institutionalisation, but also from recording errors, as when unique identification codes are omitted. In addition, some types of case may be excluded from or drop out of electronic records, as happens with unemployment benefit claims handed clerically rather than in JUVOS.

**Special methodological issues for administrative data**

Within the confines of the types of information available in administrative data, there are evidently substantial methodological advantages in terms of low collection costs, large sample sizes, representativeness and low levels of attrition. There are however other methodological issues which may require more attention if administrative data are to be used for longitudinal analysis.

**Incomplete information**

The potential value of administrative data can be impaired by incomplete information on individuals. To give a specific example, the JUVOS database classifies individuals who terminate a claim for unemployment benefit according to their destination – for example, whether they are exiting to a job or moving to a different benefit. This important information is weakened because about 25 per cent of cases are recorded as ‘destination unknown’.

These are cases where the claim has been terminated without the local office being able to obtain the destination from the individual. Gaps in information also sometimes occur in relation to classificatory information; for example, information on individual qualifications has proved difficult to collect. Depending on the purpose of an analysis, gaps such as these may reduce the effective sample size considerably and may introduce bias and non-representativeness into the remaining cases.

More complete information might be achieved if it were possible to link across datasets that are complementary. For example, information on National Insurance contributions might be used to infer that an individual has moved from unemployment into employment.

---

14 See Hales and Collins (1999) for a recent investigation.
Limits on tracking

An important limitation on any one administrative data system is that there are gaps in tracking. In the case of a system based on benefit, tracking is suspended when the claim is terminated. There is then a gap in the record until the next claim, if any, commences. Even if the main focus of an inquiry is benefits, there is a difficulty for the analysis if one knows nothing of an individual's circumstances when not claiming. Once again, the ability to link between different sources of administrative data is central to obtaining a more continuous record.

4.5 Combined or linked uses of administrative data

The combination or linking of administrative data has until now been restricted by legislation. We understand that legislative changes now under consideration may, if adopted, relax these restrictions in some ways. It is beyond our competence to comment on legal issues, and accordingly, our discussion assumes that there is no legal impediment to the linking or combination of data from different sources. We also assume that such a development for purposes of statistical information and research would provide protections for individuals that accord with ethical standards of research.

We focus specifically on the potential advantages of administrative data when used in conjunction with surveys. An example of combined survey and administrative data is the linked New Earnings Survey and JUVOS dataset (NES-JUVOS). The NES consists of earnings data, returned by employers from payroll records, of a 1 per cent random sample of individuals, which is nested within the JUVOS sample. Individuals are sampled by a fixed combination of National Insurance digits and so tend to recur in the sample from year to year. The database makes it possible to study individual wage and unemployment dynamics over a long period, with the added advantage of combined classificatory information from JUVOS and NES. This potentially valuable database is however not in the public domain.

Sample construction

Administrative databases can play an important part in constructing the samples for one-off longitudinal surveys. However, as noted in the discussion of single-purpose longitudinal studies, address details contained in some databases are of lower accuracy than might be supposed. In principle one might use historical administrative databases, and interview samples drawn from several years ago to study the long-term impacts of past events or interventions, as now revealed. This would provide a simple and low-cost method of conducting a longitudinal inquiry. The reason why this cannot be done in practice is that there is at present no means of up-dating addresses from the past. The ability to obtain current addresses from other linked databases might overcome this limitation.

Background information

Administrative databases contain historical information that may be valuable when added to current survey data, such as previous histories of unemployment or benefit receipt. Historical information of these kinds is of obvious relevance to current problems concerning the labour market, deprivation and exclusion.

Long-term outcomes

Administrative databases can be used as a means of tracking outcomes for individuals who have been surveyed at a particular point in time, provided that linking identifiers are available. The specific advantage of administrative data is for outcomes which stretch beyond
the one or two years that follow-up surveys usually cope with. With administrative data, tracking over many years becomes possible.

*Multiple outcomes*

The one-dimensional information in any one administrative database can be problematic for outcome measurement. One may over-estimate or over-interpret the importance of the data one has, neglecting the potential importance of the data one does not have. For example, research on the relationship between benefits and unemployment has tended to use exit from unemployment as the outcome, rather than exit to employment. Though the latter is the preferable measure from the viewpoint of economic theory, it has not until recently been available in benefit claim databases.

The linking of surveys with administrative sources becomes more attractive if a range of databases can be made available. In the USA, outcomes from welfare-to-work programmes have been tracked using a variety of administrative sources to give information on employment, earnings, welfare receipts, income, and criminal arrests.
5. Illustrative applications for longitudinal data

In this section we take some areas of policy interest and consider how longitudinal data might be developed for them. These examples are meant to illustrate some of the general issues we have already discussed, but they also raise some more specific points. The examples reflect our own interests and experience, and do not claim either to be areas of especially high priority (although all are important), nor in any way representative of all aspects of policy.

5.1 Incomes and low-income households

Levels of income, and how they vary over time, are of evident importance for a wide range of policy issues. There is generally a particular policy interest in low levels of income. This has strongly influenced the development of research on income dynamics, with many studies focusing upon some 'low income threshold', the frequency and recurrence of movements across this line, either down or up, and the length of time which individuals remain below the low-income threshold. The threshold is partly arbitrary, since there is no actual discontinuity in the distribution of income, and studies sometimes show results for several alternative definitions.

