

Centre for Longitudinal Studies • Institute of Education University Of London

CLS Cohort Studies Working Paper 1

**THE DESIGN AND CONDUCT OF
THE 1999-2000 SURVEYS OF THE
NATIONAL CHILD DEVELOPMENT STUDY
AND THE 1970 BRITISH COHORT STUDY**

**John Bynner
Neville Butler
Elsa Ferri
Peter Shepherd
Kate Smith**

Contents

	page
Background	4
Conceptual framework	6
<i>Life course perspective</i>	5
<i>Data specification parameters</i>	7
Applying the parameters to the major life domains	11
<i>Lifelong learning</i>	12
<i>Relationships, parenting and housing</i>	15
<i>Employment and income</i>	17
<i>Health and health behaviour</i>	19
<i>Citizenship and values</i>	22
Survey preparation	25
<i>CAPI</i>	25
<i>National Centre for Social Research (NCSR)</i>	25
<i>Questionnaire development</i>	25
<i>Pre-piloting</i>	26
<i>Paper pilot</i>	27
<i>Programme pilot</i>	28
<i>Tracing</i>	29
<i>Fieldwork</i>	31
Appendix	34

Background

Britain is widely recognised as a world leader in the production of longitudinal research resources and their use in the analysis of developmental and life course processes. This Working Paper contains the first account of the design, development and conduct of a new round of data collection for two of Britain's three national longitudinal birth cohort studies, the National Child Development Study (1958 cohort) and the 1970 British Cohort Study (BCS70).

The National Child Development Study (NCDS) began as a Perinatal Mortality Survey to examine the social and obstetric factors associated with stillbirth and infant mortality among the 17,000+ babies born in Britain in the week 3-9 March 1958 (Butler and Bonham, 1963)¹. Follow-up surveys of the whole cohort were carried out at ages 7, 11, 16, 23, and, most recently, in 1991 at the age of 33 (eg Davie, Butler and Goldstein, 1972; Fogelman, 1983; Ferri, 1993)². At that time, a special study was also undertaken of the children of one third of the cohort members, including assessments of the behavioural and cognitive development of approximately 5000 children. There have also been surveys of sub-samples of the cohort, the most recent at age 37, when information was collected on the basic skills difficulties of a representative 10 per cent sample (Bynner and Parsons, 1997)³.

The 1970 British Cohort Study (BCS70) follows a similar pattern to NCDS, taking as its subjects all 17,000+ British births in the week 5-11 April 1970, and, in the initial survey, providing insight into contemporary patterns of obstetric and neonatal care (Chamberlain et al, 1973, 1975)⁴. Subsequently, full sample surveys took place at ages 5, 10, 16 and 26 (eg Butler, Golding and Howlett, 1986; Osborne, Butler and Morris, 1984; Bynner, Ferri and Shepherd, 1997)⁵. As in NCDS, sub-samples have been studied at various ages: for example, at age 21, paralleling the NCDS survey at 37, a 10 per cent representative sample was assessed for basic skills difficulties (Ekinsmyth and Bynner, 1994; Bynner and Steedman, 1995)⁶.

From their original focus on the circumstances and outcomes of birth, the two cohort studies broadened in scope to chart all aspects of the health, educational and social development of their subjects as they passed through childhood and adolescence. In later sweeps, the information collected covered their

¹ Butler N.R. and Bonham D.G. Perinatal Mortality. Edinburgh: Livingstone, 1963

² Davie R., Butler N.R. and Goldstein H. From birth to seven. London: Longman in assoc. with National Children's Bureau, 1972;

Fogelman, K.R. (ed.) Growing up in Great Britain: collected papers from the National Child Development Study. London: Macmillan, 1983

Ferri E. (ed.) Life at 33: the fifth follow-up of the National Child Development Study. London: National Children's Bureau and City University, 1993.

³ Bynner J. and Parsons S. Use it or lose it: the impact of time out of work on literacy and numeracy skills. London: Basic Skills Agency, 1997.

⁴ Chamberlain, R., Chamberlain, C, Howlett B. and Claireaux A. British births 1970. vol. 1: the first week of life. London: Heinemann, 1973.

Chamberlain, R., Chamberlain, C, Howlett B. and Claireaux A. British births 1970. vol. 2: the first week of life. London: Heinemann, 1975.

⁵ Butler NR, Golding J and Howlett BC. From birth to five: study of the health and behaviour of Britain's five-year-olds. Oxford: Pergamon, 1986.

Osborne AF, Butler NR and Morris AC. The social life of Britain's five-year-olds: a report of the Child Health and Education Study. London: Routledge and Kegan Paul, 1984

Bynner J, Ferri E and Shepherd P (eds.) Twenty-something in the 1990s: getting on, getting by, getting nowhere. Aldershot: Ashgate, 1997

⁶ Ekinsmyth C and Bynner J The basic skills of young adults. London: Basic Skills Agency, 1994

Bynner J and Steedman J. Difficulties with basic skills: findings from the 1970 British Cohort Study. London: Basic Skills Agency, 1995

transitions to adult life, including leaving full-time education, entering the labour market, setting up independent homes, forming partnerships and becoming parents.

The Centre for Longitudinal Studies (CLS) at the Institute of Education (formerly the Social Statistics Research Unit located at City University) has been responsible for the National Child Development Study since 1985, when the study was transferred from its previous home at the National Children's Bureau. Likewise, CLS has housed the 1970 British Cohort Study since its relocation from Bristol University in 1991. A Forward Plan for the cohort studies was developed by the Director of CLS, Professor John Bynner, which sought to integrate the timing, design and analysis of future surveys of NCDS and BCS70 – taking account of the sequencing of Britain's third birth cohort study, the 1946 cohort (National Survey of Health and Development), housed at University College London. Such a programme would significantly enhance the research potential of the studies, enabling comparisons to be made between cohorts born at different times, or between different age groups at the same point in time.

Endorsement of the principles of the Forward Plan by the Economic and Social Research Council, and the Government departments, which had provided financial support for previous sweeps, resulted in an ESRC decision to fund the new surveys. Initially, this was restricted to BCS70 because of the need to update the dataset after 13 years without a comprehensive survey, but, subsequently, funding extended to a new survey in NCDS as well. ESRC contributed half the costs, and the rest came from Government departments under the coordination of the Office of National Statistics: Department for Education and Employment (the major funder), Department of Health, Department of Social Security, the Home Office, ONS themselves, the Scottish Office and the Basic Skills Agency.

The first stage in designing the new surveys was to locate the key issues and questions salient to the stage in adult life that the cohort members in each study had reached. These were formulated in terms of a 'life course' theoretical framework that focused on the factors central to the formation and maintenance of adult identity in each of the following main life domains:

lifelong learning

relationships, parenting and housing

employment and income

health and health behaviour

citizenship and values.

The next section develops this idea in general terms, while the subsequent section applies it in each of the above domains, arriving at the critical questions to be asked and the priority to be attached to them. Following that, an account is given of the actual development of the survey instruments, as well as the tracing and fieldwork procedures adopted in conducting the new surveys.

CONCEPTUAL FRAMEWORK

Life course perspective

Adult identity, comprising an individual's pre-dispositions, attitudes, behaviour and statuses, needs to be viewed both as a product of past life-course processes and a precursor of later ones. The 'life course' perspective enables us to take on board these different elements of human development, while offering a means of integrating them within a single inter-disciplinary framework.⁷ Life courses represent the transitions occurring in the different domains of life through which individuals make adjustments to changes in their social environment and through which their statuses in education, employment, relationships, family life and health are maintained and developed.

Characterising the long-term processes involved is the idea of 'accumulated risk', with its counterpart the acquisition of 'protective resources'⁸. Accumulated risk may lead to marginalisation in relation to employment, community, family life and citizenship, the main consequence of which is *social exclusion*.⁹ Adverse outcomes in adulthood, such as those concerned with education problems or ill health, can frequently be traced back to disadvantaged conditions earlier on. Such conditions may accompany birth and the early years of life, but they may also arise from later experiences and personal choices.

The lifecourse process is in part linear, but also to a degree cyclical, in the sense that experiences later in life can affect achievements earlier on. For example, unemployment can lead to a loss of cognitive skill among those whose cognitive skills were poor to begin with, thus creating a vicious circle.¹⁰

Another critical feature of the life course perspective is the *interconnectedness* of experiences and developments in these different spheres. To understand the transitions in one domain, eg employment, we need to take account of the transitions happening in other domains, eg the family. In other words, the life course is viewed *holistically* as the product of sets of interactions between biological predisposition, personal agency, and the social environment.

Personal agency is both formed by, and serves to form, adult identity. It comprises sets of resources and pre-dispositions. The *resources* reside in human capital (eg skills and qualifications), social capital (eg family and community resources), and cultural capital (eg beliefs and understandings relevant to high status occupations). Such resources are mediated in childhood through the education system, and this continues to play a part in adult life, though much learning then takes place outside formal settings. *Dispositions* comprise psychological state (eg malaise), control beliefs, including a sense of self-efficacy and internal locus of control, motivation, aspiration and self-esteem. Though dispositions are less directly targeted by education, they nevertheless are increasingly seen as a vital product of it, especially the acquisition of self confidence and self esteem, which are key protective factors in the transition from school to work and, subsequently, in gaining and maintaining employment.

⁷ Elder G.H. Children of the Great Depression: social change in life experience. University of Chicago Press, 1974; Heinz, W.R. (ed) Theoretical Advances in Life Course Research. Weinheim: Deutscher Studien Verlag, 1991.

⁸ Robins, L. and Rutter, M. (eds.) Straight and Devious Pathways from Childhood to Adulthood. Cambridge University Press, 1990.

⁹ Atkinson, A. B. and Hills, J. (eds.) Exclusion, Employment and Opportunity. CASEpaper4, London School of Economics, 1998.

¹⁰ Bynner, J and Parsons, S Use it or Lose It: The Impact of Time Out of Work on Literacy and Numeracy Skills, London. Basic Skills Agency, 1998.

The *social environment* comprises socially and culturally defined structural categories (such as gender and social class), as well as institutions (such as the family, school, labour market), and the social relationships that operate within them. A set of resources intimately connected with these relationships is *social capital*, which is increasingly seen as fundamental to the generation of human capital, as well as being a worthwhile end in itself, in terms of social cohesion. Social capital characterises communities with high levels of network membership, mutual trust and shared norms and values. Exposure in varying degrees to these manifestations of social capital will improve adults' own prospects and quality of life, and those of their children, but it will also foster their contribution to social capital. This emphasises the collective aspects of agency to set alongside those concerned with individual development and action. It also stresses the dynamics of social life. People are not only affected by the social environment, but are actively engaged in its development.

Comparing different cohorts, life course processes may also display changes in the relative balance between structural and personal agency effects. *Individualisation* is seen to arise through the transformation of employment brought about by new technology and the processes of globalisation.¹¹ Complexity of the modern life course, as far as employment is concerned, has meant that relatively straightforward routes into particular occupational niches have given way to the need for a portfolio of skills and occupational experiences, in which the very idea of a single 'occupational identity' breaks down. In consequence, the social norms and values, acquired through family and work, which structured social action in the past, have steadily weakened. Life courses become increasingly individualised as people have to rely more on their own resources, firstly to navigate their way into the labour market and subsequently to find a niche or series of niches in it.

Individualism impacts not only on employment but on family life as well. The traditional pattern of leaving the parental home to get married and set up a home and family of one's own gives way to an extended period of different kinds of living arrangements, the 'single lifestyle' and cohabitation, with postponement or rejection of the long term commitments involved in marriage and parenting. The institution of marriage is similarly less stable than it was, with around 40% of first marriages ending in divorce.

These changes are not, of course, uniform across society, showing increasing prevalence across cohorts and widening differences between social groups – *polarisation*.¹² This is why comparisons between BCS70 and NCDS are of such scientific interest. In a world where qualifications and a portfolio of relevant work experiences are critical to occupational achievement, those without these attributes get increasingly left behind. Leaving school at the minimum age, without qualifications, sets the young person off on a risk trajectory, where opportunities for secure employment and career development are limited. Casual, often part-time, and unskilled work experience, interspersed with unemployment, is part of the process of marginalisation brought about by the demand for qualifications and a steady decline in unskilled work. The consequence is seen not only in patchy employment careers but also through effects on health and family life. Early childbearing on the part of women, low self esteem and high levels of depression typically accompany the social exclusion process. High levels of smoking, drug and alcohol problems, and crime are its counterparts.

Health-related behaviours (smoking, drinking and drug-taking, for example) and healthy and unhealthy lifestyles (diet, exercise and so on) underpin morbidity and mortality later in life. The choices involved reflect *personal agency*: the individual may choose to take the risk of ill health. At the same time, health

¹¹ Beck, U. Risk Society. London: Sage, 1992.

¹² Bynner, J., Ferri, E. and Shepherd, P. Twenty-something in the 1990s: Getting on, Getting by, Getting nowhere. Aldershot: Ashgate, 1997

status itself has an effect on the life course in other domains besides health - for example, employment. This is obviously the case in relation to 'limiting life-time conditions', including various kinds of physical disability, but it also applies to the more specific childhood hazards associated with accidents and illness, and the debilitating conditions of later life such as respiratory illness. Both these components of well-being - health-related behaviours and health status - are distributed unevenly across the population, showing marked gradients by social class, gender, and locality - *health inequalities*.¹³

Finally, there are both *objective* and *subjective* features of the life course which need to be embraced. For example, in relation to the impact of social change, such as a rise in unemployment, or a new social policy initiative, how the individual appraises the situation may be just as important as the objective features of the damage or improvement the change brings about. Clearly, the reduction in income produced by unemployment may impact in a direct sense on life quality, but also important is the subjective evaluation of this change, and how the individual internalises it. A key part of this subjective component is *control beliefs*, which are the extent to which individuals believe they are in a position to control events, rather than be controlled by them.¹⁴

Data specification parameters

The main focus of the new surveys is to elucidate the latest stage of the life course of cohort members, revealing the pathways taken to current statuses, the influences to which they had been subjected and their outcomes in terms of current identities. Survey data in these areas provide bankable information for modelling the future, as well as for explaining the present in terms of the past. The life course perspective outlined above provided a set of parameters for the central design task of the new surveys. This was to identify the key variables which best represent the different elements of the life course process in each life domain and across life domains, and to find the optimum way of operationalising these through survey questions and other measures. The parameters included:

Structural categories: gender, social class, religion, and geographical location.

Institutions: family, education system, employment system, trades unions, political system, health and welfare services.

Social relations: assessed in terms of the embeddedness of individuals in social networks, in education, at work, in the family and at leisure.

Social roles, statuses and values: these components of individual *identity* embrace roles such as parent, spouse, employee and community member. They also include position in the social hierarchy, as identified with occupation/social class, or in the family - provider, carer etc. They underpin, and are underpinned by, social and political attitudes and values, including gender values, and membership of associations including Trade Unions, churches, local community groups and political parties.

Health and lifestyle: Distinct from, but connected to, the range of more socially designated statuses are those to do with health and physical development and functioning. These both impact on, and arise from, experience in the labour market and need to be incorporated. They comprise the record of mental

¹³ Power, C. Manor, O. and Fox, A.J. *Health and Class: the Early Years*. London: Chapman and Hall, 1991.

¹⁴ Heinz, W. R. (1999) *From Education to Work: Cross National Perspectives*, Cambridge University Press. Diewald, M. *Continuities and Breaks in Occupational Careers and Subjective Control: The Case of the East German Transformation*. In Bynner, J and Silbereisen, R. K.(eds.) *Adversity and Challenge in Life in the New Germany and England*, Basingstoke: Macmillan, 2000.

and physical illnesses, symptoms, accidents and the treatments received for them and their outcomes. They also bring in the topic of 'lifestyles', in which smoking, alcohol and drug use, diet and exercise are all manifestations with health-related aspects.

Continuity and comparability

By 1999, when fieldwork was to begin, the members of both cohorts were well into adulthood, with BCS70 subjects aged 29 and their NCDS counterparts having reached 41. This pointed to a high degree of commonality between the two cohorts in the topics to be covered and the kinds of questions to be asked. Common questioning was, of course, also an important consideration in relation to the potential of the new data for inter-cohort comparisons.

On the other hand, each study had specialised somewhat in the past, focusing to a certain extent on different features of childhood and adolescence, and adopting its own distinctive methods of data collection in earlier sweeps (for example, diaries and medical measurements in BCS70, parental and school data and much information on educational performance in NCDS). To realise the cohort studies' value for analytic purposes, variables need, whenever possible, to be repeated across surveys. Information that serves primarily the cross-sectional purpose of estimating current population parameters must take second place to that which captures the dynamics of developmental processes observed across time. Event histories play an important role in charting transitions in the key domains. From the point of view of internal consistency and continuity, therefore, there was also a case for maintaining the distinction between the studies' data specification requirements. What was needed was a balance between replication of variables *across* the two cohorts and repetition of variables *within* each study from one survey to the next.