To provide a basis for studying income dynamics, a dataset must observe income in a consistent way for individuals at several times. This is no mean task, since income comes from many sources (earnings from one or multiple jobs, benefits in-work and out-of-work, pensions, property rents, interest payments, and so on). Also, while some studies have tried to disregard income that is not directly paid to an individual, the most usual practice is to take account of the whole household income and to attribute this to the individual, after taking account of the number of people who share in the income. This increases the complexity of the task.

There are several methodological issues of particular importance for studies of income dynamics. Missing information is one. The BHPS, for example, obtains complete household income information for only 60 per cent or so of the sample each year. Some sub-groups, notably self-employed people, have a particularly high incidence of missing data.

Another issue is the extent to which period-to-period transitions can be artefactual, as a result of definitional issues or errors in the data. Some exploratory analysis has shown quite large disparities between what appear to be moves into or out of low-income conditions, and respondents' reports of whether they feel better off or worse off than at the previous interview. This may in part arise because of mis-reporting, if for example an individual forgets a source of income at one interview but reports it at the next. It may also arise through natural variability in the flow of income, which evens out over a period that is longer than the period defining the analysis.

One of the main interests in income dynamics has been to identify the circumstances which 'trigger' a transition into or out of the low-income category. To analyse this, information is needed on changes in circumstances, which can be timed reasonably accurately against the income transitions. For example, loss of employment (either by the individual or by other household members), separation from a partner, and entry into retirement are common triggers for a movement into the low-income category, and these are the kinds of events which a dataset should be able to identify. One may also seek to decompose income transitions by considering which elements of income have fallen or risen in conjunction with the transition. Analysis at this detailed level may also help to distinguish between artefactual and real income transitions, particularly if transitions are followed over more than two periods.
How do various types of longitudinal datasets help to solve these various problems, or to advance this area of inquiry in other ways?Sophisticated methods of imputation for missing income data have been developed with the BHPS, but users have freedom either to use these or to adopt other methods of their own. In addition, the BHPS has detailed information on the likely 'triggers' of income transitions, and in many instances their dates; it can be used for detailed income decompositions; and it has subjective report and attitudinal data on financial circumstances. Because of these advantages, the BHPS is the main site for income dynamics research in Britain. However, a limitation is the collection of income data annually, since attempts to time the changes in income and in the triggers of change, within the year-to-year interval, must rely on the respondent's recall.

The collection of income data from all waves of the QLFS could complement the BHPS and lead to important advances in the field, although it must be appreciated that income data from the former are necessarily less detailed and less complete than from the latter. At present, the QLFS parallels the BHPS in collecting income data at two waves separated by one year, albeit in every quarter. More robust quarterly income data should produce better estimates of total income over a year within the QLFS, and would offer an opportunity to cross-calibrate income estimates from the quarterly and annual interviewing series. Methodological work of this kind is important if income dynamics are to become an increasingly important instrument for policy analysis. The explanation of income transitions or fluctuations would also benefit, because quarterly collection would help to identify triggers for income changes.

According to Jenkins (1998), it is essentially for cost reasons that the collection of income data in the QLFS is restricted to two waves. However, quarterly longitudinal collection of income data could be cost-neutral if the income questions were asked only of a random sub-sample of respondents.

Many of the elements of income data reside in administrative databases, and we know that Scandinavian countries have brought such sources together sufficiently to conduct income analysis. To the extent that a similar approach is technically feasible in Britain (and it could be a large project to establish whether it is), this would remove some of the problems resulting from missing income data and inaccurate recall or reporting. Of course, it could not remove all sources of missing data or error, since not all income sources would be covered in administrative data, not all income would be declared, and records contain errors. The experience of the Scandinavian countries on these issues would be of interest.

With a new focus on social exclusion, several new topics in income dynamics are likely to come to the fore. One topic is the extent to which recurrent or persistent low income is associated with loss of participation in various forms of social life and participation. This is the type of question to which the BHPS, with its wide-ranging subject-matter, is well suited. However, the BHPS does not provide adequate coverage of subgroups such as ethnic minorities, where the topic may be particularly relevant.

Another topic is the identification of the long-term antecedents of persistent or recurrent low income. This is the type of question to which long-term cohort studies might be expected to make a distinctive contribution. Although the NCDS and the BCS do not have much income data in the early years, they could still be used to investigate other types of early influence, such as early attainment.

5.2 Longitudinal data on youth transitions

There are several sources of longitudinal survey data on young people.

a) One of the most extensive, for England and Wales, is the Youth Cohort Survey (YCS). Since 1985 this has followed successive cohorts of young people over the first two and a half years after the end of compulsory education, tracking their progress in both
education and the labour market. For similar purposes, Scotland has its own regular follow-up survey of young people using a somewhat different methodology.

b) The British birth cohort studies, discussed in section 4(2), are one of the few longitudinal data sources for young people that cover the whole of Great Britain including Scotland. However members of the most recent of these (BCS70) are currently aged 29, and grew up in a system that is now much changed.

c) The British Youth Panel, begun in 1994 as a supplement to the BHPS, also covers England, Wales and Scotland. It collects information annually from 11-15 year olds who, from age 16 onwards, may go on to be interviewed as part of the BHPS. However sample numbers (around 750 in total in any one wave) are too small to be of great use to policy makers, and could not be increased within the current overall design of the BHPS.

d) The QLFS is used by the DfEE as one of the data sources for its cross-sectional figures on the activities of 16-18 year olds. Though it offers much smaller sample numbers for young people than YCS, and a potential follow-up period of at most one year, its high data quality and detailed information on labour market activities suggest that it might be more extensively exploited for the study of some aspects of youth transitions.

There are currently quite limited national administrative sources of longitudinal data on young people. The DfEE's data base of qualifications, combining results from the separate examination boards to produce an individual record for each candidate, allows GCSE results to be linked to A/AS level results, but holds very little other information (DfEE 1995). In addition, there are tracking systems based on individual identifiers within FE colleges and within universities, enabling entry qualifications to be linked to progress within that institution and to the student's subsequent destination. However there is no general system for tracking young people between educational institutions, much less between education, jobs and training programmes, or of linking qualifications obtained from different awarding bodies. Perhaps the biggest information gaps relate to vocational qualifications, for which there are numerous separate awarding bodies, and to young people not in work, education or training, who are rarely eligible for social security benefits and therefore tend to fall between the various sources of tracking information.

Better links between administrative databases could bring real benefits in this area. Annual cross-sectional estimates of young people's participation in education and training are assembled by the DfEE from a variety of sources that present problems of reconciliation (see for example DfEE 1996). Moreover, these estimates relate only to England, because the Welsh Office is responsible for collecting information from schools in Wales, and uses slightly different data collection instruments. Administrative data relating to Scotland, which has its own education system, are of course entirely separate. The Careers Services also produce an annual report based on their own records of the first destinations of young people after 16, but this is purely cross-sectional and has very little information on other variables of potential interest.

There are currently a number of attempts at both national and local level to create tracking systems for young people. For example,

a) Warwickshire LEA is using unique identifiers to follow students within the county boundaries through from school to the end of university;

b) the University of Greenwich is running an ESRC-funded pilot project using fuzzy matching techniques to track students in higher education;

c) Careers Services have a new statutory obligation to trace 17 and 18 year olds within their own area and are establishing ways of sharing information between areas;

d) the DfEE is exploring the feasibility of creating a national linked database on 16-21 year olds.

All these are brave attempts to get round the basic obstacle to a national tracking system, which is that Britain (unlike Finland, which tracks all young people) has no single and unique person identifier used in all registrations.
In the absence of a national tracking system, policy makers have currently no option but to rely on the YCS as the main source of information on youth transitions. YCS offers a large random sample, with achieved sample sizes at the first sweep of each cohort (in the spring following the end of compulsory education) that range from 14,000 up to 25,000. A second sweep takes place two years later, and two cohorts have had or will have a third sweep in the early twenties. At each sweep, detailed information is collected on qualifications, education and training courses, employment, GST, earnings and social security benefits, together with a month-by-month retrospective diary of activities.

The main weaknesses of the YCS arise from the fact that, in order to achieve such large samples while avoiding massive fieldwork costs, data are collected by post via a fairly long and complex self-completion questionnaire. This, together with the questionnaire's concentration on qualifications and training, produces a high level of non-response and attrition that heavily biases the achieved sample against low achievers and the disaffected. Although a sophisticated weighting system corrects estimates based on the whole sample for this bias (Lynn et al. 1994), difficulties arise in analyses focussing on particular sub-groups, like truants or those with poor examination results, where actual sample sizes may be much smaller than the weighted number, and where those who have completed the questionnaires may not be representative of the sub-group as a whole, in terms of their unobserved characteristics.

The postal method has other disadvantages. The data on qualifications are not always of high quality (Bradley 1996), and YCS does not cover pupils in special schools, presumably because many would find it difficult to complete the questionnaire. Home background data are limited, with much missing information on parental occupation and qualifications, and there is substantial item-level missing data throughout the questionnaire because of the difficulties that some young people have in following routing instructions.

A switch to personal interviewing could remedy some of these problems, though this would increase costs a great deal. In addition, data on GCSEs, A/AS Levels and GNVQs could be made more accurate if the information were merged in from administrative databases. This would follow the practice of the Scottish School Leavers Surveys, and would have the added benefit of simplifying and shortening the questionnaire. However there would still be problems with vocational qualifications, for which no comprehensive database exists.

In designing survey instruments for very long-term longitudinal studies, there is often a trade-off between achieving comparability over time and incorporating changes to increase data quality and make the survey more relevant to current issues. The pace of policy developments regarding 16-18 year olds has caused some dilemmas of this kind in YCS, though many cross-cohort comparisons can be made.

There is strong policy interest in interventions in Year 11 or earlier that might help to increase the number of young people continuing in education or training after 16. YCS has always relied on recall for information on such interventions, and has even collected data on attitudes and evaluations retrospectively, though these retrospective reports are clearly likely to be influenced by actual outcomes. A survey sweep before cohort members reach minimum school leaving age would be of value, though again costly. The timing of the second sweep (in the spring of the third post-compulsory year) is also not ideal for some purposes. It coincides with the 'gap' year, and so misses a number of future entrants to higher education, and also takes place before many of those aiming for vocational qualifications of NVQ Level 3 through Modern Apprenticeships and similar schemes have completed their training.