Age, cohort and period effects

The data also needed to embrace the common elements of all life histories, such as those determined by biologically-based developmental processes (*age effects*) and their variability in terms of timing and duration. It was equally important to encompass those life course elements that are culturally determined, and how these differed between different socially defined sub-strata – including socially excluded groups. Such life course elements are subject to two other types of external influence. The effects of social change are embraced by comparisons between different cohorts – *cohort effects*. Societal influences at the time of data collection, which are held constant for any given survey, reflect *period effects*. Taken together with geographical location, these constitute the temporal and spatial context (time-space co-ordinates) of our data collection.

Space-time co-ordinates

These comprise information about the year of study (1999-2000), and the geographical location of each individual cohort member. In relation to time of study (period), we needed to consider external events that could influence cohort members, eg Government legislation, health scares and so on. In relation to space, it was necessary to consider external pressures on individuals which are regionally based, eg reductions in employment, and legislation with differential effects across regions, such as devolution. Thus, we needed to take on board distinctions to do with rural/urban, North/South, Scotland/England/Wales and so on. With respect to age, we needed to separate out the developmental components that are biologically based (eg physical ageing) from those that are more socially determined (eg age at marriage, or at birth of first child). We may nevertheless see the former as defining certain boundaries for, or driving, the latter, eg the age-defined period of the life course when having children is possible.

Harmonisation

A final consideration in setting out the principles guiding the design of the new surveys was the move begun in the 1990s towards the harmonisation or standardisation of the information collected for Government social surveys. In developing the new instrumentation, therefore, reference was also made to ways in which questions were formulated in surveys such as the British Household Panel Study and the General Household Survey, as well as, for maximising the opportunities for inter-cohort comparisons, the National Survey of Health and Development (1946 cohort).

APPLYING THE PARAMETERS TO THE MAJOR LIFE DOMAINS

The above parameters provided guidelines for examining the relevance of possible questions or measurements to each of the life domains, and formed the basis of a grid for filtering out the salient from the irrelevant questions. This was the starting point for a lengthy, systematic process to develop the new data collection instruments; a process which involved not only the Cohort Studies team in CLS, but drew in contributions from a large number of experienced users of the data from the wider research community.

Collaborative network of researchers

Close consultation between the CLS team responsible for the cohort studies and their users and beneficiaries has been a hallmark of each stage of the studies' history. This process has been maintained and enhanced as NCDS and BCS70 move forward in tandem, both as an essential part of the development of the new surveys, and to set in place a structure for consultation over the fullest possible exploitation of the new datasets, and the coherent planning of subsequent sweeps.

A consultative conference was held at the Institute of Education on the 26th March 1998, to engage the experience and expertise of the research and policy communities in designing the questionnaire content for the new surveys. Seven advisory groups were formed, one for each of the major topic areas to be covered.

NCDS/BCS70 Advisory groups

lifelong learning

employment and income

family, parenting and housing

health

citizenship and values

child development and education

methodology

(The *child development and education* group was formed in the hope of obtaining sufficient additional funding to carry out the first follow-up of the children of a one in three sample of the NCDS cohort (first assessed in 1991), and the first survey of a similar sample of the children of the 1970 cohort. Unfortunately, such funding was not forthcoming in time for this work to be included in the new adult surveys, although it remains a key objective in the Forward Plan for the future of the cohort studies.)

Each group appointed a co-ordinator and was served by a member of the CLS team, who facilitated liaison among its members. Following the initial meeting, group members exchanged ideas and information via email and/or meetings. Details of group membership, notice of planned meetings, and reports on the activities of each group were posted on a special website.

Given the extensive coverage of the new surveys, and the wide-ranging interests represented within each group, it was inevitable that the volume of suggested questions far exceeded the available questionnaire space. Consequently, group members were asked to assign priority levels to the proposed questions, on the basis of the criteria contained within the conceptual framework for the surveys, namely theoretical importance, continuity with previous sweeps, and relevance to time, context, and age of the cohort members.

Written advice on the content of the surveys was provided to the CLS team by each advisory group between March and June 1998. These reports were collated by CLS and circulated for information to all members of the network. The advice received was invaluable to the CLS team in developing the new instrumentation. Copies of the draft questionnaires produced at the next stage were also circulated to all advisors before versions for piloting were finalised.

The following sections present the result of applying the life course perspective, and the developmental process outlined above, to each of the major life domains. The main topics and themes included in the survey are outlined, as is the rationale for their inclusion. A document containing all the specific questions included in the survey is available as an appendix to this Working Paper. A summary of question content is supplied in the appendix to this paper.

LIFELONG LEARNING

Learning Goals

All learning can serve two kinds of functions. In economic terms, it serves production as a form of human capital accumulation, with direct benefits both for the individual and for society. It also serves consumption as an enjoyable and fulfilling activity in its own right. In discussions of the 'Learning Society' these two functions typically become confused. However, they define the different types of goals to which learning is directed and the new surveys embrace both of them.

Life stage and context

The 1958 cohort reached 16 at a time when around two thirds of young people left school at that age to get a job – ideally with an apprenticeship in a skilled trade attached, or with high quality training and clear career progress in a white collar job. Of those staying on, only about one in eight went on to higher education. Unemployment was not experienced until the slump of the 1980s put many cohort members out of work and forced them to re-train, often through government-supported schemes like TOPS. However, the cohort had all experienced paid work, which in employers' eyes was an important attribute of employability.

By the time the 1970 cohort reached 16, half were still leaving school to seek work, but substantial proportions were compelled to enter Government training schemes, YTS, or become unemployed. Apprenticeships had largely disappeared. Their careers were likely to be more chequered than those of the 1958 cohort, with training, qualifications and work experience becoming at an increasing premium in getting employment. Marginalisation to the periphery of the labour market and unemployment were increasingly the fate of those without them. The proportion continuing to higher education - one in four – was double that in the 1958 cohort.

We needed to reflect these differences in eliciting cohort members' learning careers. By age 41, the 1958 cohort's interest in learning would be more directed towards updating and leisure pursuits. At 29, sizeable proportions of the 1970 cohort would still be engaged in human capital accumulation through gaining qualifications. Those on the margins, including single mothers on benefit, would be the target of the Government's 'New Deal' packages of training and supported employment. Both cohorts were likely to be benefiting from the Government's response to the consultations on the White Paper on life-long learning - *The Learning Age*, intended to remove perceived obstacles to learning, increase participation and facilitate continued learning throughout life.

Obstacles, opportunities and outcomes

Those who have had the most education and gained the most from it in terms of qualifications continue to participate most and to get most out of it. Obstacles in the way of the others begin in school with failure to acquire the basic skills and the learning skills that build upon them. This is likely to be followed by lack of exam passes, reduced motivation and limited opportunities to pursue further studies. Family and child care responsibilities, the cost of courses or, when at work, simply the lack of training and encouragement to train, are likely to compound the difficulties the potential learner faces. The *Access* movement, which began in the late 1980s, and the *Open College Networks*, which grew out of them, offered a means of removing these obstacles, as well as helping people to make the best choices among the learning opportunities available. On the other hand, people who fail to engage with the formal educational system may acquire various kinds of experience and skills through work or leisure. Data collection, therefore, needed to address both the widest range of obstacles to learning and the widest range of opportunities taken to do it in formal and informal settings. It also needed to address what the learner gained from his or her learning experiences. This goes beyond the more easily recognised outcomes such as the qualifications gained, to embrace subjective appraisals of skills and other gains acquired.

Previous approach to data collection

Previous sweeps, of NCDS particularly, focused on the production side of learning, questioning cohort members about the qualifications they achieved at school, subsequent courses leading to qualifications and work-based training courses. Learning done under what was called in the past ‘non-vocational’ adult education received relatively little attention and informal learning or self-directed learning through books, television and computers, was not examined at all.

In relation to the learning record itself, there were two kinds of problems. First, the years in which qualifications were obtained were not routinely recorded – only what qualifications had been achieved by the time of each survey. Second, in order to restrict the amount of data collected in NCDS5, only the courses leading to the *two highest* qualifications were covered in any detail, and in the case of work-related training courses (taking *three days or more*), only the *three most recent* courses. It was clear from other indicators that this approach seriously under-represented the amount of formal learning that had been going on.

Continuity from NCDS4 to NCDS5 was weakened by the fact that in NCDS4, the major emphasis was on different forms of employment, as opposed to staying on in education. Apprenticeship was a major focus. Unemployment was virtually unknown, and youth training had not begun. This framework extended beyond the period of its relevance into the early 20s, when work-related training, for example, would have been highly relevant, but was not studied in a way comparable to its subsequent study in NCDS5. In fact, the only work-related training courses examined were those lasting 14 days or more. This time, questions were asked on the order in which they occurred after leaving school: first, second and so on.

BCS70 suffered from the problem that the basic design of the survey at 16 assumed that cohort members were still at school. Even though the data collection extended into a period when many had left, this was not covered in any detail, especially the youth training that many would have entered in 1986. In the 26-year postal survey, with limits on questionnaire space, only statuses could be elicited, eg qualifications achieved and retrospective counts of the amount of training and number of courses leading

to qualifications that respondents had taken. This meant that there was a gap in the learning record to be made up.

Data collection strategy

Learning Record

It would clearly have been impossible, especially for NCDS, to collect a complete record of all learning experiences since 16. As an alternative, following the well-developed approach in the *National Adult Learning Survey*, we could get a record of qualifications achieved, including type of qualification, subject, date and provider from age 16 upwards. We also recorded in more detail, over the period since the preceding survey, all major *learning episodes* – formal and informal. For each major episode identified by the respondent, we established where it took place, the learning medium (eg formal/self-directed/computer/television), how long it lasted, what was achieved, who paid for it and the cohort member's subjective appraisal. For BCS70, learning episodes included YTS schemes.

Skills acquired

We pioneered through NCDS5, with replications in later surveys, the self-reported 'skills inventory'. This recorded whether the person thought they had each of a number of skills, where it was applied and whether it had improved or deteriorated over the last 10 years. This approach was repeated, drawing on the more developed list of work-related skills used in the *Skills Survey* carried out through the Centre for Labour Market Research at the University of Leicester.

Basic Skills

The new surveys continued with detailed questioning on cohort members' basic skills problems, including where the problems arose, and what they had tried to do about them through remedial courses etc. In addition, basic skills problems in relation to helping children learn to read and with their subsequent school work were also included. Perceptions of different kinds of learning provision and motivation to engage generally with remedial education were also covered.

Attitudes to learning

Two types of attitude questions asked in the Adult Learning Survey were drawn on for the new surveys. The first asked people to describe themselves in terms of their situation, ability and motivation in relation to learning, and obstacles to their learning. The second was of the more standard 'Likert' type, comprising a number of statements about employers' attitudes to learning, the value of qualifications, interest in learning and its value etc., to each of which the respondent indicated agreement or disagreement on a five-point scale.

Benefits and Constraints

Finally, we needed to get an overall appraisal of the learning experience to date: what benefits the cohort members felt they had obtained overall from the major learning episodes they had had, and what, if any, disadvantages had accrued from them. We also wanted to get an overall picture of the sorts of barriers they encountered in undertaking learning, and how they overcame them.

RELATIONSHIPS, PARENTING AND HOUSING

The changing demography of family life

By the time of the new surveys, the NCDS cohort at age 41 would (females at least) be approaching the completion of their fertility, and the BCS70 cohort at 29 would be continuing their transitions to partnerships and parenthood. Both cohorts, but particularly the younger group, will have made these transitions during a period when this particular life domain has been characterised by change and instability. Among the key demographic trends which will have been influential are the increase in cohabitation, the postponement or even rejection of marriage, the later average age at first birth (although with an increase in the number of births to very young mothers) and a continuing rise in the number of births outside marriage – currently around one in three, although over half of these are registered jointly by cohabiting couples. Alongside these trends is the increasing fragility of partnerships and family life: in the most recent surveys of both cohorts (at 33 for NCDS and 26 for BCS70), sizeable proportions had already experienced the breakdown of a cohabitation or marriage, bringing up children single-handed or with new partners, or missing out on active parenting altogether through loss of contact or co-residence with their offspring.

The cohort members as parents

Both cohorts will have made the transition to parenthood during a period of increasing social concern about parents' 'performance' and changing social policy expectations regarding parental responsibility (eg withdrawal of benefit for under 25's; parental fines for misdemeanours, proposals for 'parenting contracts' etc). An important question for the new surveys, therefore, concerned how the cohort members, as parents, pass on to their children the product of their own life course experiences to date. There is already evidence from NCDS that there is an *accumulated risk* of disadvantaging factors in the cohort members' own families of origin – eg parental divorce, family disruption, which is linked to adverse adult outcomes such as early partnership and parenthood, relationship breakdown and low educational and occupational achievement.

Further questions arise about how these earlier experiences influence the ways in which the cohort members themselves raise their own children, and how these children are, in turn, affected by such antecedents. It was considered particularly important to collect information about how the cohort members undertake their parenting tasks. This needed to cover both objective and subjective aspects, and include data on their attitudes, values and behaviour in respect of their own and their partner's parenting roles. Such information would be essential in investigating the contrasting theoretical views that individuals internalise and reproduce the model of parenting experienced in their family of origin, or that many people consciously reject this model and adopt opposite approaches themselves.

The importance of fathers

It was also important that the data collected on parenting behaviour and attitudes pertain to *fathers* as well as *mothers*. This is critical to an analysis and understanding of changing gender roles in the family, as well as correcting the almost universal bias in previous research in this area, including earlier sweeps of the cohort studies themselves. In NCDS5 for, example, information about family life and parenting was collected from female cohort members or the wives/partners of cohort fathers.

Cohort members and their own parents

For the NCDS cohort, it was necessary to have a clear idea of what a longitudinal study of individuals in early middle age should be about. For most, their 'social destinations' were arrived at some time ago and would probably not have changed much. In mid-life, it is appropriate to focus on family life from the standpoint of cohort members' identities as both *parents* and *children*. Intergenerational links were seen as a key area of enquiry, both from the point of view of salient social issues at this stage in the life course, and of what was now a three-generational longitudinal dataset. Recent sweeps of the 1946 study (NHSD) included questions on cohort members' contact with their own parents, and feelings and concerns about them. Replicating such questions with NCDS would permit comparisons of different cohorts' current and future relationships with their parents, including their views on their roles and responsibilities in relation to elderly parents and their care needs.

For BCS70, similar questions should be asked about contact with own parents, but for them, the focus should be on their parents as sources of *support* to them in their transitions to independent living, parenthood and so on.

Data collection strategy

Stability and change in cohort members' relationships and family lives

The new surveys continue to chart the relationship and family formation histories of the members of each cohort, through event history data covering partnerships (including cohabitations and marriages, separations and divorces), pregnancies and births (as well as infertility and voluntary childlessness). As in previous sweeps, details are also obtained of everyone living in the cohort members' current household and their relationship to the cohort member. This contributes to the picture of second and subsequent families, which is completed by information about cohort members' (and partners') children living elsewhere.

The subjective dimension to partnership and marriage also needed to be covered, with information about the cohort members' feelings about their current relationships. Details were also required about the circumstances of the breakdown of earlier partnerships and the nature of any subsequent contact, especially with the other parent of cohort members' children.

Cohort members as parents

Although much in-depth questioning about parenting behaviour and experiences would have been addressed to those in the one in three sample had this sub-study taken place, it was important to obtain some information in this area at what was for both cohorts a key stage in the life course. The respective roles of mothers and fathers in terms of responsibilities for work and child care is an important issue in contemporary family life. Other salient topics included their involvement with their children's schools and helping with their education, and information about family life such as joint activities. From the subjective viewpoint, the cohort members' attitudes towards having and raising children, and their views and concerns about their children's needs and development, formed another relevant dimension.

It was also important to obtain details about periods in which the cohort member had been caring for children as a lone parent, plus information about contact between children and absent parents, and whether or not financial support was received.

Cohort members and their own parents

The 1946 cohort study (NSHD) provided a valuable source of questions covering the topics of contact and relationships with own parents and support *to* ageing parents among the NCDS cohort. For the 1970 cohort, further questions were developed to examine ways in which the younger group had in the recent past continued to receive support *from* their parents.

Social networks and emotional support

As well as investigating living arrangements and personal relationships, it was necessary to widen this area of enquiry to look at the cohort members' wider social networks. This required questioning about social activities with friends and about sources of emotional support.