The DfEE is currently considering the feasibility of setting up a large cohort study of 13-19 year olds, using personal interviews, which would overcome many of these limitations of the YCS. The expense of such a study might however mean that it would be limited to a 'one-off' inquiry rather than becoming a source of continuous information.
5.3 Longitudinal data on ethnic minorities

While it is known that members of certain ethnic minority groups experience various forms of disadvantage, less is known about the underlying processes. More evidence is needed to ascertain the relative importance of socio-economic, environmental, biological and cultural factors in contributing to ethnic minorities’ higher risk of adverse health and socio-economic outcomes. Longitudinal data on individuals’ life experiences are needed to build up a profile of the factors that may be involved at various life stages and transitions. Clarifying the complex trajectories leading to disadvantage may help to suggest where preventive measures might be targeted. Longitudinal data make it possible to look at the persistence of forms of disadvantage among ethnic minority groups, and to learn more about the cumulative effects of different types of disadvantage, which may form 'vicious circles' leading to social exclusion.

There are four special issues that need to be fully considered in the design of instruments for collecting longitudinal data on ethnic minority groups. These are diversity, mobility, spatial dispersion and communication.

Diversity

Recent research on ethnic minorities has shown clear differences in levels of disadvantage between different ethnic groups. The PSI 4th National Survey of Ethnic Minorities showed that, on most indicators, Pakistanis and Bangladeshis were the most disadvantaged, followed by Caribbeans, with Indians generally experiencing less disadvantage, compared with whites. This survey only compared outcomes for the major ethnic groups. Other work has shown that there are significant differences within these groups, for example, between Indians of different religious backgrounds. There are also other, smaller ethnic minorities, whose members might be expected to experience high levels of disadvantage. These include groups where a high proportion are recent refugees, such as Somalis. Putting all the groups together underestimates the extent of disadvantage experienced by certain ethnic minorities while overestimating that experienced by others. Current longitudinal data sources do not generally contain sufficiently large sample sizes to allow differentiation by ethnic group.

Another element of ethnic diversity that is particularly important for the design of longitudinal surveys concerns differences in the demographic profile of ethnic groups, arising from different times of arrival in Britain. For example, the first wave of Caribbean migration to Britain came in the late 1950s and 1960s, and most of the migrants were then young people of working age. Many are now in retirement or approaching retirement age, but the Caribbean group does not yet include as many very elderly people as does the white population. The children of these migrants would now be at an age where they have families of their own, in their 30s and 40s. Peaks and troughs in the age distribution of ethnic groups need careful consideration when birth or other types of cohort studies are planned, as these peaks and troughs come in different places for different ethnic groups.

Diversity within ethnic groups has a variety of sources. Surveys need to contain sufficiently large samples of each ethnic group to allow distinctions between the experience of men and women, between different age groups, and between recent migrants and established British citizens. Also, the experiences of first generation migrants from a particular ethnic group are unlikely to be similar to those of the second generation, who have been brought up in Britain. An alternative is for surveys to focus on key groups of particular interest. For example, the Scottish Twenty-07 longitudinal survey included a sample of 173 South Asians aged 30 to 40, for comparison with the 35 year old non-Asian cohort. A sample of 14 to 15 year old Asians and non-Asians was also taken from the same local area, making it possible to compare outcomes and assess inequalities between different age cohorts.
Mobility

Longitudinal data are particularly useful for studies of mobility, an issue of particular interest for those studying the experiences of ethnic minorities. For example, there is interest in the extent of downward occupational mobility experienced by migrants living and working in a new culture. Policy issues concerning racism and discrimination are also partly issues of mobility, and the ways in which members and institutions of the host society may block such mobility for minority groups. Ethnic minority individuals and households may experience high rates of 'cultural' mobility, as migrant groups become more integrated into the host society. Changes in language use, attitudes and cultural practices may contribute to changing socio-economic and health outcomes, for different ethnic groups. We do not currently have the data to compare the life histories of ethnic minority individuals, by ethnic group, age and time of migration.

Recent migrants present particular problems for survey designers. They may move frequently from one temporary domicile to another, making them difficult to trace, and they may be absent from administrative records or lack identifiers such as National Insurance numbers. These kinds of problems are likely to contribute to initial under-sampling and high sample attrition.

Spatial dispersion

Ethnic minority groups tend to be concentrated in particular geographical areas. This means that a nationally representative sample of individuals will not be representative of the ethnic minority population unless this is specifically designed. Samples of ethnic minorities will need to take heed of the geographical distribution of the minority populations they intend to cover.

Another possible result of the concentration of ethnic minorities in disadvantaged local areas is the confusion of ecological effects with the effects of minority membership. Recent research has suggested that health outcomes, for example, vary between local areas, when other variables are controlled. It is important that longitudinal studies should take account of location and of geographical mobility. Surveys comparing different ethnic groups living in the same area (like the Scottish Twenty-07 and 11-16 longitudinal studies comparing South Asians and non-Asians living in particular districts of Glasgow) may be useful in identifying spurious ecological effects.

Communication

The most obvious communication issue that arises when surveying ethnic minority groups is that of language. Many members of ethnic minorities will not speak English as a first language. If they are to be represented in surveys, it is important to translate the survey instruments into appropriate languages, and to recruit interviewers fluent in those languages.

Language difficulties probably contribute to the attrition of respondents from ethnic minority groups that leads to response bias in all of the current major longitudinal studies. However, they are clearly not the only cause, since Afro-Caribbeans, who mainly speak English as their first language, have consistently high rates of attrition. More attention needs to be paid to other communication issues that might contribute to low response among ethnic minority groups. It would be helpful to know, for example, whether members of ethnic minorities see survey questions as less relevant to them, and to look at the effects of using ethnically matched interviewers. It would also be helpful to know the extent to which including questions of particular relevance to ethnic minorities might help to raise response rates.

It would be helpful to seek the advice of successful survey designers from other countries, who have managed to achieve high minority retention rates in longitudinal surveys. An example is the USA’s National Longitudinal Survey of Youth (NLSY). This survey consists
of three separate random samples designed to represent the population of individuals born between 1957 and 1964. The second of these three samples was designed to oversample Hispanic, black and other economically disadvantaged youth. The NLSY has a high retention rate and high participation of ethnic minority groups.

**Types of longitudinal data for research on minorities**

We now consider how far the special requirements outlined above can be accommodated within various types of longitudinal data, and what particular difficulties occur in each case.

BHPS, the major general purpose longitudinal survey in Britain, is based on a random sample of the national population. It contains only small numbers of ethnic minorities. For example, the first wave respondents included 67 black Caribbeans and 52 people of Pakistani or Bangladeshi origin. Numbers in the ethnic minority groups have fallen in subsequent waves. This means that there are insufficient numbers to allow comparisons of individuals and households from the different minority groups. There are also insufficient numbers to carry out sub-group analysis within the minority groups.

Some issues which the BHPS covers well, such as changes in household composition and in economic circumstances, are of particular importance in the study of ethnic minority outcomes. For example, there are stereotypical assumptions about the effects of high rates of lone parenthood among younger Caribbean women, and about the effects of extended family organisation among Asian groups, that would merit exploration if sufficient evidence were available. It could be helpful to have an ethnic minority booster to BHPS to collect this type of evidence. However, it is not obvious how the spatial concentration of ethnic minorities could be reconciled with the structure of the general population sample already adopted there. A separate longitudinal survey with oversampling of ethnic minorities would probably be a sounder approach, as in the case of the USA's second NLSY study that was mentioned above.

We have also mentioned the Scottish Twenty-07 and 11-16 studies, which included booster samples of South Asians from two age cohorts. This type of longitudinal locality study can take advantage of the fact that ethnic minorities tend to be highly concentrated in a limited number of areas, such as parts of London and the West Midlands. It cannot however provide a complete picture, and it remains an open question whether the spatial concentration of minorities will diminish in the future.

Birth cohort data have the potential to provide a very full picture of influences from the earliest years, including biological, cultural and personality data. However sample numbers for ethnic minorities are small, especially in the earlier cohorts. In NCDS, immigrants born abroad in the same week as the original cohort were added to the 16-year old sweep, but high attrition rates meant that many were lost again. The BCS includes more ethnic minorities than NCDS but still not enough for disaggregation by ethnic group.

An ethnic minority booster to the proposed Millennium birth cohort would be particularly valuable for comparing the lifetime experiences of white and ethnic minority groups. It might also be helpful to include adequate samples of some of the particularly vulnerable recent migrant refugee groups, so that their progress could be traced. Difficulties in maintaining contact with families from these groups would need to be specifically addressed if the exercise was to be worthwhile.

Apart from birth cohorts, other types of cohort study are likely to illuminate the progress and the disadvantages of ethnic minority members passing through various life-stages and transitions. A young school-age cohort, e.g. of 7 year olds, would have the advantage of providing quickly available information on issues such as racial awareness among young people, and the experiences of young people from various ethnic minority groups. Again, the movement from education into the labour market is one of the key life transitions. Despite
their generally high participation rates in post-compulsory education, young people from ethnic minorities are at particular risk of social exclusion and disadvantage when they eventually enter the labour market. An ethnic minority booster on this topic, attached to the Youth Cohort Surveys, could provide the opportunity to study how members of different ethnic minority groups negotiate this transition. It would also be helpful to include questions particularly relevant to young people from ethnic minorities, such as experience of racial harassment or discrimination.

There is considerable policy interest in older workers and the transition to retirement. We do not know of any longitudinal sample allowing for special examination of the circumstances of older ethnic minority people, as they pass through these stages. Given the demographic profile of certain ethnic minorities in Britain, this might be a particularly opportune time to institute such a study.