Housing

Each follow-up survey of both NCDS and BCS70 has collected detailed information about the cohort members' immediate material environment in terms of the type and tenure of their current accommodation. The new surveys also obtain subjective information about satisfaction with present home, area and transport, and any plans to move. In line with earlier sweeps, housing histories, covering all addresses lived at since the previous survey, are also updated, including details of any periods of homelessness among the 1970 cohort.

EMPLOYMENT AND INCOME

The changing labour market

Perhaps more than any other area of investigation, the domain of employment has been profoundly affected by social change in recent years, most notably in the form of labour market transformations. These include deregulation and the introduction of flexible work patterns, the decline of manufacturing and the rise of service industries -with their respective effects on male and female employment, the increasing influence of the global economy, the emasculation of trade unions and the decline of collective bargaining. Among the most striking trends in the area of employment has been the increase in the number of women, especially mothers, in the labour force.

Current and past experience in the labour market is a key domain in the life course perspective shaping the long-term follow-up of both the NCDS and BCS70 cohorts. For the 1970 cohort, it was clear that the new survey would need to chart the pathways into employment from the time the cohort members left full-time education. For this cohort, who were 16 in 1986, the timing and nature of the transition to employment will have been very different from that of the 1958 cohort. As noted in the earlier section on Lifelong Learning, the labour market which confronted those born in 1970 was characterised by increasing insecurity and short term contracts. Their employment experiences are also likely to have been more directly influenced than those of the 1958 cohort by new policy initiatives to support employment and reduce welfare dependency (eg Welfare to Work). Analysis of postal survey data on current employment in 1996 indicated considerable polarisation between those whose fund of human capital, acquired through education, was associated with high incomes and high status occupations, and those operating on the margins of the economy. It was thus essential to chart the labour market histories of the 1970 cohort to relate their employment transitions to other life course experiences such as education, training and family formation.

For the NCDS cohort, new data would be required to update their labour market experiences since they were last surveyed in 1991. The 30s is a period in the life cycle when career progress is likely to be consolidated, but also a time when the dependency of young families, and the income required to sustain living standards, are likely to be maximised. An important issue for the new survey, therefore, was how these factors interlinked in the context of the 1990s labour market. In both cohorts, event history data would be needed to document stability and change in employment and unemployment, and enable these measures to be linked to changes in other life domains such as relationships and family, housing, physical and mental health, and the acquisition of skills.

In addition to new data on employment, it was also important to include questions about income. Information on current household income, from employment and other sources such as welfare benefits, is a key indicator of material wellbeing and quality of life and has been a routine element of data collection in all adult surveys of the NCDS cohort. Details of income from employment were also obtained from the 1996 postal survey of the 1970 cohort.

Data collection strategy

Job histories

At NCDS5, this was obtained from the cohort member 'Your Life' self-completion questionnaire, with some additional information about each job collected during the Cohort Member Interview. In the new surveys, all information is collected from the cohort member during the interview, which should result in improved quality data. Details are obtained of the dates of starting and ending each job since 1991 for the NCDS cohort, and 1986 for the members of BCS70 (ie up to 13 years). In addition, enough information is recorded about the nature of each job for social class of occupation to be coded.

In the 1991 NCDS survey, employment event histories relating to cohort members' current partners were also included, with information obtained from the partners themselves via a self-completion questionnaire. This data was not of high quality, however, and for the new surveys it was decided to ask only for details of partners' current jobs, and to collect this from the cohort members during the interview.

For both cohorts, event history data is also obtained covering periods of unemployment or other spells out of the labour force (for NCDS since 1991, for BCS70 since leaving full-time education).

Occupational status

Occupation needs to be coded in such a way that continuity or mobility within and across generations can be identified. The most up to date occupational coding scheme that is still in use in most survey organisations is that produced by the Office for National Statistics based on the 1991 census. However, in the intervening period, there has been a major review of occupational classifications, especially in relation to social class. A new version will shortly become available. In the time period of the current survey, which involves coding during the course of fieldwork itself, it was not possible to apply the new classification. However, the raw data obtained in the interview about cohort members' occupations will be available for researchers to apply the new classification once it is published later this year.¹⁵

¹⁵ Rose, D. (1995). ESRC Review of OPCS Social Classifications. London: Office of Population Censuses and Surveys.

Other aspects of current/most recent job

In addition to the details of current job which were needed for occupational coding (repeat measures from NCDS5), a number of additional questions were included which focused more on features of employment and work which impinged on other life domains. The criterion for identifying such measures was whether they were of relevance to analyses of the influences on the adult identities of the cohort members and/or other aspects of their development and functioning (eg as parents), as opposed to being of interest to studies more directly concerned with the impact of changing labour market conditions. The topics included working patterns (hours, times worked), receipt of fringe benefits, pension arrangements, and ‘family friendly’ working practices.

Subjective aspects

Individuals’ subjective evaluation of their employment circumstances is a crucial aspect of how these objective measures link with other aspects of their lives and wellbeing. Thus, new questions were included relating to job satisfaction, insecurity, and whether the current job was considered demanding or stressful. Two types of ‘Likert’ type attitude questions were also included in the new surveys; the first focusing on ‘attitudes to paid work’ (previously used at NCDS5) and the other on ‘children and working mothers’, drawing on established questions from the 1946 cohort and other surveys.

Household income

Information on earnings from current employment was obtained in the same way as at NCDS5. Details of other sources of household income was also repeated as at NCDS5, with appropriate updating to take account of changes in benefit and taxation systems – eg Jobseekers’ Allowance, Working Families Tax Credit.

Subjective information on income and financial situation

Information on feelings of financial hardship was sought from the parents of the NCDS cohort at several of the early follow-ups, and proved a useful measure in many studies at that time. It has also been used in more recent analysis, and would seem to be an important component of the subjective/individual agency aspect of the life course process. New questions were introduced to cover this topic, including expectations of future financial circumstances.

HEALTH AND HEALTH BEHAVIOUR

Both NCDS and BCS70 have their origins in a primarily medical concern with the health and welfare of the mother and her child during pregnancy and childbirth. Whilst both cohort studies have broadened their scope through subsequent sweeps, health and health behaviour have remained key aspects. A substantial amount of information has been gathered in these areas from parents and cohort members at each follow-up (and various sub-studies) via interview and self-completion, through medical examination and from medical records. This has provided the data for a wealth of analytic research, focusing in particular on: the natural history of specific conditions, including mental health; reproductive health; the association between health problems and social, educational, and economic circumstances; and cross-cohort comparisons. In addition, evidence based on the longitudinal cohort study data had a major influence on the Acheson Report on health inequalities.

Prospective and retrospective approaches

A major objective in designing the health questions for the new surveys was, as in the past, to obtain data which would inform both prospective and retrospective analysis of the natural history and trajectories of health and disease, and the search for causal factors. The appropriate data would act as outcome measures at present, and also as risk factors and baselines against which change can be measured at a later date.

The main hypotheses to be tested are that poor early life physical development and poor social circumstances are associated with reduced mental and physical functioning in adulthood, and that these adult outcomes are associated also with continuing intergenerational poor social circumstances, low educational attainment, low social class occupation, high stress employment, high stress family circumstances, adverse life events, poor health-related behaviour (nutrition, alcohol, smoking and exercise), and with raised risk of overweight/obesity.

Starting from this perspective, it was clearly essential that any new material collected should have longitudinal relevance. This could take retrospective or prospective format and could involve the gathering of data banked for future analysis later in the life-course. For example, data could be collected on a range of adverse practices or experiences and positive health-promoting behaviours which may be hypothesised as evolving as key precursors of future onset of disease/health outcomes or premature death. Further hypotheses could relate to the extent to which early adult or midlife health outcomes are associated with continuing and new social and physical advantages/disadvantages, positive/adverse life events, and positive/negative health related-behaviours and beliefs. The key questions to be addressed in all these cases concern the routes and antecedents to outcomes, both positive and negative.

Inter-cohort comparisons

The importance of continuity within cohorts in the selection of new health measures was matched by the need for replication between them, in order that cross-cohort mapping could maximise the value of longitudinal data already gathered. For example, because the 1970 cohort was born at a time when greater proportions (especially women) were entering higher/further education, when work was increasingly sedentary, nutritional choice available on wider scale, exercise increasingly fashionable, alcohol consumption rising, smoking falling in some sections of the population, and social mobility falling, inter-cohort comparisons will be extremely important in showing how these changes affect health, particularly health function, as it relates to early life risk factors.

Data collection strategy

The potential quantity of questions on health and health-related behaviour – both for repeating previously collected measures and for introducing new items – was considerably greater than for any other domain. Applying the parameters indicated earlier, a high proportion of questions were repeated from NCDS5, with some modifications where appropriate. In identifying new questions, other survey studies were consulted for appropriate measures eg NSHD (the 1946 cohort study) the Health Survey of England, the Whitehall Study, and the British Crime Survey.

A very large number of health topics have been covered in the interview, broadly following the pattern of previous adult sweeps, especially NCDS5. The following paragraphs list the main topic headings, and indicate where repeat measures or new questions are used. As with other domains, the full list of the questions are included in the appendix to this Working Paper.

Health Status

A single question asking for assessment of health status on a four point scale rating from 'excellent' to 'poor' has been widely used in research and shown to correlate with a large number of other variables. It was therefore important to repeat it in the current survey. The other aspect of health status is longstanding chronic illness or disability. Here, the focus is on distinguishing between those conditions acquired later in life and disabilities existing since birth. Previous surveys have covered this aspect of health status in considerable detail. The new data will offer the opportunity of investigating continuities and discontinuities.

Mental health and psychopathology

This again has proved a major area for research in the past and questions have been repeated on depressive/behavioural disorders and consultations with GPs. The Malaise Inventory, which can be used to identify people with a potential for clinical depression, has also been repeated, this time alongside the other major indicator of psychological state, the Cornell General Health Questionnaire. Both of these latter inventories, Malaise and the GHQ, are used in self completion mode.

Reproductive Health

The majority of women in the 1958 cohort were, at 41, nearing completion of their child bearing histories, which made it a particularly useful point at which to update the record. Those in the 1970 cohort, at age 29, were entering the peak childbearing period. The new surveys presented the first opportunity to collect detailed information about their child bearing. This covered details of each pregnancy, including those not reaching full term, and information on births, including birthweight. Birthweight relative to gestational age has been shown in numerous studies to be a key predictor in later development and health risks. Questions on current use of contraception were also included.

Respiratory function

As in previous sweeps, a considerable number of questions addressed respiratory problems, including asthma, cough, and wheeze. In the great majority of cases, continuity with earlier sweeps was maintained through repeat questioning; although suggestions from advisers led to some additions or improvements, eg questions on allergic rhinitis.

Visual acuity and hearing problems

Early NCDS and BCS70 sweeps conducted direct measures of special senses, including vision and hearing (visual acuity tests and audiograms). The new surveys obtained self-reported information on both sight and hearing, since no adult data had been collected for the 1970 cohort and not since NCDS4 for the 1958 cohort.

Hospitalisation

The record of hospital visits involving an overnight stay or day surgery is updated in the new surveys, from 1991 from NCDS cohort members and from 1986 for the BCS70 cohort. The information includes the number, time and reason for all such episodes.

Accidents/assaults

In addition to information about hospitalisation or other medical treatment, more detailed questioning in this area has been undertaken to make good past deficiencies. The aim was to gain a complete record of injuries occurring through accidents and assaults.

Health related and lifestyle behaviour:

As previously, the major areas covered are smoking and drinking using standard questioning techniques, including the Cage inventory in the case of alcohol consumption and abuse. Self-completion information about drugs taken is also explored, as well as diet and nutritional practice, again by self-report. Finally, there are questions on exercise and physical activity.

Physical and biometric measurements

There is good evidence to suggest that self reported heights and weights give good approximation to physical measurements of the kind that were used in NCDS5. In order to save time and money, in this survey we have returned to self assessments, as they were conducted at age 23 in NCDS4.

Specific Conditions

Although many of these will be relatively rare, especially for the 1970 cohort, identification of them is an important tool for screening out cohort members for detailed investigation. There is a potential for following up particular reported conditions through a further interview, medical tests and, with the cohorts' members permission, existing medical records via the GP. As before, information in the new surveys is of two kinds: a checklist is used to identify a number of common conditions raising from bronchitis through hernia to high blood pressure, with questioning directed at the condition as occurred in the last 12 months and whether a doctor has been seen, and, in the case of women, whether the condition had been related to the experience of pregnancy. In relation to a number of medical conditions, more detailed questioning has been pursued in order to gauge more precisely the status of the condition. The major areas for this are respiratory illness, migraine and back pain.

CITIZENSHIP AND VALUES

This domain is concerned with assessing cohort members' identities, especially in their exercise of the functions of citizenship or their exclusion from it. It embraces the social capital they have access to, which underpins their citizenship, and their attitudes and values.

Citizenship

Citizenship is an over-arching concept, extending across all the domains of life, which is usually identified with rights in three areas. These comprise civil rights (to own property, to freedom of association), political rights (to participate in political power), social rights (to a standard of living). The rights are coupled with obligations, which are mainly exercised through various forms of *participation*, eg the act of voting. Lack of participation in the different domains of life rises from society-imposed or self-imposed exclusion.

The context of citizenship, and its expression through politics, has been experienced differently by the two cohorts. The 1958 cohort left school at 16, when Labour was still in power. The ‘winter of discontent’, at the tail end of the Callaghan government in 1979, was experienced when they were five years into employment. This laid the foundation for the Thatcherite revolution, in which individualism and anti-collectivist stances took over from the welfarism (sociotropic values) that had dominated politics until then. In contrast, the 1970 cohort left school six years into the Thatcher Government and until 1997 had no experience of any other political regime. It was this cohort which were both more cynical about politics, yet also led the shift away from support for the Conservatives, particularly women, a majority of whom supported Labour for the first time. Two and a half years into the Blair Government, it is difficult to say what continuities and discontinuities we are likely to find, particularly as the new Government appears to be taking on much of the individualistic and broad market-oriented stance of its predecessor.

Social capital

The main base for active citizenship, and the antidote to exclusion from it, is the accumulation of *social capital*. Social capital resides in the sets of relationships in which the individual is engaged through the community and is identified with a common sense of purpose, high levels of trust, and networks that lead to various sorts of opportunities. Added to it, and in some ways more significant in relation to later achievements, is *cultural capital*. This comprises the values and understandings which underpin educational achievement, and the labour market success that is likely to arise from it as vested in the *human capital* the person acquires. Much of the generation and reproduction of social and cultural capital comes through the family, which can be seen as investing such capital in its children, with a view to ensuring their future and, indirectly, that of the parents themselves.

Attitudes and values

These are of two kinds of attitudes: those concerned with the self-concept and personality, such as motivation, self-efficacy and locus of control (self-attitudes), and those concerned with presentation of self to other people through expression of evaluative beliefs (social attitudes). Values are the expression of the deeper pre-dispositions that produce a degree of consistency in attitudes – ‘left–right’ and so on. Both kinds of attitudes and values mediate action, and arise from it. Thus, parents’ educational aspirations for their children will only convert into educational achievement on the part of the children themselves if the children adopt positive attitudes to learning and become motivated to achieve. Similarly, support for a particular political viewpoint aligning with the platform of one party rather than another, is likely to be realised in voting for that party. On the other hand, apathy towards politics, and cynicism about political institutions and politicians, underpin the decision whether to vote, and, consequently, whether to exclude or include oneself in the political process.

To gain the advantage of attitudes and values as mediators we needed repeated measures of those used earlier. At the same time, it was important to tap into new attitudes and values likely to be of significance in people’s’ lives in the future. To explain the development of the life course in any of our domains – health, family, employment and so on - we also needed to include subjective appraisals expressed through attitude measures.

Data collection strategy

In NCDS, attitudes and values were covered only for the first time in the fifth sweep in 1991. Accordingly, their use as mediators of circumstances on behaviour has not been possible so far.

Expressions of citizenship, too, were investigated only for the first time in NCDS5. However, BCS70, at age 16, did include a wide range of attitude measures, and the 26- year postal survey drew on attitude measures from NCDS5.

Attitudes

The self-completion questionnaire employed at NCDS5 yielded ten attitude dimensions:

support for work ethic
support for authority
support for traditional marital values
permissiveness about work and family
opposition to family life
political cynicism
left-right beliefs
support for sex equality
environmentalism
anti-racism.

Two criteria informed decisions about measures for the new survey. First, we needed to decide the dimensions of values and attitudes that were of interest and ensure that the minimum number of items were included in the questionnaire to measure them. The new surveys repeat the above measures in self-completion format, with some modification to the items covering family life. In addition, new items are introduced to tap attitudes to learning and to information technology.

As indicated in the relevant preceding sections, attitude items were also included as subjective appraisals in all the domains of life covered in the survey, ie family, employment, learning.

Self attitudes

As before, too, the new self-completion questionnaire also covers self-attitude measures, including self-efficacy, locus of control, partnership satisfaction (drawn from the Locke Wallace scale) and general life satisfaction.

Political participation and preference

Questions used in previously on political participation, e.g. voting behaviour and political party affiliations are repeated in the interview for the new survey.

Community involvement

A large number of questions were devoted to this measurement in NCDS5, and, as an essential area for social capital appraisal, there was value to be gained from repeat measurements. These include questions relating to membership and active involvement in trade unions, other community organisations and religious groups.

SURVEY PREPARATION

The preceding sections have described the rationale for undertaking new surveys of the 1958 and 1970 cohorts, and the parameters which guided decisions about the data to be collected. This section presents a technical account of the preparation for, and conduct of, the actual survey. It begins with the development of the new instrumentation, then goes on to describe the tracing programme undertaken to locate as many cohort members as possible. Details are then given of the way in which the major fieldwork operation was carried out.

CAPI

A major innovation in the conduct of the new surveys was the adoption of computer-assisted, rather than paper-based, methods of data collection. The use of CAPI (Computer Assisted Personal Interviewing) serves to improve the quality of the data collected, by simplifying the conduct of the interviews with complex filter structures. It also provides for the rapid production of clean data, because of the facility to edit data on entry. In addition, the employment of CAPI simplifies coding of information about occupation and answers to other open-ended questions where responses can be keyed in during the interview. The same method, CASI (Computer Aided Self-Interviewing), was employed with the self-completion instrument, when the laptop computer was handed over to the cohort members themselves.

The advantages gained from the move to computer-assisted data collection will be data of a higher quality, available earlier, and at a lower unit cost than for earlier surveys.

National Centre for Social Research (NCSR)

The new surveys were funded under the auspices of the Joint Centre for Longitudinal Research, which includes the Centre for Longitudinal Studies, the International Centre for Health and Society at University College London, and the National Centre for Social Research (NCSR), Britain's leading independent survey organisation.

The project was carried out by CLS and NCSR in partnership. NCSR (formerly Social and Community Planning Research) had previous experience of the cohort studies, having been part of the fieldwork consortium for NCDS5 in 1991. The specific responsibilities of the NCSR team in relation to the new surveys included:

- developmental work on survey instruments for use in the field
- organisation and conduct of piloting
- development of the CAPI application
- briefing (and debriefing) of all interviewers
- conduct of the main fieldwork, including sending advance letters to cohort members, tracing, contacting and interviewing, and sending thank-you letters
- initial post-field data preparation, including coding, data entry, and editing.

Questionnaire development

Two types of survey instrument were developed for use in the new follow-ups. The first was, as in previous sweeps, a main interview schedule designed to gather information about event histories and current circumstances in all the life domains covered by the study. Applying the parameters for data

collection outlined in previous sections, the question content was largely (around 90%) identical for the two cohorts. The main difference lay in the period to be covered by the event histories: for NCDS this (almost always) went back to the fifth sweep in 1991; for the 1970 cohort, the key date was the last full interview survey in 1986.

The second, a self-completion questionnaire, focused on attitudes, relationships and some aspects of health and health-related behaviour. A similar self-completion instrument (What Do You think?) had been sent out in advance of the interview at NCDS5 in 1991, and was a success with cohort members, who were pleased to have the chance to express their views. However, we could not exclude the possibility that its length put some people off filling it in or taking part in the interview itself, and it was decided on this occasion to administer it in conjunction with the main interview.

Pre-piloting

The discussions and reports from the Advisory Groups indicated a number of new areas of questioning, most particularly in the areas of family life and relationships. These included topics relating to *parenting behaviour and attitudes* on the part of cohort members. Although an extensive review of earlier NCDS and BCS70 sweeps, and other relevant surveys (eg NSHD, BHPS and the Australian Parenting-21 study) revealed a number of possible items, it was decided that some exploratory work was needed to identify the most meaningful areas of questioning. It was also considered important to extend questions about parenting to fathers as well as, conventionally, mothers only, and this needed to be tested out. Another new topic for this survey was *relationship with own parents*. There was some evidence from piloting work in NCDS5 that this was a sensitive area of questioning. Such considerations also pointed to the need for preliminary work to test out some of the ideas and issues involved.

Qualitative interview

The development of questions in these areas began by drawing up a list of topics to be explored in a semi-structured, qualitative interview. The aims of the interview were to:

explore the meanings and relevance of the topics

assess the sensitivity of the topics

identify recurring themes with a view to constructing appropriate questions and response categories.

In the sphere of parenting behaviour and attitudes, the following topics were focused on:

- parental involvement in children's learning and education
- control and autonomy (including discipline, rules and expectations of self-care)
- parent-child relationships (emotional warmth and distance)
- time spent with children
- stresses and satisfactions of being a parent

Questions about relationship with own parents centered on:

- the nature of current relationships (frequency of contact, emotional closeness)
- expectations of future help parents will need and willingness to give any such help
- help received from parents

The instrument was initially tested with staff in CLS who were parents acting as respondents. The interviews were taped, and following this initial work revisions were made to the schedule. Three experienced qualitative interviewers were then employed to carry out further interviews, which took

place between December 1998 and March 1999. They were briefed by members of the CLS team. After each had carried out a few interviews, an initial debriefing took place and some revisions were made to the schedule. A final debriefing took place at the end of the assignments.

The interviewers were encouraged to find respondents themselves, through local schools or other contacts. However, they did not carry out any interviews with respondents who were known to them. Interviews were carried out either at the home of the respondent or at a mutually convenient site such as place of work. Respondents were paid £10 for each completed interview. Where the contact was made through a school, a further £1 per interview was donated to the school fund. Respondents were reported to be happy to answer the questions and all interviews embarked upon were completed.

The interview sample consisted of 14 mothers and eight fathers aged 28-32 or 38-42 (corresponding to the cohort members' ages) and living with their children. They lasted on average one and a half hours. All were tape-recorded, and notes were made and discussed by the CLS team. The results revealed no difficulty in putting questions about parenting to fathers, or to questioning either age group about their relationship with their own parents. The material was used as the basis for devising new survey questions in these areas, and also produced additional information which could contribute usefully to the design of planned surveys of the children of the cohorts.

Paper pilot

Following the recommendations from the Advisory Groups and the pre-pilot exercise, first drafts were produced of the interview schedule and self-completion questionnaire. Before embarking upon the major task of converting these instruments into Blaise code, as required by CAPI, a paper pilot was conducted by three NCSR interviewers. Its main purpose was to establish whether the draft instruments fitted into the costed time limit of 85 minutes for the main interview and 15-20 minutes for the self-completion questionnaire, and also to identify any problem of 'flow', question wording, pre-codes, filters and so on.

Again, cohort members were not involved in this exercise. Each interviewer were asked to obtain four interviews with a mixture of male and female respondents in the appropriate age bands, including some with children. The CLS and NCSR teams carried out joint briefings of the interviewers, who carried out the work over the weekend 9-11 July 1999. The interviewers completed Pilot Evaluation Forms, which recorded the length of each section of interview and gave feedback on other aspects of the interview. A joint debriefing session was held shortly afterwards.

The interviewers carried out a total of 12 interviews with non-cohort members. The characteristics of the respondents are shown below (from table provided by NCSR July, 1999).

Table 1: Sample for paper pilot interviews

Characteristics	Number Interviewed
Male	5
Female	7
Age 27-33	6
Age 37-43	6
Working	11
Children Under 16	8

The results of the pilot interviews indicated that the main interview needed to be cut by approximately 20 minutes, and the self-completion by 5-10 minutes. As each section had been timed, it was possible to identify those which appeared disproportionately lengthy. In addition, the piloting provided useful insight into the ordering and suitability of the questions and indicated where modifications could improve the instruments.

Table 2: Paper pilot: timings for interview and self-completion

	Length (minutes)
Interview	
Shortest	73
Longest	141
Mean	103
Median	101
Self-completion	
Shortest	25
Longest	85
Mean	35
Median	28.5

Programme pilot

A further pilot was carried out between 8 and 19 September 1999. It was based on two computer-assisted survey instruments: the interview (CAPI) and the self-completion questionnaire (CASI). Again, the pilot was designed to establish the time taken to administer the (hopefully) near final draft instruments; and to identify any problems with flow, application of definitions, question wording, recall of dates or past events, pre-codes, routing, etc. The content of both instruments was based on a substantial review of those used for the first pilot survey in order to reduce the burden on both interviewer and respondent. This was carried out alongside the major task of converting traditional, paper-based instruments to computer-assisted interview and self-completion.

This pilot was also a ‘dress rehearsal’, and designed to test the various administrative aspects of the planned surveys, including advance letters; briefing; interviewer respondent-contacting and tracing; and the flow of information between the CLS tracing team, the NCSR field office and interviewers.

In order to mirror the circumstances of the main survey as nearly as possible, advance letters were mailed to all NCDS and BCS70 cohort members with a confirmed address in 11 selected postcode districts (in Blackburn, Brighton, Bristol, Cambridge, Cardiff, Croydon, Darlington, Leicester, Motherwell, Sheffield and Wolverhampton). One NCSR interviewer from each area was briefed on all aspects of the survey, including tracing, contacting and administering the interview and self-completion. They were asked to obtain at least 5 interviews (and self-completions) with at least 5 NCDS and 5 BCS70 cohort members.

The pilot again revealed a relatively small number of problems with flow, definition, question wording, recall, precodes and routing. It also indicated that the tracing and contacting procedures worked

effectively, and highlighted the extent and nature of the burden imposed on the CLS and NCSR teams by the exchange of information between office and field. Most importantly, as the table below indicates, it showed that the survey instruments were no longer overlength.

Table 3: Programme pilot: timings for interview and self-completion

	Minimum (minutes)	Mean (minutes)	Median (minutes)	Maximum (minutes)
Interview:				
Target	-	75	-	-
Actual:				
<i>All</i>	41	75	71	178
<i>NCDS</i>	41	73	66	138
<i>BCS70</i>	51	78	75	178
Self-completion:				
Target	-	20	-	-
Actual:				
<i>All</i>	11	19	18	48
<i>NCDS</i>	11	20	18	48
<i>BCS70</i>	11	19	18	40

Tracing

An essential requirement for successful surveys of the NCDS and BCS70 cohorts is an up-to-date address file. Experience suggests that, once traced, cohort members are more likely than a general population sample to agree to provide information. For this reason, considerable efforts have been made in recent years to maintain ongoing contact with the study subjects. This has been achieved mainly through the annual mailing of a birthday card to all those with a known address. Unfortunately, an important minority of cohort members remains untraced, and a special effort was needed in the new survey programme to locate as many as possible. This would serve not only to maximise response, but also to minimise response bias.

The tracing exercise for the new surveys was carried out by a team based in CLS, comprising two permanent and up to six temporary staff. Early work to recruit and train appropriate staff ensured that there was a small core of experienced workers. Systems and procedures were kept under constant review; the retraining of staff undertaken as databases and systems were modified in the light of experience, and new sources of address information utilised.

The work of the tracing team was based on the computerised NCDS and BCS70 address databases. These databases had been enhanced and established on a network server, so that they were accessible to all members of the tracing team. In order to ensure the security of the confidential address data, this network was isolated, completely independent of the other computer networks within CLS and the

Institute of Education. To ensure the safety of new information on the whereabouts of cohort members, backups were taken twice daily, and the performance of the systems regularly monitored.

The work of the tracing team built on the wealth of experience gained during the NCDS 1981 and 1991 follow-ups, and the postal survey of BCS70 in 1996. In addition to the existing address databases, use was made of a variety of sources of information, as indicated below.

Table 4: Sources of address information

- Annual birthday card mailings
- Address and contact address information provided by cohort members in the past
- Other information contained in study records
- Telephone number databases
- Postcode databases
- Electoral register databases
- National Health Service Central Register records of NHS registration, emigrations and deaths
- Health Authorities address records
- Driver and Vehicle Licensing Agency address records
- Ministry of Defence records
- Media appeals
- Interviewer detective work during fieldwork

Because of the more extended lead time to the new surveys than originally anticipated (due to funding uncertainty), the tracing exercise proved more successful than in any previous adult sweep of either NCDS or BCS70. Table 5 below summarises the state of the NCDS and BCS70 address databases on 27 October 1999, shortly before the main surveys began. The current whereabouts of over 14,000 members of each cohort were known, and the majority of addresses had been confirmed since tracing for the new surveys began. The efforts of the tracing team continued once fieldwork started, and were supplemented where necessary by interviewers seeking to establish the location of the cohort member, speaking to neighbours and others and following up leads as appropriate. It was particularly pleasing that the numbers of traced subjects was roughly the same in each cohort, thus achieving one of the major objectives of the new project – inputting both longitudinal studies on the same footing from now on.

Table 5: Summary of the state of the address databases at 27 October 1999 prior to the start of fieldwork

	NCDS	BCS70
	Status at 27/10/99 a	Status at 27/10/99 b
TRACED		
Potential respondents:		
Confirmed address	12794	12986
Forces (confirmed)	6	19
Parental address	85	378
Temporary address	16	11
	12901	13394
Others:		
Emigrated (confirmed)	281	246
Refusal	815	284
Proxy refusal	44	54
Death	247	109
	1387	693
Total traced	14288	14087
UNTRACED		
Potential respondents: ^{c, d}		
Forces, not confirmed	23	14
Demolished	75	4
Gone away, etc	1635	1409
Untraced/unconfirmed	326	1088
	2059	2515
Others		
Emigrated (unconfirmed)	113	93
Total untraced	2172	2608
All cohort members	16460	16695

NOTES

a. NCDS: 96% of addresses were confirmed in 1998/99

b. BCS70: 93% of addresses were confirmed in 1998/99

c. A significant number of these would be traced on the address databases of the Health Authorities. First returns to mailings seemed to confirm this (NCDS=853 (70%) confirmed addresses; BCS70=855 (62%))

d. Additional tracing would rely on media appeals and interviewer tracing

Source: CLS tracing records

Fieldwork

Given the unprecedented size of the sample to be interviewed for the simultaneous survey of both cohorts), the fieldwork, which began in early November, 1999, was carried out in a series of six waves. With the exception of the final one, each wave broadly covered all areas of the country and targeted some 2,500 members of each cohort. The final wave would contain the vast majority of 'outliers' (addresses in the Highlands and islands), as well as movers who could not be picked up in earlier waves.

Table 6: Approximate dates of fieldwork waves

Wave	1st briefing	fieldwork
1	29/10/99	29/10-13/12
2	29/11/99	29/11-10/01
3	06/01/00	06/01-14/02
4	31/01/00	31/01-13/03
5	28/02/00	28/02-10/04
6	03/04/00	03/04-18/05

Each fieldwork wave began with the mailing of an advance letter, advising the cohort members selected for inclusion in the wave that an interviewer would be calling shortly. This was followed by a series of face-to-face interviewer briefings, during which the background and purpose of the survey was explained; careful instructions were given on the contact, tracing and other administrative procedures; the survey content was outlined; and the interviewers were given an opportunity to conduct a full ‘dummy’ interview. This was designed to take them through all the main sections of the CAPI and CASI instruments, and to highlight areas of questioning and/or issues of definition where the surveys differed from others on which the interviewers may have worked. To minimize travel costs, briefings were conducted in different parts of the country. Each was conducted by a member of the CLS team and a member of the NCSR team, and attended by 10-12 interviewers.

Table 7 : Interviewer briefings

Wave 1
Birmingham (9/11/99; 10/11/99) Bristol (1/11/99; 8/11/99) Edinburgh (3/11/99; 8/11/99) Leeds (4/11/99) Liverpool (2/22/99; 9/11/99) London (29/10/99; 1/11/99; 2/11/99; 3/11/99; 4/11/99; 5/11/99) Newcastle (5/11/99)
Wave 2
Birmingham (6/12/99) Bristol (2/12/99; 3/12/99) Glasgow (1/12/99) Leeds (3/12/99) Liverpool (30/11/99) London (29/11/99; 30/11/99; 1/12/99; 2/12/99)
Wave 3
Bristol (6/1/00) Leeds (7/1/00) London (6/1/00; 7/1/00) Birmingham (10/1/00)
Wave 4
Bristol (2/2/00) Leeds (3/2/00) London (1/2/00; 2/2/00)
Wave 5
Leeds (29/2/00) London (29/2/00)
Wave 6
London (3/4/00)

Each wave of interviewing was carried out over a period of 6-8 weeks. The mailing of the advance letter generated numerous calls to the CLS Tracing Team from cohort members who had a new address, name, or telephone number; or from those who, for example, would be unavailable for interview in the next few weeks. This vital information was passed to the NCSR field office, in order that interviewers could be advised accordingly. As the fieldwork progressed, information came back from the field about the success of contacting respondents, including identification of addresses that were out of date. In cases where movers could be traced in the field by an interviewer, the new address would be visited by the interviewer or, if some distance away, passed to another interviewer to pursue. For movers where no new address could be found, information was returned via the NCSR field office to the CLS Tracing Team where further work could be done to try to trace the cohort members. In the order of 800-1,000 addresses come back from each wave of fieldwork on this basis, which meant that, as with NCDS4 and NCDS5, CLS needed to maintain a reduced, but consistent, tracing operation throughout the period of fieldwork.