Administrative data include records on large numbers of people from ethnic minorities. For example, a recent monthly extract from the Employment Service’s Labour Market System (LMS) recording people of working age but over 25 years, produced nearly 269,000 records, of which 4,500 were classified as Black Caribbean, 2091 Black African, 1737 Indian, 1577 Pakistani, 551 Bangladeshi and 426 Chinese. The potential of such data for analysis as well as for use as a sampling base is increasingly being recognised. However, there is a problem with recording ethnic group – for this particular extract, 35 per cent of the sample fell into the category “no ethnic group selected”. Non-selection of an ethnic group appeared to be more of a problem than clients’ refusal to specify an ethnic group, with fewer than 3 per cent categorised as “prefer not to say”. It would be worthwhile attempting to improve the recording of ethnic group data among those producing administrative records.
6 Extension and improvement of longitudinal data

In this section we discuss some of the ways in which policy needs could be better served through the enlargement or improvement of existing data, or through the addition of new sources of data. These suggestions are put forward to aid further discussion rather than as strongly argued recommendations. In several places, we point to the need for further investigation or review of particular issues.

6.1 General purpose longitudinal surveys

The policy value of a general-purpose longitudinal survey like the BHPS will depend in part on how many *ad hoc* inquiries it can satisfy. Policy interests frequently focus on minority groups, or unusual circumstances, especially those that are involved in disadvantage. The higher initial cost of a larger general-purpose survey should be considered against the larger number of inquiries which can then be accommodated, the savings in avoiding fresh survey costs or other one-off inquiries, and the savings in lead-time which may also be achieved. There is, in our view, a *prima facie* case for a substantially larger BHPS that would be able to meet a wider range of requirements.

We do not think it appropriate for us to suggest what the increased sample size should be nor how or when it should be introduced. Such decisions would require a far more detailed review and analysis than the present one. A relevant exercise would be to take a list of topics of current interest to government, specify a relevant sample, and tabulate the numbers of cases found under this sample definition during (say) waves 1-5 of the BHPS. One could then see how much practical difference would be made by a given increase in sample size.

To achieve an increased use of general-purpose surveys, there may be a need to raise awareness in government of the scope for secondary analysis to meet *ad hoc* requirements. Although the evidence is anecdotal, we feel that there is some tendency to initiate new surveys without giving full consideration to information in existing sources. This however applies not only to longitudinal data but also to national cross-sectional and repeat cross-section surveys. Documentation and access issues, which would help users become more aware of the possibilities, will be considered later in this section.

Another issue, which our earlier discussion highlighted, was that of additional sub-samples to improve coverage and representativeness. The BHPS has already been augmented by a subsidiary sample of low-income households\(^{15}\). This however appears a separate issue from that of overall sample size. Larger sample size increases the flexibility of a general-purpose survey, while additional sub-samples cater for specific requirements. The case for any additional sub-sample would need to be made separately.

Longitudinal surveys require especially high response rates. For the future development of longitudinal studies, we need fuller knowledge of why US surveys have done so well in this respect, and what could be done to improve response rates in Britain. Survey response is currently an area of lively professional interest and debate, and the ESRC/ONS review might seek to capitalise on this.

There is potential value in the QLFS in the role of a general-purpose longitudinal survey. The large sample size, and the quarterly interviews, offer something which is distinctive and complementary to BHPS. Setting aside purely technical issues, the great advance would be in collecting wage and non-wage income data at every interview instead of, as now, only twice for each person. Even if only a proportion of the sample were treated in this way, so as to make the change cost-neutral, the gain in information would be valuable.

\(^{15}\) This has been part-funded by Eurostat through ONS. However, data from the supplementary sample are not available as part of the BHPS dataset.
6.2 Long-term cohorts

There is already a government financial commitment, via the ESRC, to a Millennium birth cohort. The wide consultations already taking place should, hopefully, ensure that a correspondingly wide framework of household economic and social descriptors is gathered during the early period of the new birth cohort study. It is on this that the value of the study will depend, during later decades, in regard to many future policy issues that cannot now be foreseen.

The lack of a birth cohort study during the 1980s is a limitation on across-period comparisons. The ESRC and ONS have pointed to the possibility of an additional cohort study of children at school, with the aim of partially closing this gap. What age should this cohort be? The younger the cohort, the sooner will across-period comparisons become available, and the earlier into the educational process those comparisons will penetrate. This makes a case for a cohort from primary education (for example, aged 10 in 2000). But there is also current policy interest in the period around the completion of compulsory schooling. At present this period is studied chiefly through the YCS, a large-sample study which suffers from non-response. A higher-quality survey of this period would provide a long-term comparative resource for the proposed Millennium birth cohort study, and the DfEE is currently exploring the feasibility of following a new cohort through from age 13 to 19, using face-to-face interviews.

The advent of a new birth cohort study, and possibly a school cohort study as well, coupled with the ESRC’s support for extending the NCDS and BCS to 2003, should give a powerful stimulus to research. From a policy viewpoint, there may be an opportunity to harness these developments to learn more about long-term processes of social exclusion. This requires encouragement to the research community to make more use of the full time-spans offered by previous birth cohorts. There are also related issues of theory and methodology, in which ESRC could foster interest.

6.3 Special-purpose longitudinal studies

Special-purpose studies provide longitudinal data that cannot be supplied by general-purpose datasets. But ‘one-off’ longitudinal studies are costly. One should be sure that there is not an existing data source which meets the need, and one should give some thought to how special-purpose studies could themselves be exploited more.