Response in the interview reached a very high level. Evidence from Wave 1 and other early waves suggested that the surveys were going to be highly successful. As the fieldwork progressed, figures indicated that interviewers were achieving an interview with over 90 per cent of those contacted for both NCDS and BCS70. It is also important to note that many of those not contacted by interviewers during any one wave (for example, because they were ill, away, or had changed their address) could be re-allocated for interview at a later date.

Data was returned from the field (via modem), and coding and residual editing undertaken before clean data was passed to CLS for merging with existing cohort study data, longitudinal editing and documenting.

New datasets

As with previous sweeps of the cohorts, the new cross-sectional and longitudinal datasets will be deposited at the ESRC Data Archive at the University of Essex, from where they will be available for secondary analysis. This will bring to completion one of the most successful follow-ups in the birth cohort studies that has ever been carried out. It is particularly pleasing to be able to anticipate confidently the re-establishment of the BCS70 cohort as one of the major British longitudinal research resources.

Contacting the NCDS or BCS70 User Support Group

Further information about any aspect of NCDS or BCS70 may be obtained by contacting the respective *User Support Group* at the address below.

Centre for Longitudinal Studies
Institute of Education
20 Bedford Way
London WC1H 0AL

Telephone: +44 (0)20 7612-6860
Fax: +44 (0)20 7612-6880
Email: cohort@cls.ioe.ac.uk
WWW: <http://cls.ioe.ac.uk/>

APPENDIX 1: NCDS/BCS70 Follow-ups 1999/2000 Summary of content

INTERVIEW	
Household Grid	
Household grid	Fairly standard grid. Gathers details of sex, age, relationship to respondent, marital status
Ethnicity	Based on the new question being developed for the 2001 Census
Language spoken in the home	Language usually spoken
Housing	
Current address	Rooms, heating, housing benefit, tenure
Intentions to move	Moving, future tenure, why moving
Property inheritance	Ever inherited house/flat, when
Homelessness	Experience of, number of times, date of last, why, where stayed, duration, applied to LA
Housing history	Works backwards from last address. Seeks details of: dates, tenure, living arrangements, reasons for moving.
Relationships	
Marital status	Legal status, prior cohabitation, dates of marriage/cohabitation, age, marital status of partner
Relationship history	Marriage/cohabiting. Works backwards from last relationship. Seeks details of: dates, sex, age, marital status, marriage, separation/divorce
Children	
Pregnancy history	Works backwards from last pregnancy conceived/ fathered. Seeks details of: outcome, name, sex, birth weight, DOB, problems, smoking in pregnancy, other parent, whereabouts, support and circumstances of absent children
Lone parenthood	Periods of ≥ 1 month. Works forward for up to 4 periods. Seeks details of dates and numbers of children
Infertility	Ability to have and plans for (more) children
Adopted children	For up to 4 children. Seeks details of age of child on adoption and nature of adoption
Partner's children from a previous relationship	Number, whether seen or visit
Children over 16	For up to 4 children. Seeks details of: name, age, economic, marital and parent status
Family activities	Things done as a family
Demands of parenting	Physical/emotional/time demands, worries, closeness of family
Family, Social Relationships & Support	
Contact with family	Parents alive/contact/close/divorce/worries/help. Contact with siblings/in-laws/grandparents
Emotional support	Sources/nature of emotional support
Family Income	
Other Income	Income from benefits and regular income form other sources received by respondent and/or partner
Financial situation	Organisation of household money and money problems
Employment	
Economic activity	Fairly standard question
Current job	Seeks details of: earnings, hours, fringe and Other benefits, pensions, prospects
Other paid work	Weekly earnings from odd/casual jobs
Currently unemployed	How became unemployed and job search
Labour market histories	Periods in a job or not in a job lasting ≥ 1 month. Works backwards. Seeks details of: circumstances and dates. If in job gathers details of job, employer, responsibilities (for SOC and SIC coding)
Partner's job	Age left full-time education, economic status, job, employer, responsibilities (for SOC and SIC coding), earnings

Lifelong Learning	
Qualifications	Age left full-time education and educational and vocational qualifications held (subjects, grades, dates awarded, where studied)
Current course for qualification	Qualification, subject(s), date started, where studied
Assessment of current/most recent course	Why taken and expected/experienced benefits
Other courses and training	Number, why taken and expected/experienced benefits
No formal learning	Reasons why no learning
Learning overview	Useful and enjoyable periods of learning
Contact with information technology	Use of computers at home and at work
Literacy and numeracy	Problems with reading, writing and maths, implications and courses to improve
Health	
General health	Self-assessed health and experience of list of conditions, including age at onset and contact with doctor
Long-term health conditions	Details of longstanding illnesses, etc (including limiting impact and age at onset), impact on employment, registered disabled.
Respiratory problems	Coughing, phlegm and shortness of breath
Mental health	Experience of mental health problems, including age at onset
Seeing and hearing	Problems with sight/eyes and with hearing
Other conditions	Details of other health conditions requiring regular medical supervision
Accidents/injuries	Works backwards. Details of accidents/ injuries/assaults (age, why admitted, out/in-patient, type of injury). Nature of any permanent disability resulting from any accident/etc
Hospital admissions	Works backwards. Age and why admitted
Smoking	Smoking habit of respondent and partner
Drinking	Alcohol consumption in last 7 days, other aspects of drinking behaviour
Diet	Frequency of consumption of types of food, vegetarian or other special diets
Exercise	Exercise at work and in daily life
Height and weight	Self-reported height and weight and assessment of weight
Citizenship and Values	
	Involvement with organisations, voting behaviour and intentions, political alignment, trade union membership, religion, newspaper readership, car ownership, values, political activity

SELF-COMPLETION	
<i>Your views</i>	Attitude statements
<i>How you get on with your husband, wife or partner</i>	Includes Locke-Wallace
<i>Some more of your views</i>	Attitude statements
<i>How you feel</i>	Malaise Inventory
<i>Your skills</i>	How good at skill/is skill used at work
<i>More of your views</i>	Attitude statements
<i>How you feel about your life so far</i>	GHQ 12
<i>More of your views</i>	Attitude statements
<i>School exclusion and truancy</i>	Number of temporary/permanent suspensions/exclusions; frequency of truancy
<i>Contact with the police and crime</i>	Number of times moved on, questioned, warned, taken to police station, cautioned, found guilty by a court
<i>Use of illegal drugs</i>	Whether tried number of specific drugs ever/in last 12 months

NCDS/BCS70 1999-2000 Follow-ups

***Guide to the Combined
Dataset***

(Revised December 2002)

**Compiled by on behalf of the
Joint Centre for Longitudinal Research Team**

by

Peter Shepherd

JOINT CENTRE FOR LONGITUDINAL RESEARCH

The 1999-2000 follow-ups of the National Child Development Study (NCDS) and the 1970 British Cohort Study (BCS70) were carried out under the auspices of the **Joint Centre for Longitudinal Research**. The JCLR exists to promote and facilitate the widest possible use of longitudinal studies, and the development of expertise in longitudinal research. It is a partnership of three institutions with unparalleled experience in the conduct of longitudinal surveys and the analysis of longitudinal data:

- **Centre for Longitudinal Studies (CLS)**, Institute of Education, University of London
- **International Centre for Health and Society (ICHS)**, University College Medical School, London
- **National Centre for Social Research (NatCen)**

The work of designing and conducting the NCDS and BCS70 1999-2000 Follow-ups was carried out by a joint team drawn from two of the JCLR partners – CLS and the National Centre. The key members of the team are listed below:

CLS NCDS/BCS70 Team

Research design and implementation:	John Bynner, Peter Shepherd, Elsa Ferri, Kate Smith, Neville Butler
Cohort Studies Administrator:	Denise Brown
Tracing Unit:	Kevin Dodwell, Farnaz Farahmand, (and temporary tracing staff to numerous to name here)
IT Team:	Mahmood Sadigh, Mos Mojaddad

National Centre NCDS/BCS70 Team

Design and implementation of surveys:	Debbie Collins, Kavita Deepchand, Rory Fitzgerald, Jane Perry
Field administration:	Elaine Iffland, Theresa Patterson, Glenis Naudo, Jill Penn, Maureen Slater
CAPI/CASI programming:	Michael Hart, Kevin Palmer, Steve Kelly

The Centre brings together expertise the conduct of longitudinal research and the potential for major scientific advances at the interface of education, medicine and social science.

ACKNOWLEDGEMENTS

We wish to acknowledge the support for these two follow-ups of our principal funders: the Economic and Social Research Council; Government Departments and Agencies (Office of National Statistics, Department for Education and Employment, Department of Social security, Department of Health, Scottish Executive, Basic Skills Agency); and the International Centre for Child Studies

The work could not have been carried out successfully without the involvement of over a hundred advisors drawn from researchers, policy makers and funders, who we consulted throughout the design of the surveys.

The names of NCDS/BCS70 Advisors, and an indication of their contribution may be found in the following document, which also accompanies the data deposit:

The design and conduct of the 1999-2000 surveys of the National Child Development Study and the 1970 British Cohort Study.

REVISED GUIDE (December 2002)

This Guide was revised to accompany the deposit with the UK Data Archive of revised data for the NCDS and BCS70 1999-2000 Follow-ups.

The revision of the data was undertaken in the Centre for Longitudinal Studies by Brian Dodgeon and Rosemary Creeser.

PREFACE

This document has been prepared to accompany the initial deposit, with the UK Data Archive at the University of Essex, of data from the most recent follow-ups of two continuing, multidisciplinary, national, longitudinal studies – the National Child Development Study (NCDS) and the 1970 British Cohort Study (BCS70).

The other elements of the deposit, to which reference will be made throughout this document, are identified below. Users are advised that they will need to consult all elements of the documentation to gain a full understanding of the data.

NCDS/BCS70 Deposit: Elements

Title	Format
NCDS and BCS70 1999-2000 Follow-ups: Initial Cross-sectional Data (June 2001)	SPSS
The design and conduct of the 1999-2000 surveys of the National Child Development Study and the 1970 British Cohort Study	Word
NCDS/BCS70 1999-2000 Follow-ups: CAPI Documentation	Word
NCDS/BCS70 1999-2000 Follow-ups: Technical Report	Word
NCDS/BCS70 1999-2000 Follow-ups: Interactive Data Dictionary for Combined NCDS/BCS70 SPSS dataset (based on the SPSS Data Dictionary)	Idealist for Windows
NCDS/BCS70 1999-2000 Follow-ups: Guide to the Combined Dataset (June 2001)	Word
NCDS Publications list 2001	Word
BCS70 Publications List 2001	Word

REVISED GUIDE (December 2002)

This Guide was revised to accompany the deposit with the UK Data Archive of revised data for the NCDS and BCS70 1999-2000 Follow-ups. The main revisions relate to:

- BCS70 identifiers – Users are now urged to use the variable KEY wherever possible. Further details are available in:
CLS Cohort Studies Data Note 1: Longitudinal Linkage in BCS70: Rationalising Case Identifiers
- Updates of the NCDS and BCS70 pregnancy histories to correct known errors. Further details are available in:
CLS Cohort Studies Data Note 2: Pregnancy Histories in the Combined NCDS/BCS70 1999/2000 Data
- Updates of the NCDS and BCS70 `grid to correct known errors. Further details are available in:
CLS Cohort Studies Data Note 3: Household Grid Variables in the Combined NCDS/BCS70 1999/2000 Data

CONTENTS

	<i>Page</i>
Joint Centre for Longitudinal Research	i
Acknowledgements	li
Preface	lii
Introduction	1
NCDS/BCS70	1
Advisors	2
Survey instruments/CAPI	2
Content of surveys	3
Fieldwork	6
Data coding and editing	7
Timetable	8
NCDS/BCS70 Data	9
Variable names	9
Variable labels	10
Value labels	11
Missing values	11
Variable order	11
Values	11
Consistency	11
Some useful variables	13
Self-completion	14
Proxy interview	15
Identifiers	16
Reference dates for retrospective data/histories	17
Date of interview	17
CAPI Derived variables	19
Further assessment of data quality	22
Household	24
Housing	25
Relationships	26
Children	30
Family Social Relationships & Support	32
Family Income	33
Employment	34
Lifelong Learning	39
Health	63
Citizenship and Values	82
Self-completion (covering attitudes/especially sensitive topics)	83
Appendices (for details see overleaf)	93

Appendices		Page
Appendix 1	Pregnancy History Variables	A2
Appendix 2	Proxy Interview Variables	A5
Appendix 3	NCDS/BCS70 Advisors contributing to quality assessment of NCDS/ BCS70 data	A10
Appendix 4	Earnings and related variables	A11
Appendix 5	Basic Skills and other variables	A57
Appendix 6	Problems with reporting of GCSEs	A96
Appendix 7	Highest Qualification	A138
Appendix 8	Attitude Scales	A163

INTRODUCTION

This document has been prepared to accompany the initial deposit, with the UK Data Archive at the University of Essex, of data from the most recent follow-ups of two continuing, multidisciplinary, national, longitudinal studies – the National Child Development Study (NCDS) and the 1970 British Cohort Study (BCS70). The follow-ups took place between November 1999 and September 2000, and it is noteworthy that this was the first time that both cohorts had been surveyed at the same time. They were designed and implemented jointly by the Centre for Longitudinal Studies of the Institute of Education, University of London (CLS), and the National Centre for Social Research (NatCen), on behalf of the Joint Centre for Longitudinal Research. The work was mainly funded by the Economic and Social Research Council, but important contributions were also made by a number of government departments, and by the Basic Skills Agency.

NCDS and BCS70

The National Child Development Study (NCDS) started life as the Perinatal Mortality Survey and examined the social and obstetric factors associated with stillbirth and infant mortality among over 17,000 babies born in Britain in the week 3-9 March 1958. Since this first study the whole cohort have been surveyed on five other occasions in order to monitor their health, education, social and economic circumstances. These surveys were carried out in 1965 (age 7), 1969 (age 11), 1974 (age 16), 1981 (age 23) and 1991 (age 33). As part of the 1991 survey, a special study was also undertaken of the children of one third of the cohort members, including assessments of the behaviour and cognitive development of approximately 5,000 children. There have also been surveys of sub-samples of the cohort, the recent occurring in 1996 (age 37) when information was collected on the basic skills of a representative sample of 10 per cent of cohort members.

The 1970 British Cohort Study (BCS70) was designed along similar lines to the NCDS, surveying over 17,000 babies born in Britain in the week 5-11 April 1970. Since the birth survey there have been four other major data collection exercises in order to monitor their health, education, social and economic circumstances. These were carried out in 1975 (age 5), 1980 (age 10), 1986 (age 16) and 1996 (age 26). As in NCDS, subsamples have been studied at various ages: for example at age 21, paralleling the NCDS survey at age 37, a 10 per cent representative sample was assessed for basic skills difficulties.

From their original focus on the circumstances and outcomes of birth, the two cohort studies have broadened in scope to map all aspects of health, education and social development of their subjects as they passed through childhood and adolescence. In latter sweeps, the information collected has covered their transitions into adult life, including leaving full-time education, entering the labour market, setting up independent homes, forming partnerships and becoming parents.

The latest rounds of data collection for NCDS and BCS70 took place in 1999/2000 (NCDS cohort member were 41/42 and BCS70 cohort members were 29/30). The main aim of these most recent surveys was to explore the factors central to the formation and maintenance of adult identity in each of the following domains:

- Lifelong learning
- Relationships, parenting and housing
- Employment and income
- Health and health behaviour
- Citizenship and values

Further details of this 'life course' theoretical framework and its use in the development of question areas for the most recent sweeps of NCDS and BCS70 are contained in the following, which also accompanies the data deposit:

The design and conduct of the 1999-2000 surveys of the National Child Development Study and the 1970 British Cohort Study.

Advisors

As with previous NCDS and BCS70 follow-ups, the surveys were designed in collaboration with advisors drawn from researchers, policy makers and funders. Following an initial meeting, held in March 1998, a number of advisory groups were formed, one for each of the proposed major topic areas to be covered by the new surveys – see below.

NCDS/BCS70 Advisory Groups

Child development & education
Citizenship and values
Employment and income
Family, parenting and housing
Health
Lifelong learning
Methodology

After the initial meeting, members of each group exchanged ideas via email and/or meetings and, ultimately provided written advice on the content of the survey instrumentation (June-September 1998). Subsequently, members of the advisory groups were updated on the development of the instrumentation, and the progress of the surveys. Latterly, a number of advisors have been involved in the initial assessment of the quality of the information obtained during the survey.

Further details of the advisory groups, and their consultations and advice are contained in the following, which also accompanies the data deposit:

The design and conduct of the 1999-2000 surveys of the National Child Development Study and the 1970 British Cohort Study.

Survey instruments/CAPI

Information was gathered from NCDS and BCS70 cohort members by interview and using self-completion questionnaire. A major innovation in the conduct of the new surveys was the adoption of computer-assisted, rather than paper-based, methods of data collection. The use of CAPI (Computer Assisted Personal Interviewing) serves to improve the quality of the data collected, by simplifying the conduct of the interviews with complex filter structures. It also provides for the rapid production of clean data, because of the facility to edit data on entry. In addition, the employment of CAPI simplifies coding of information about occupation and answers to other open-ended questions where responses can be keyed in during the interview. The same method, CASI (Computer Assisted Self-Interviewing), was employed with the self-completion instrument, when the laptop computer was handed over to the cohort members themselves.

Following development work - including: pre-piloting qualitative interviews and a pilot based on traditional paper interview schedules and self-completion questionnaire all based on non-cohort member samples; and a CAPI/CASI 'dress rehearsal' pilot based on a cohort member sample - the

survey instruments were programmed into a single CAPI/CASI instrument using Blaise 3 which was suitable for both NCDS and BCS70 cohort members. This was possible because over 90 per cent of questions were common to both cohorts. The major differences between the content of the NCDS and BCS70 surveys were:

- Reference dates for retrospective questions/histories. These were March 1991 for NCDS, and April 1986 for BCS70 - although NCDS cohort members who did not take part in the 1991 follow-up (NCDS5), were asked for details of qualifications gained since March 1981 (see below).
- The inclusion of additional questions for NCDS cohort members dealing with:
 - Children over the age of 16 years
 - Children absent from the household, but who were living with cohort member in 1991

A short proxy interview was also included in the Blaise program for use where the cohort member was unable to understand or respond to questions put by the interviewer, or to the self-completion. Questions were put to a family member or carer.

Details of the survey instruments and their development are to be found in the following, which also accompanies the data deposit:

NCDS/BCS70 1999-2000 Follow-ups: CAPI Documentation

NCDS/BCS70 1999-2000 Follow-ups: Technical Report

Content of surveys

As noted above, the survey instrumentation was developed in consultation with those who have been involved with the design and analysis of earlier NCDS and BCS70 surveys, other research advisors and funders; and in accordance with the following principles:

- Relevance to the stage of life reached
- Continuity with previous surveys
- Comparability across NCDS and BCS70
- Compatibility with other surveys (eg: BHPS, the General Household Survey and the (US) National Longitudinal Survey of Youth)

A summary of the topics covered by the surveys is given below.

Survey topics

Household
Housing
Relationships
Children
Family Social Relationships & Support
Family Income
Employment
Lifelong Learning
Health
Citizenship and Values
Self-completion (covering attitudes/especially sensitive topics)

A more detailed summary of the content of the surveys is also given below, and full details of the content of the survey instruments are to be found in the following, which also accompanies the data deposit:

NCDS/BCS70 1999-2000 Follow-ups: CAPI Documentation

NCDS/BCS70 Follow-ups 1999/2000 Summary of the contents of the surveys

INTERVIEW	
Household Grid	
Household grid	Fairly standard grid. Gathers details of sex, age, relationship to respondent, marital status
Ethnicity	Based on the new question being developed for the 2001 Census
Language spoken in the home	Language usually spoken
Housing	
Current address	Rooms, heating, housing benefit, tenure
Intentions to move	Moving, future tenure, why moving
Property inheritance	Ever inherited house/flat, when
Homelessness	Experience of, number of times, date of last, why, where stayed, duration, applied to LA
Housing history	Works backwards from last address. Seeks details of: dates, tenure, living arrangements, reasons for moving.
Relationships	
Marital status	Legal status, prior cohabitation, dates of marriage/cohabitation, age, marital status of partner
Relationship history	Marriage/cohabiting. Works backwards from last relationship. Seeks details of: dates, sex, age, marital status, marriage, separation/divorce
Children	
Pregnancy history	Works backwards from last pregnancy conceived/fathered. Seeks details of: outcome, name, sex, birth weight, DOB, problems, smoking in pregnancy, other parent, whereabouts, support and circumstances of absent children
Lone parenthood	Periods of ≥ 1 month. Works forward for up to 4 periods. Seeks details of dates and numbers of children
Infertility	Ability to have and plans for (more) children
Adopted children	Age of child on adoption and nature of adoption
Partner's children from a previous relationship	Number, whether seen or visit
Children over 16 (NCDS only)	Name, age, economic, marital and parent status
Absent formerly resident children (NCDS only)	Name, age, whereabouts, economic, marital and parent status; CM contact and payments for children absent from the household, but who were living with CM in 1991
Family activities	Things done as a family
Demands of parenting	Physical/emotional/time demands, worries, closeness of family
Family, Social Relationships & Support	
Contact with family	Parents alive/contact/close/divorce/worries/help. Contact with siblings/in-laws/grandparents
Emotional support	Sources/nature of emotional support
Family Income	
Other Income	Income from benefits and regular income form other

	sources received by respondent and/or partner
Financial situation	Organisation of household money and money problems
Employment	
Economic activity	Fairly standard question
Current job	Seeks details of: earnings, hours, fringe and Other benefits, pensions, prospects
Other paid work	Weekly earnings from odd/casual jobs
Currently unemployed	How became unemployed and job search
Labour market histories	Periods in a job or not in a job lasting ≥ 1 month. Works backwards. Seeks details of: circumstances and dates. If in job gathers details of job, employer, responsibilities (for SOC and SIC coding)
Partner's job	Age left full-time education, economic status, job, employer, responsibilities (for SOC and SIC coding), earnings
Lifelong Learning	
Qualifications	Age left full-time education and educational and vocational qualifications held (subjects, grades, dates awarded, where studied)
Current course for qualification	Qualification, subject(s), date started, where studied
Assessment of current/most recent course	Why taken and expected/experienced benefits
Other courses and training	Number, why taken and expected/experienced benefits
No formal learning	Reasons why no learning
Learning overview	Useful and enjoyable periods of learning
Contact with information technology	Use of computers at home and at work
Literacy and numeracy	Problems with reading, writing and maths, implications and courses to improve
Health	
General health	Self-assessed health and experience of list of conditions, including age at onset and contact with doctor
Long-term health conditions	Details of longstanding illnesses, etc (including limiting impact and age at onset), impact on employment, registered disabled.
Respiratory problems	Coughing, phlegm and shortness of breath
Mental health	Experience of mental health problems, including age at onset
Seeing and hearing	Problems with sight/eyes and with hearing
Other conditions	Details of other health conditions requiring regular medical supervision
Accidents/injuries	Works backwards. Details of accidents/ injuries/assaults (age, why admitted, out/in-patient, type of injury). Nature of any permanent disability resulting from any accident/etc
Hospital admissions	Works backwards. Age and why admitted
Smoking	Smoking habit of respondent and partner
Drinking	Alcohol consumption in last 7 days, other aspects of drinking behaviour
Diet	Frequency of consumption of types of food, vegetarian or other special diets
Exercise	Exercise at work and in daily life
Height and weight	Self-reported height and weight and assessment of weight
Citizenship and Values	
	Involvement with organisations, voting behaviour and intentions, political alignment, trade union membership, religion, newspaper readership, car ownership, values, political activity

SELF-COMPLETION	
Your views	Attitude statements
How you get on with your husband, wife or partner	Includes Locke-Wallace
Some more of your views	Attitude statements
How you feel	Malaise Inventory
Your skills	How good at skill/is skill used at work
More of your views	Attitude statements
How you feel about your life so far	GHQ 12
More of your views	Attitude statements
School exclusion and truancy	Number of temporary/permanent suspensions/exclusions; frequency of truancy
Contact with the police and crime	Number of times moved on, questioned, warned, taken to police station, cautioned, found guilty by a court
Use of illegal drugs	Whether tried number of specific drugs ever/in last 12 months

Fieldwork

The main fieldwork for the NCDS and BCS70 surveys began following the first interviewer briefing at the end of October 1999. It was conducted in a series of 6, overlapping waves as shown below.

Dates of Fieldwork Waves

Wave	1st Briefing	Fieldwork
1	29/10/99	29/10-13/12
2	29/11/99	29/11-10/01
3	06/01/00	06/01-14/02
4	31/01/00	31/01-13/03
5	28/02/00	28/02-10/04
6	03/04/00	03/04-25/09

With the exception of the last wave, each wave broadly covered all areas of the country and targeted some 2,500 members of each cohort (5,000 in total). The last wave contained the vast majority of outliers (addresses in the Highlands and islands), as well as movers and others who could not be interviewed in waves 1-5.

Each wave began with the mailing of an advance letter advising the cohort members selected for inclusion in the wave that an interviewer would be calling shortly. This was followed by a series of face-to-face interviewer briefings, held in different parts of the country, during which: the background and purpose of the survey was explained; instructions were given on the contact, tracing and other administrative procedures; the survey content was outlined; and the interviewers given an opportunity to conduct a full 'dummy' interview (designed to take them through all the main sections of the CAP/CASI instruments, and to highlight areas of questioning and/or issues of definition, where the surveys differ from other surveys on which the interviewer may have worked/be working).

During fieldwork, National Centre interviewers administered the CAPI/CASI instruments, after carrying out any necessary tracing to establish the whereabouts of cohort members. The interviewer tracing supplemented the efforts of the small Tracing Team maintained by CLS during the preparations for and conduct of the survey.

Between follow-ups efforts are made by CLS, through the mailing of an annual birthday card and other activities, to maintain contact with as many members as possible of both cohorts. But, unfortunately, at any one time an important minority of NCDS/BCS70 cohort members remains untraced, and considerable efforts have been made before and during the surveys to locate as many

as possible of the untraced. This serves not only to maximise response, but also to minimise response bias.

Prior to fieldwork, the CLS Tracing Team attempted to obtain a current address for as many cohort members as possible. The work of the team built on experience gained during the NCDS 1981 and 1991 follow-ups, and the BCS70 1996 Postal Follow-up to trace attempt to trace as many cohort members as possible. It made use of a variety of sources of information, as indicated below both before and during fieldwork.

During the period of fieldwork, which ended on 25 September, the efforts of the tracing team were supplemented, where necessary, by interviewers who sought to establish the whereabouts of the cohort member, speaking to neighbours and others, and follow-up leads as appropriate. Where the interviewer failed to find the cohort member, information was passed back to the CLS tracing team for further investigation.

Sources of address information during tracing

- Annual birthday card mailings
- Address and contact address information provided by cohort members in the past
- Other information contained in study records
- Telephone number databases
- Postcode databases
- Electoral register databases
- National Health Service Central Register records of NHS registration, emigrations and deaths
- Health Authorities address records
- Driver and Vehicle Licensing Agency address records
- Ministry of Defence records
- Media appeals
- Interviewer detective work during fieldwork

Further details of the fieldwork and tracing are to be found in the following, which also accompanies the data deposit:

NCDS/BCS70 1999-2000 Follow-ups: Technical Report

Data coding and editing

Data were returned from the field (via modem) and coding and residual editing undertaken

A major advantage of the use of CAPI and CASI is the reduced need for post-fieldwork editing – the majority of checks for validity, range and consistency can be incorporated into the CAPI/CASI program. Inevitably, however, there were checks, which were overlooked, or not initially thought necessary. These checks were incorporated into the DP activities undertaken by the National Centre after the survey.

The NCDS/BCS70 interview and self-completion include a number of open-ended questions where the verbatim answers of cohort members are keyed by interviewers, and a rather larger number of questions where precodes are provided for answers but provision is also made to record additional information where then precode 'other' is used. Following the start of the surveys, these questions were reviewed by the CLS team in order to determine the priorities for coding, and to identify the appropriate coding frames. Where possible, coding frames that had been employed for earlier NCDS/BCS70 surveys were adopted, although it was usually necessary to include additional codes. In other instances, it was necessary to develop a coding frame from scratch. Coding was undertaken by the National Centre and CLS, with the latter being responsible for coding of health and related

problems using the WHO *International Classification of Diseases*, and OPCS *Classification of Surgical Procedures and Operations*.

Further details of the editing and coding are to be found in the following, which also accompanies the data deposit:

NCDS/BCS70 1999-2000 Follow-ups: Technical Report

Timetable

An indication of the overall timetable for the NCDS and BCS70 follow-ups is given below with reference to a number of key events and activities mentioned above.

It should be noted that the full funding necessary to carry out both follow-ups was not finally secured until early 2000, and that the main fieldwork period was extended in order to ensure that as many cohort members as possible had an opportunity to participate.

Key event/activity	Date
Start of survey design and tracing of cohort members	January 1998
Consultative conference – advisory groups established	March 1998
Written advice received on content from advisory groups	June – September 1998
Development of instrumentation begins	July 1998
Further consultation with advisors	July 1998-July 1999
Qualitative pre-piloting undertaken	December 1998-February 1999
First, paper pilot (non-cohort sample)	July 1999
Development of CAPI/CASI program	July-October 1999
Development of instrumentation completed	August 1999
Second, CAPI/CASI pilot - 'dress rehearsal' (cohort sample)	September 1999
First briefings for main surveys	October 1999
Main fieldwork for both follow-ups begins	November 1999
Coding frames for open answers agreed, coding begins	March 2000
Last briefings for main surveys	April 2000
Meeting with advisors to report progress	May 2000
Fieldwork ends/tracing of cohort members ends	September 2000
Coding and editing completed/Last data transferred to CLS	December 2000
Initial assessment of data by CLS and advisors begins	January 2001
Deposit of initial cross-sectional data for both follow-ups	July 2001

NCDS/BCS70 DATA

The main data for the recent NCDS and BCS70 follow-ups is supplied to the UK Data Archive in the form of a single, combined SPSS dataset. This holds 5,188 variables for a total of 22,680 cases – 11,419 NCDS and 11,261 BCS70.

As the table below shows, the bulk of these (over 98%) where represent full interviews accompanied by a self-completion. The majority of the remainder are full interviews, where the self-completion was not answered, although there are also a small number of incomplete or partial interviews. Short proxy interviews were undertaken with a family member or carer where the cohort member was unable to understand or respond to questions put by the interviewer or to the self-completion. On a small number of occasions where the cohort member could understand and respond with the aid of an interpreter, an interview was attempted where a family member or carer was able to act as an intermediary.

NCDS/BCS70: Full, partial and proxy interviews

	NCDS	BCS70	Total
Full interview & self completion	11,281	11,116	22,397
Full interview, no self-completion	94	88	182
Partial interview	13	22	35
Proxy interview	31	35	66
Total	11,419*	11,261*	22,680

- 10 NCDS and 7 BCS70 interviews where carried out with the aid of an interpreter (normally another family member)

Source: Crosstabulation of UNOUT x SAMPLE x INTWHO

Variable names - The variable names on the dataset are those automatically allocated by the CAPI program (Blaise 3). Within the Blaise, each question has a variable name (rather than number), made up of a maximum of 8 characters, and this is used to determine the variable name on the dataset.

Where the question is repeated (eg: the same question is asked for each birth, relationship, job, qualification, etc reported), Blaise automatically allocates a number suffix (eg: name, name2, name3, name4). Unfortunately, where the variable name in the Blaise program was originally more than 6 characters long, Blaise truncates the name to allow for the suffix. As a result, there is not always a simple match between the Blaise program documentation and the data.

A somewhat extreme example of the range of variable names that may be encountered is given in the table below. In the interview, the question: "Who is the other parent of (name of baby)?" was repeated for each child conceived in each pregnancy. In Blaise, and in the CAPI documentation, this question has the name 'WHOPARB'. In the dataset, variable names are reserved for a maximum of 5 children conceived in each of a maximum of 8 pregnancies. Only for the first baby reported as conceived in the first reported pregnancy does the variable have the name WHOPARB. Blaise allocates modified variable names for each of the 39 other variables which identify the "other parent". It is important to note that information on pregnancy history was gathered by starting with the most recent pregnancy.

A full list of the variables relating to pregnancy history is given in Appendix 1.

Details of the CAPII program are to be found in the following, which also accompanies the data deposit:

NCDS/BCS70 1999-2000 Follow-ups: CAPI Documentation

Variable names for the repeated question: “Who is the other parent of (name of baby)?”