Following up our earlier comments, we suggest that there could be some procedure to consider whether inquiries of a longitudinal nature can be satisfied from existing sources of data, such as the QLFS, BHPS, the birth cohort studies, or the repeated cross-section surveys. Making full use of the existing sources is a good way to ensure their continuing vitality and policy relevance, as well as offering potential cost savings by avoiding new surveys. Perhaps ONS could provide and promote a formalised advisory service to potential users that would help them to review existing sources.

Secondary use of single-purpose studies has been and continues to be at a low level. Government might get more out of its currently funded surveys and evaluation studies, if efforts could be made to get them documented, anonymised and placed in the public domain more rapidly. This needs supplementary funding. Our experience has been that government departments have seen the potential advantages, but in practice have found it irresistible to make cuts in something which has no immediate return, while researchers will always prefer to spend their own budgets on primary analysis than on preparing their data for others. Probably no progress will be made unless there is a separate fund to support this type of activity.
Where one-off studies are required, there might sometimes be scope for reducing the cost of longitudinal components. More consideration could be given to the use of telephone interviewing – perhaps not for initial interviews, but to add or extend the longitudinal dimension. This may not be effective in studies of financially deprived groups, with low levels of telephone ownership, but with telephone ownership reaching around 95 per cent of households overall, there must be many studies where it is an appropriate and cost-effective method for longitudinal follow-up.

In evaluation studies, perhaps the largest single advance would be to have substantially longer follow-up periods, so that long-term outcomes of policies and programmes could be assessed. Major progress depends upon being able to track outcomes through administrative databases, and link these outcomes to the primary sample data.

6.4 Administrative databases

We urge that a high priority is given to the development of longitudinal datasets from administrative sources, provided of course that legal and ethical requirements can be properly satisfied. The potential advantages in making more use of administrative data include large cost savings, large improvements in data quality, especially with regard to economic data, and (as noted above) longer periods of follow-up.

A thorough assessment of the possibilities is needed, resulting in clearer guidance to commissioning departments and external contractors about the availability and use of administrative data sources. Part of this might consist of a cross-departmental review of relevant administrative data sources, many of which may not be widely known. Some coverage of this ground may already be available from the recent DETR/Oxford University review of the index of local deprivation, which has considered numerous sources of data.

There may also be a need to consider some technical issues in administrative data, which could become important when the data are used for new purposes. This includes, for instance, the issue of incomplete data in the primary sources. It may also include issues of consistency across databases when information from more than one source is to be merged.

Some other countries, including Denmark, Sweden, and the USA, have placed much reliance on administrative data for economic and social research. A review of their practices could be helpful for future developments here.

6.5 Some additional points

An important issue for making longitudinal datasets accessible is, as already noted, documentation. Of course this is true of any survey, but the complexity of the longitudinal datasets imposes additional requirements. With rapidly advancing Internet technology, it should be possible to shift documentation to on-line sources, so as to make all information available through computer-aided search facilities. This has been achieved for the BHPS, and the providers of other major datasets should emulate the Website offered by ISER.

Whether there is a need for a high level of documentation for special-purpose, one-off longitudinal studies is more dubious. The high learning costs of using complex datasets usually deter all but the original researchers from attempting secondary analysis from these sources. The question is whether investment in documentation would significantly change matters. But, without trying, one cannot tell, since outsiders will certainly not use datasets that are undocumented. Additionally, it is important that datasets made available for secondary use should include derived variables as well as source data.
Documentation is clearly an important issue if administrative databases are to come into wider use for policy research. Such experience as we have suggests that existing documentation assumes insider knowledge of the system, and tends to be unsatisfactory for outside users. Extracts from administrative databases should ideally be supported by fresh documentation of the selected variables.

Another entry barrier to the larger longitudinal datasets is the complexity of the data structures in which they are stored. New users must invest time in understanding the data structures before they can make any practical progress. Some programming is then usually required in order to extract subsets of variables and re-structure them for a particular purpose. With the aid of technology, it should soon be possible to remove this barrier. In the next few years, PCs will have cheap hard disks with capacity much greater than the last generation of supercomputers. It will then become possible to ignore data storage costs and provide datasets as simple 'flat' structures. Accessing the dataset for a particular inquiry will simply be a matter of selecting variables. In addition, the rapidly advancing potential of communications through the Internet and Websites will make it possible for large databases to be accessed on-line (the 'Extranet' concept). There is a need to exploit these technical possibilities to the full, since they could greatly reduce the entry barriers to using existing datasets.
Summary and conclusions

In the past two or three decades, some of the more predictable or 'standard' features of people's lives have given way to a greater fluidity and diversity. Off-quoted examples are in working careers and in marital careers. In addition, there is either a real increase in, or at least an increased awareness of, the interlinked nature of social disadvantage. Changes such as these have created new issues for public policy. Policy concepts such as social exclusion, employability and joined-up government are striving to re-integrate the agenda. These are some of the underlying pressures that create new requirements for data that will track individual change over time.

The majority of government's requirements for monitoring social change continue to be met by repeated cross-sectional surveys. However, we have identified four types of need for longitudinal data that are distinct from cross-sectional data.

- Descriptive information relating to life-patterns
- Descriptive information relating to life-transitions
- Information required to explain and interpret social change or social problems
- Information required to evaluate policies or programmes.