The question is repeated for each child conceived (maximum=5) in each pregnancy (maximum=8)

NB: Pregnancy 1 is the most recent pregnancy

Pregnancy	Baby	Variable name	Pregnancy	Baby	Variable name
1	1	WHOPARB	5	1	WHOPAR33
	2	WHOPARB2		2	WHOPAR35
	3	WHOPARB3		3	WHOPAR37
	4	WHOPARB4		4	WHOPAR39
	5	WHOPARB5		5	WHOPAR41
2	1	WHOPARB6	6	1	WHOPAR43
	2	WHOPARB7		2	WHOPAR45
	3	WHOPARB8		3	WHOPAR47
	4	WHOPARB9		4	WHOPAR49
	5	WHOPAR11		5	WHOPAR51
3	1	WHOPAR13	7	1	WHOPAR53
	2	WHOPAR15		2	WHOPAR55
	3	WHOPAR17		3	WHOPAR57
	4	WHOPAR19		4	WHOPAR59
	5	WHOPAR21		5	WHOPAR61
4	1	WHOPAR23	8	1	WHOPAR63
	2	WHOPAR25		2	WHOPAR65
	3	WHOPAR27		3	WHOPAR67
	4	WHOPAR29		4	WHOPAR69
	5	WHOPAR31		5	WHOPAR71

Variable labels – The variable labels included on the dataset were also initially derived from the CAPI program. In exporting the SPSS dataset from Blaise, labels based on the wording of questions were automatically allocated. Subsequently, these have been individually reviewed and, where necessary, modified in an effort to ensure that labels are comprehensible and accurate.

Again, particular problems occurred where a question was repeated (eg: the same question is asked for each birth, relationship, job, qualification, etc reported). When initially created, the Blaise-generated dataset had identical labels for each repeat of the question. In revising these labels, efforts have been made to indicate which variables relate to which birth, relationship, job, qualification, etc.

Again, an example based on the repeated question: “Who is the other parent of (name of baby)?” is given below. Within the label, the “(P1)”, “(P2)”, etc identify the first reported pregnancy, second reported pregnancy, etc.; and the “baby1”, “baby2”, etc identify the first, second, etc baby reported as conceived. It is important to note that information on pregnancy history was gathered by starting with the most recent pregnancy. Similar conventions are used for the other histories within the dataset.

Variable labels for repeated question: “Who is the other parent of (name of baby)?”

NB: Pregnancy 1 is the most recent pregnancy

Pregnancy	Baby	Variable name	Label
1	1	WHOPARB	(P1) Who is baby1s other parent
2	1	WHOPARB6	(P2) Who is baby1s other parent
3	1	WHOPAR13	(P3) Who is baby1s other parent
4	1	WHOPAR23	(P4) Who is baby1s other parent
5	1	WHOPAR33	(P5) Who is baby1s other parent
6	1	WHOPAR43	(P6) Who is baby1s other parent
7	1	WHOPAR53	(P7) Who is baby1s other parent
8	1	WHOPAR63	(P8) Who is baby1s other parent

Value labels – The value labels are also similarly derived from the Blaise program and have similarly been reviewed and, where necessary, modified in an effort to ensure that labels are comprehensible and accurate.

Missing values – In general, the use of CAPI/CASI has meant that missing data is less common than in earlier NCDS/BCS70 surveys. Missing values are not identified as such within the initial dataset (“declared missing” within SPSS), nor are they labelled. However, “refused”, “don’t know”, “not answered” and “not applicable” have been given consistent values and should be readily distinguishable:

Missing values (unless otherwise labelled)

7, 97, 997, 9997, 99997, 999997	=	Refused
8, 98, 998, 9998, 99998, 999998	=	Don't know
9, 99, 999, 9999, 99999, 999999	=	Not answered
. (sysmis)	=	Not applicable

Variable order – The order in which variables appear in the dataset will broadly follow the order of sections, and of questions within sections of the survey instruments.

However, the order is determined by the structure of the Blaise program, which does not necessarily hold each question in the order in which they are put to the respondent. As the example given in the table below illustrates, the sequence of variables in the dataset relating to drug use does not follow the same order as the questions on drug use in the self-completion (and as shown in: National Centre for Social Research NCDS/BCS70 Team (2000) *NCDS/BCS70 CAPI Documentation* which also accompanies the data deposit). This change in order is typically, but not exclusively associated with question sequences which are repeated to produce grid-like data structures (eg: birth, relationship, job, qualification histories, etc).

Values – As this was a CAPI/CASI survey, the values should be within the specified range for each variable.

Consistency – Again, the use of CAPI/CASI should ensure that all filters have been correctly followed.

Further details of the content of the data set can be found by generating an SPSS ‘data dictionary’. An interactive version of this which facilitates key word/phrase searches of the content of the dataset also accompanies the data deposit:

*NCDS/BCS70 1999-2000 Follow-ups: Interactive Data Dictionary for
Combined NCDS/BCS70 SPSS dataset (based on the SPSS Data Dictionary)*

EXCLUDED VARIABLES

A number of variables have been removed from the dataset originally derived from the CAPI/CASI program in order to ensure that the anonymity of cohort members is preserved.

Self-completion (CASI) and dataset order of variables relating to drug use

Self-completion (CASI) order		Dataset order		
Variable	Label	Variable	Sequential position	Label
CANNABIS	(SC) Have you ever tried cannabis?	DRUG	6360	(SC) Name of 1st other drug not mentioned
ECSACY	(SC) Have you ever tried ecstasy?	DRUG12	6365	(SC) Taken 1st other drug in the last 12 mths?
AMPHET	(SC) Have you ever tried amphetamines?	MORE31	6366	(SC) Taken more than 1 other drug?
LSD	(SC) Have you ever tried LSD?	DRUG2	6367	(SC) Name of 2nd other drug not mentioned
POPPER	(SC) Have you ever tried amyl nitrate?	DRUG13	6372	(SC) Taken 2nd other drug in the last 12 mths?
MAGMUSH	(SC) Have you ever tried magic mushrooms?	MORE32	6373	(SC) Taken more than 2 other drugs?
COCAINE	(SC) Have you ever tried cocaine?	DRUG3	6374	(SC) Name of 3rd other drug not mentioned
TEMAZ	(SC) Have you ever tried temazepan?	DRUG14	6379	(SC) Taken 3rd other in the last 12 mths?
SEMERON	(SC) Have you ever tried semeron?	MORE33	6380	(SC) Taken more than 3 other drugs?
KETAMINE	(SC) Have you ever tried ketamine?	DRUG4	6381	(SC) Name of 4th other drug
CRACK	(SC) Have you ever tried crack?	DRUG15	6386	(SC) Taken 4th other drug in the last 12 mths?
HEROIN	(SC) Have you ever tried heroin?	MORE34	6387	(SC) Taken more than 4 other drugs?
METHAD	(SC) Have you ever tried methadone?	DRUG5	6388	(SC) Name of 5th other drug
OTHDRUG	(SC) Have you tried any other illegal drugs?	DRUG16	6393	(SC) Taken 5th other drug in the last 12 mths?
DRUG	(SC) Name of 1st other drug not mentioned	MORE35	6394	(SC) Taken more than 5 other drugs?
DRUG12	(SC) Taken 1st other drug in the last 12 mths?	DRUG6	6395	(SC) Name of 6th other drug
MORE31	(SC) Taken more than 1 other drug?	DRUG17	6400	(SC) Taken 6th other drug in the last 12 mths?
DRUG2	(SC) Name of 2nd other drug not mentioned	MORE36	6401	(SC) Taken more than 6 other drugs?
DRUG13	(SC) Taken 2nd other drug in the last 12 mths?			[Other CASI variables here]
MORE32	(SC) Taken more than 2 other drugs?	CANNABIS	6526	(SC) Have you ever tried cannabis?
DRUG3	(SC) Name of 3rd other drug not mentioned	ECSACY	6527	(SC) Have you ever tried ecstasy?
DRUG14	(SC) Taken 3rd other in the last 12 mths?	AMPHET	6528	(SC) Have you ever tried amphetamines?
MORE33	(SC) Taken more than 3 other drugs?	LSD	6529	(SC) Have you ever tried LSD?
DRUG4	(SC) Name of 4th other drug	POPPER	6530	(SC) Have you ever tried amyl nitrate?
DRUG15	(SC) Taken 4th other drug in the last 12 mths?	MAGMUSH	6531	(SC) Have you ever tried magic mushrooms?
MORE34	(SC) Taken more than 4 other drugs?	COCAINE	6532	(SC) Have you ever tried cocaine?
DRUG5	(SC) Name of 5th other drug	TEMAZ	6533	(SC) Have you ever tried temazepan?
DRUG16	(SC) Taken 5th other drug in the last 12 mths?	SEMERON	6534	(SC) Have you ever tried semeron?
MORE35	(SC) Taken more than 5 other drugs?	KETAMINE	6535	(SC) Have you ever tried ketamine?
DRUG6	(SC) Name of 6th other drug	CRACK	6536	(SC) Have you ever tried crack?
DRUG17	(SC) Taken 6th other drug in the last 12 mths?	HEROIN	6537	(SC) Have you ever tried heroin?
MORE36	(SC) Taken more than 6 other drugs?	METHAD	6538	(SC) Have you ever tried methadone?
		OTHDRUG	6539	(SC) Have you tried any other illegal drugs?

Some useful variables

Variables included on the initial dataset, which may be of particular value to users are identified below.

Some useful variables

Information	Variables
Survey elements completed	UNOUT, INTWHO
NCDS or BCS70 respondents	SAMPLE
Self-completion	UNOUT, CASIINT (see below)
Proxy interviews	UNOUT, IFPROXY (see below)
Identifiers	NSERIAL/KEY/BSERIAL (see below)
Sex	DMSEX, CMSEX
Ethnic group	ETHNIC
Date of interview	INTDATE/DATEINT (see below)
Derived variables within CAPI	See below

Additional information about the self-completion, proxy interviews, identifiers and derived variables used within CAPI is given below, along with additional guidance for users.

Self-completion

The self-completion (CASI) was administered at the end of the interview. The interviewer handed the laptop computer used for the interview to the cohort member and explained how they should complete the questionnaire. Where the cohort member was unable or reluctant to use the laptop, the interviewer assisted, and if necessary administered the self-completion as an interview.

The variables which hold the data for the self-completion are identifiable through their labels – all begin with the endorsement “(SC)”. These variables are identified below.

Variables which hold data for the Self-completion (CASI) (ordered left to right)

DRUG	DRUG12	MORE31	DRUG2	DRUG13	MORE32	DRUG3
DRUG14	MORE33	DRUG4	DRUG15	MORE34	DRUG5	DRUG16
MORE35	DRUG6	DRUG17	MORE36	LR1	AR1	E1
A1	C1	PC1	L1	MOR1	LR2	WM1
IT1	WE1	MOR2	HAPPYREL	OUTTOG	OUTALONE	POUTALON
SAMEPART	WISHP1	WISHP2	WHOCOOKS	WHOSHOPS	WHOCLEAN	WHOWASH
WHODIY	WHOCASH	WHOTENDS	WHOTEACH	WHOCARES	A2	MOR3
LR3	AR2	C2	L2	PC2	LR4	E2
A3	WM2	AR3	GHQ1	GHQ2	GHQ3	GHQ4
GHQ5	GHQ6	GHQ7	GHQ8	GHQ9	GHQ10	GHQ11
GHQ12	SKILL1A	SKILL1B	SKILL2A	SKILL2B	SKILL3A	SKILL3B
SKILL4A	SKILL4B	SKILL5A	SKILL5B	SKILL6A	SKILL6B	SKILL7A
SKILL7B	SKILL8A	SKILL8B	SKILL9A	SKILL9B	MAL01	MAL02
MAL03	MAL04	MAL05	MAL06	MAL07	MAL08	MAL09
MAL10	MAL11	MAL12	MAL13	MAL14	MAL15	MAL16
MAL17	MAL18	MAL19	MAL20	MAL21	MAL22	MAL23
MAL24	LR5	AR4	IT2	MOR4	A4	E3
WE2	LR6	WM4	L3	C3	A5	SUSPSCH
NUMSUSPS	EXCLSCH	NUMEXCLS	TRUANT	POLICE1	POL1NUM	POLICE2
POL2NUM	POLICE3	POL3NUM	POLICE4	POL4NUM	POLICE5	POL5NUM
COURT	COURTNUM	CANNABIS	ECSACY	AMPHET	LSD	POPPER
MAGMUSH	COCAINE	TEMAZ	SEMERON	KETAMINE	CRACK	HEROIN
METHAD	OTHDRUG	WM3	LR7	MOR5	PC3	L4
AR5	MOR6	A6	IT4	WE3	WM5	C4
IT5	EFFICAC1	EFFICAC2	EFFICAC3	LIFESAT1	LIFESAT2	THANK

Further details of the self-completion are to be found in the following, which also accompanies the data deposit:

NCDS/BCS70 1999-2000 Follow-ups: Technical Report

Proxy interview

As noted above, where the cohort member was unable to understand or respond to questions put by the interviewer or to the self-completion, short proxy interviews were undertaken with a family member or carer. The variables which hold the data for the proxy interview are identifiable through their labels – all begin with the endorsement “(Proxy)”. These variables are identified below.

Variables which hold data for the Proxy Interview (ordered left to right)

PROXTHK	PROXYWHO	PROXYAC1	PROXYAC2	PROXYAC3	PROXYAC4	LSIIMWK
LSIAGE	LSILIM	MORE11	LSIIMWK2	LSIAGE2	LSILIM2	MORE12
LSIIMWK3	LSIAGE3	LSILIM3	MORE13	LSIIMWK4	LSIAGE4	LSILIM4
MORE14	LSIIMWK5	LSIAGE5	LSILIM5	MORE15	LSIIMWK6	LSIAGE6
LSILIM6	MORE16	LSIIMWK7	LSIAGE7	LSILIM7	MORE17	LSIIMWK8
LSIAGE8	LSILIM8	MORE18	LSIIMWK9	LSIAGE9	LSILIM9	MORE19
LSIIMW10	LSIAGE10	LSILIM10	MORE20	HOAGE	HOSUP1	HOSUP2
HOSUP3	HOMORE	HOAGE2	HOSUP4	HOSUP5	HOSUP6	HOMORE2
HOAGE3	HOSUP7	HOSUP8	HOSUP9	HOMORE3	HOAGE4	HOSUP10
HOSUP11	HOSUP12	HOMORE4	HOAGE5	HOSUP13	HOSUP14	HOSUP15
HOMORE5	HOAGE6	HOSUP16	HOSUP17	HOSUP18	HOMORE6	HOAGE7
HOSUP19	HOSUP20	HOSUP21	HOMORE7	HOAGE8	HOSUP22	HOSUP23
HOSUP24	HOMORE8	HOAGE9	HOSUP25	HOSUP26	HOSUP27	HOMORE9
HOAGE10	HOSUP28	HOSUP29	HOSUP30	HOMORE10	YEARIN	YEARM
TENURE	MARSTAT	PREGPROX	PROXYTYP	PROXYECO	CJSUP	CJEMPS
AGELFTED	ACTAGEL	FURTHED	LFTMORED	EDQUALS	EDQTYP01	EDQTYP03
EDQTYP04	EDQTYP05	EDQTYP06	EDQTYP07	EDQTYP08	EDQTYP09	EDQTYP10
EDQTYP11	EDQTYP12	EDQTYP13	VOCQUALS	VOCTYP01	VOCTYP02	VOCTYP03
VOCTYP04	VOCTYP05	VOCTYP06	VOCTYP07	VOCTYP08	VOCTYP09	VOCTYP10
VOCTYP11	LSIANY	LSIREG	HEIGHT	HTMETRES	HTCMS	HTFEET
HTINCHES	WEIGHT	WTKILOS	WTSTONES	WTPOUNDS	IFPROXY	NMPRX
NMPRX2	NMPRX3	NMPRX4	NMPRX5	NMPRX6	NMPRX7	NMPRX8
NMPRX9	NMPRX10	SEXPRX	SEXPRX2	SEXPRX3	SEXPRX4	SEXPRX5
SEXPRX6	SEXPRX7	SEXPRX8	SEXPRX9	SEXPRX10	AGEPRX	AGEPRX2
AGEPRX3	AGEPRX4	AGEPRX5	AGEPRX6	AGEPRX7	AGEPRX8	AGEPRX9
AGEPRX10	MSPRX	MSPRX2	MSPRX3	MSPRX4	MSPRX5	MSPRX6
MSPRX7	MSPRX8	MSPRX9	MSPRX10	RELPRX	RELPRX2	RELPRX3
RELPRX4	RELPRX5	RELPRX6	RELPRX7	RELPRX8	RELPRX9	RELPRX10
CMSEXPX	BDAT1PX	BDAT2PX	CMNAMPRX	RENAMPRX	RESIDPRX	NORMPRX

Further details of the proxy interview are to be found in Appendix 2, and the following, which also accompanies the data deposit:

NCDS/BCS70 1999-2000 Follow-ups: Technical Report

Identifiers

NCDS and BCS70 have unique identifiers which appear on the datasets already lodged with the UK Data Archive. These identifiers are also to be found on the new dataset.