The distinction between description and explanation is an important one. Only longitudinal data will provide the requisite descriptive information about the events and stages in individuals' lives and how these are linked. However both cross-sectional and longitudinal data have roles to play in explanation, and longitudinal data do not necessarily make the task of explanation easier. Such data are inherently complex and our ability to collect them is far in advance of our ability to analyse them rigorously and interpret them convincingly.

Turning to longitudinal data sources, we have distinguished four main types:

- General purpose longitudinal sample surveys
- Long-term cohort studies (notably birth cohorts)
- Special-purpose longitudinal sample surveys
- Administrative databases.

This division is largely a reflection of how longitudinal datasets have developed in Britain. The third category is rather miscellaneous in character, since it also includes 'one-off' longitudinal surveys and evaluation studies as well as shorter-term cohort studies. The recognition of administrative databases as a source of longitudinal data is particularly important. In some other countries, including Denmark, Sweden and the USA, administrative data play a major part in social and economic research and their potential as longitudinal data sources might be developed much further in Britain.

All sources of longitudinal data confront a range of methodological problems. Some of these - topic coverage, sample size and representativeness, sub-group coverage, and response rate - are shared with cross-sectional data sources, but are made more difficult because of the cost of repeated data collection. Additionally, longitudinal data have the particular problem of sample attrition over time; combating attrition is of crucial importance to those surveys that continue over many years. There is currently growing concern about survey response rates in Britain.

Administrative data sources (if available for research) provide very large numbers at low cost, and permit long-term tracking without attrition in the normal sense. However they cover too narrow a range of information to be regarded as a replacement for surveys. Rather, surveys and administrative data are to be seen as complementary. In addition, administrative databases have their own problems when used for research, such as incomplete recording or documentation which is not designed for external use, and these will need further work.
After reviewing the main sources of longitudinal data, we also briefly considered three particular applications for longitudinal data: income dynamics, youth transitions, and ethnic minority disadvantage. These are illustrative examples and are not meant to indicate the range of potential applications, which is evidently much wider. Collectively the examples illustrated the potential value of increased longitudinal data in a diversity of policy areas, but they also illustrated the need for more methodological improvement.

On the basis of our review, we have put forward a variety of suggestions for the extension, improvement and wider application of longitudinal data. This range of suggestions is not comprehensive, and it doubtless reflects bias towards our own areas of research, and ignorance of the needs of some other areas. Even so we are conscious that the list is over-long from a practical viewpoint. From the list of suggestions, we would select the following as probably the most important. (Note that the following list is not arranged in any order of priority.)

a) It should be possible to increase the use of general purpose longitudinal surveys such as the British Household Panel Survey for ad hoc policy inquiries. In our view the sample size of the BHPS (5000 households) is a serious constraint on its flexibility to meet ad hoc needs, which frequently focus on sub-groups. We hope that means will be found to enlarge the BHPS sample (or a successor sample) in future years.

b) Although the use of existing survey sources to meet the needs of ad hoc inquiries is inevitably limited by topic coverage, this is not necessarily a serious limitation. Surveys such as the BHPS, the Quarterly Labour Force Surveys (QLFS), and the birth cohort studies, contain wide ranging information including many core items which are important to most social and economic research. We suggest that an initiative might be taken to increase the awareness of these possibilities among potential users, and perhaps also to assist them to map their requirement into existing datasets.

c) The QLFS is well known as the leading general survey on the labour market, but its potential as a longitudinal survey, to study short-term transitions, has been little exploited. The QLFS could have increased value as a longitudinal dataset if questioning about wages and non-wage income was included at all waves, instead of just at the initial and final waves as at present. This would have great value for the study of income dynamics, even if only implemented for a proportion of the whole sample.

d) Longitudinal surveys require especially high response rates. While there have been some notable achievements in maintaining long-term samples, the overall picture regarding response rates and attrition is less impressive in Britain than in the USA. There is also currently a widespread concern among social researchers about the difficulty of achieving high response rates. This is an issue that could profitably be included in further review of longitudinal data development.

e) Important advantages for both cost and quality would be obtained through the development of longitudinal datasets from administrative sources, and their linking to survey data. In doing so, legal and ethical requirements must of course be met. Evaluation studies of public programmes, contributing to 'evidence based policy', would particularly benefit by the use of administrative data links to provide long-term tracking of post-programme outcomes.

f) Issues of documentation and accessibility, while appearing somewhat mundane, are crucial if more value is to be gained from existing longitudinal datasets. Here advancing information and communications technology is opening up valuable possibilities. Some mechanisms, including separate funding for on-line documentation and access, may need to be provided in order to exploit these technical possibilities more fully.
References


### Publications in the GSS Methodology Series

<table>
<thead>
<tr>
<th>Report</th>
<th>Title</th>
<th>Publisher</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dave Elliot. Software to weight and gross survey data with applications to the EC Household Panel and Family Expenditure Surveys.</td>
<td>Government Statistical Service</td>
<td>1997</td>
</tr>
<tr>
<td>6</td>
<td>Michael A Baxter. Interpolating annual data into monthly or quarterly data.</td>
<td>Government Statistical Service</td>
<td>1998</td>
</tr>
</tbody>
</table>

Available only from the Office for National Statistics – see page 2 for contact points. Price £5 unless otherwise stated. Copies are available free of charge to GSS members.