- **NCDS** – The NCDS identifier is given by the variable NSERIAL. This may be used to match the new data with existing datasets where the same unique identifier will have the variable name SERIAL.
- **BCS70** – Two BCS70 identifiers appear on the dataset:
 - **KEY** – Whenever possible this variable should be used to match the new data with existing BCS70 datasets (*ie*: any dataset which holds the identifier KEY). It may also be used to match with existing datasets in which the unique identifier is given by a combination of ‘CHES number’ and ‘Twincode’. The former will normally be given the variable name: CHESNO; whilst the latter may variously appear as: TC, CTC, TC2, TC10, etc. To facilitate the match it is important to know that:

$$\text{KEY} = (\text{CHESNO} \times 10) + \text{TC}$$

$$\text{CHESNO} = \text{trunc}(\text{KEY}/10)$$

$$\text{TC} = \text{mod}(\text{KEY}, \text{CHESNO})$$

In the initial deposit of data for the 1999-2000 follow-up some values of KEY were missing or erroneous. These problems have now been corrected. For further details see:

CLS Cohort Studies Data Note 1: Longitudinal Linkage in BCS70: Rationalising Case Identifiers

- **BSERIAL** – This was the identifier used during the 1999-2000 follow-up. It may be used to match the new data with existing BCS70 datasets which only use the identifier SERIAL (*eg*: data from the BCS70 1996 Postal Follow-up obtained from the UK Data Archive prior to January 2003).

A NOTE OF CAUTION

The initial dataset may be matched with data from earlier NCDS and BCS70 surveys using the unique identifiers included. However, it is important to note that, to date, efforts have concentrated on an internal (cross-sectional) review of the quality of the data, and although longitudinal linkage has been made, there have only been limited efforts to validate the link through longitudinal editing. Users merging new and old data are strongly advised to carry out their own checks on the validity of the longitudinal link. They should report the details of any problems encountered to the *User Support Group* via email (cohort@cls.ioe.ac.uk).

Reference dates for retrospective data/histories

Although NCDS and BCS70 are prospective longitudinal studies, the gap between follow-ups has ensured that each includes a number of retrospective questions, which focus on experience since the previous follow-up. The recent follow-ups were no exception; retrospective information was gathered on pregnancies, relationships, jobs, qualifications and health problems since the last major follow-ups as follows:

- **NCDS:** the reference date was March 1991, representing the time of the last major follow-up in 1991 (NCDS5) when the members of the cohort were aged 33 years.

This means that the retrospective histories only cover ages 33 to 41 or 42 (depending on date of interview). Similar histories for ages 16 to 33 are to be found in the NCDS4 and NCDS5 datasets. There is one exception. For those not interviewed at NCDS5, information about qualifications was gathered with reference to March 1981, representing the time of the 1981 follow-up (NCDS4) when the cohort were aged 23 years. In order to compile a complete history from age 16 it will be necessary to link data from the NCDS5 and NCDS6 surveys (and for some it will also be necessary to link data from the NCDS4 survey).

- **BCS70:** the reference date was April 1986, representing the time of the major follow-up in 1986 when the members of the cohort were aged 16 years.

This means that the retrospective data gathered during the recent follow-up provides histories covering the ages 16-29 or 30 (depending on date of interview). The BCS70 1996 Postal Follow-up was not used to define the reference date because it included very few retrospective questions, and because response to the survey was limited by the need to plan and implement the survey in a limited time interval. This was a consequence of the nature of the funding available.

Date of interview

As note above, fieldwork for the NCDS/BCS70 follow-ups took place between November 1999 and September 2000. Each interview included on the dataset was date-stamped by the laptop used by the interviewer to administer the CAPI/CASI instruments. Although, interviewers were asked at the start of the Blaise program to correct any erroneous date or time, subsequent checking revealed that a small number of interview dates were clearly wrong. Fortunately, it has been possible to correct the majority of these with reference to other survey records.

However, the dates for 9 interviews remain to be resolved. These are identified below, together with details of the fieldwork wave in which it the address was issued to interviewers. It will be seen that all but one interview was carried out as part of wave 3 which began in January 2000. Whilst it is tempting to make an educated guess about the correct date of interview, experience suggests that it is better to await confirmation.

Updates for these cases will be provided as soon as they are available.

Interview dates: Unresolved errors

	NSERIAL/ BSERIAL	INDATE	DATEINT	Wave	Wave start (approx)
NCDS	422081J	15061999	15061999	3	January 2000
	518148Y	20122000	20122000	1	November 1999
	860002E	09071999	09071999	3	January 2000
	860006P	03071999	03071999	3	January 2000
	860008T	28061999	28061999	3	January 2000
BCS70	01467071	27061999	27061999	3	January 2000
	01610017	03101999	03101999	3	January 2000
	03022037	12071999	12071999	3	January 2000
	04012039	21061999	21061999	3	January 2000

Queries about these problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

CAPI Derived variables

Routing respondents through the interview required the derivation of a number of summary variables within CAPI program. As the table below shows, these provide information in a number of areas, about the household, relationships, children, housing, economic activity and qualifications, which may be of value to users.

Some useful CAPI Derived variables

Variable	Label
HHSIZE	Total Number of people in hhld (Derived)
AGEP	Age of parner [derived]
SEXP	Sex of partner [derived]
NUMADCH	No: of adopted kids in hhld [derived]
ANYCHD	Whether any kids in hhld [derived]
CHD16F	Any kids aged 0-16yrs in hhld [derived]
CHD13F	Any kids aged 0-13yrs in hhld [derived]
CHDAGE3	Any kids aged 3+yrs in hhld [derived]
CHDAGE4	Any kids aged 4-15yrs in hhld [derived]
CHD5_16	Any kids aged 5-16 yrs in hhld [derived]
CHD16	Any kids 16+yrs in hhld [derived]
CHD0_6	(CM NCDS) No: of own kids in hhld aged 0-6yrs [derived]
OWNCHILD	Whether CM has any of their own kids in hhld [derived]
FATHIN	Whether CMs natural father lives in hhld [derived]
MOTHIN	Whether CMs natural mother lives in hhld [derived]
NSPOUSE	Whether CMs spouse in hhld [derived]
NPART	Whether CMs partner in hhld [derived]
NUMROOMS	Total no. rooms (derived)
MOVOUT-MOVOUT25	(A1-A25) Date moved out (derived)
MOVIN2-MOVIN25	(A2-A25) Date moved in (derived)
LIVIN-LIVIN9	(EP1-EP9) Date relationship started (derived)
LIVOUT-LIVOUT9	(EP1-EP9) Date relationship ended (derived)
NUMP	Total number of babies in ref period [derived]
LONESTRT-LONESTRT3	Date 1st-3rd period lp started (derived)
LONEEND-LONEEND7	Date 1st-3rd period lp ended (derived)
ABCHAGE-ABCHANGE8	Age of eldest-8th eldest child living elsewhere (derived)
STRTJOB-SRTJO10	PrevAct1-10: Date started activity (derived)
CSTRTJOB	Date started current activity (derived)
NUMGCSE	Total no. of GCSE quals CM has (derived)
NUMOLVL	Total no. of GCE O Levels CM has [derived]
NUMCSE	Tot no. of CSEs:EDCSE1+EDCSE2 [derived]
NUMASLVL	Total no. of A/S levels CM has (derived)
NUMGCSAS	Total no. of GCE A level/S level quals CM has (derived)

All CAPI derived variables include the word 'derived' within the label. A full list of CAPI derived variables is given below.

More details of the CAPI dereived variables are to be found in the following, which also accompanies the data deposit:

NCDS/BCS70 1999-2000 Follow-ups: CAPI Documentation

CAPI Derived Variables

HHSIZE	Total Number of people in hhld (Derived)	MOVIN3	(A3) Date moved in (derived)
BIRTHDAY	CM day of birthday (derived)	MOVOUT3	(A3) Date moved out (derived)
AGEP	Age of parner [derived]	MOVIN4	(A4) Date moved in (derived)
SEXP	Sex of partner [derived]	MOVOUT4	(A4) Date moved out (derived)
NUMADCH	No: of adopted kids in hhld [derived]	MOVIN5	(A5) Date moved in (derived)
ANYCHD	Whether any kids in hhld [derived]	MOVOUT5	(A5) Date moved out (derived)
CHD16F	Any kids aged 0-16yrs in hhld [derived]	MOVIN6	(A6) Date moved in (derived)
CHD13F	Any kids aged 0-13yrs in hhld [derived]	MOVOUT6	(A6) Date moved out (derived)
CHDAGE3	Any kids aged 3+yrs in hhld [derived]	MOVIN7	(A7) Date moved in (derived)
CHDAGE4	Any kids aged 4-15yrs in hhld [derived]	MOVOUT7	(A7) Date moved out (derived)
CHD5_16	Any kids aged 5-16 yrs in hhld [derived]	MOVIN8	(A8) Date moved in (derived)
CHD16	Any kids 16+yrs in hhld [derived]	MOVOUT8	(A8) Date moved out (derived)
CHD16N01	Child reference number [derived]	MOVIN9	(A9) Date moved in (derived)
CHD16N02	Child reference number [derived]	MOVOUT9	(A9) Date moved out (derived)
CHD16N03	Child reference number [derived]	MOVIN10	(A10) Date moved in (derived)
CHD16N04	Child reference number [derived]	MOVOUT10	(A10) Date moved out (derived)
CHD16N05	Child reference number [derived]	MOVIN11	(A11) Date moved in (derived)
CHD16N06	Child reference number [derived]	MOVOUT11	(A11) Date moved out (derived)
CHD16N07	Child reference number [derived]	MOVIN12	(A12) Date moved in (derived)
CHD16N08	Child reference number [derived]	MOVOUT12	(A12) Date moved out (derived)
CHD16N09	Child reference number [derived]	MOVIN13	(A13) Date moved in (derived)
CHD16N10	Child reference number [derived]	MOVOUT13	(A13) Date moved out (derived)
ADPTNU01	Adopted child reference number [derived]	MOVIN14	(A14) Date moved in (derived)
ADPTNU02	Adopted child reference number [derived]	MOVOUT14	(A14) Date moved out (derived)
ADPTNU03	Adopted child reference number [derived]	MOVIN15	(A15) Date moved in (derived)
ADPTNU04	Adopted child reference number [derived]	MOVOUT15	(A15) Date moved out (derived)
ADPTNU05	Adopted child reference number [derived]	MOVIN16	(A16) Date moved in (derived)
ADPTNU06	Adopted child reference number [derived]	MOVOUT16	(A16) Date moved out (derived)
ADPTNU07	Adopted child reference number [derived]	MOVIN17	(A17) Date moved in (derived)
ADPTNU08	Adopted child reference number [derived]	MOVOUT17	(A17) Date moved out (derived)
ADPTNU09	Adopted child reference number [derived]	MOVIN18	(A18) Date moved in (derived)
ADPTNU10	Adopted child reference number [derived]	MOVOUT18	(A18) Date moved out (derived)
CHD0_6	(CM NCDS) No: of own kids in hhld aged 0-6yrs [derived]	MOVIN19	(A19) Date moved in (derived)
OWNCHILD	Whether CM has any of their own kids in hhld [derived]	MOVOUT19	(A19) Date moved out (derived)
FATHIN	Whether CMs natural father lives in hhld [derived]	MOVIN20	(A20) Date moved in (derived)
MOTHIN	Whether CMs natural mother lives in hhld [derived]	MOVOUT20	(A20) Date moved out (derived)
NSPOUSE	Whether CMs spouse in hhld [derived]	MOVIN21	(A21) Date moved in (derived)
NPART	Whether CMs partner in hhld [derived]	MOVOUT21	(A21) Date moved out (derived)
LINENO	Line no. of person from hhld grid [derived]	MOVIN22	(A22) Date moved in (derived)
SEX	CM gender [derived]	MOVOUT22	(A22) Date moved out (derived)
AGE	CM age last birthday [derived]	MOVIN23	(A23) Date moved in (derived)
H SIZE	No: of people in hhld [derived]	MOVOUT23	(A23) Date moved out (derived)
YEARNOW	Year interview took place (derived)	MOVIN24	(A24) Date moved in (derived)
NUMROOMS	Total no. rooms (derived)	MOVOUT24	(A24) Date moved out (derived)
MOVOUT	(A1) Date moved out (derived)	MOVIN25	(A25) Date moved in (derived)
MOVIN2	(A2) Date moved in (derived)	MOVOUT25	(A25) Date moved out (derived)
MOVOUT2	(A2) Date moved out (derived)	LIVIN	(EP1) Date relationship started (derived)

CAPI Derived Variables (continued)

LIVOUT	(EP1) Date relationship ended (derived)	LINEGRI9	9th adopted child listed in hhld grid (derived)
LIVIN2	(EP2) Date relationship started (derived)	LINEGR10	1st child 16+ in hhld grid (derived)
LIVOUT2	(EP2) Date relationship ended (derived)	LINEGR11	2nd child 16+ in hhld grid (derived)
LIVIN3	(EP3) Date relationship started (derived)	LINEGR12	3rd child 16+ in hhld grid (derived)
LIVOUT3	(EP3) Date relationship ended (derived)	LINEGR13	4th child 16+ in hhld grid (derived)
LIVIN4	(EP4) Date relationship started (derived)	LINEGR14	5th child 16+ in hhld grid (derived)
LIVOUT4	(EP4) Date relationship ended (derived)	LINEGR15	6th child 16+ in hhld grid (derived)
LIVIN5	(EP5) Date relationship started (derived)	LINEGR16	7th child 16+ in hhld grid (derived)
LIVOUT5	(EP5) Date relationship ended (derived)	ABCHAGE	Age of eldest child living elsewhere (derived)
LIVIN6	(EP6) Date relationship started (derived)	ABCHAGE2	Age of 2nd eldest child living elsewhere (derived)
LIVOUT6	(EP6) Date relationship ended (derived)	ABCHAGE3	Age of 3rd eldest child living elsewhere (derived)
LIVIN7	(EP7) Date relationship started (derived)	ABCHAGE4	Age of 4th eldest child living elsewhere (derived)
LIVOUT7	(EP7) Date relationship ended (derived)	ABCHAGE5	Age of 5th eldest child living elsewhere (derived)
LIVIN8	(EP8) Date relationship started (derived)	ABCHAGE6	Age of 6th eldest child living elsewhere (derived)
LIVOUT8	(EP8) Date relationship ended (derived)	ABCHAGE7	Age of 7th eldest child living elsewhere (derived)
LIVIN9	(EP9) Date relationship started (derived)	ABCHAGE8	Age of 8th eldest child living elsewhere (derived)
LIVOUT9	(EP9) Date relationship ended (derived)	STRTJOB	PrevAct1: Date started activity (derived)
NUMP	Total number of babies in ref period [derived]	STRTJOB2	PrevAct2: Date started activity (derived)
LONESTRT	Date 1st period lp started (derived)	STRTJOB3	PrevAct3: Date started activity (derived)
LONEEND	Date 1st period lp ended (derived)	STRTJOB4	PrevAct4: Date started activity (derived)
LONESTR2	Date 2nd period of lone parenthood started (derived)	STRTJOB5	PrevAct5: Date started activity (derived)
LONEEND4	Date 2nd period of lone parenthood ended (derived)	STRTJOB6	PrevAct6: Date started activity (derived)
LONESTR3	Date 3rd period of lone parenthood started (derived)	STRTJOB7	PrevAct7: Date started activity (derived)
LONEEND7	Date 3rd period of lone parenthood ended (derived)	STRTJOB8	PrevAct8: Date started activity (derived)
LINEGRID	1st adopted child listed in hhld grid (derived)	STRTJOB9	PrevAct9: Date started activity (derived)
LINEGRI2	2nd adopted child listed in hhld grid (derived)	STRTJO10	PrevAct10: Date started activity (derived)
LINEGRI3	3rd adopted child listed in hhld grid (derived)	CSTRTJOB	Date started current activity (derived)
LINEGRI4	4th adopted child listed in hhld grid (derived)	NUMGCSE	Total no. of GCSE quals CM has (derived)
LINEGRI5	5th adopted child listed in hhld grid (derived)	NUMOLVL	Total no. of GCE O Levels CM has [derived]
LINEGRI6	6th adopted child listed in hhld grid (derived)	NUMCSE	Tot no. of CSEs:EDCSE1+EDCSE2 [derived]
LINEGRI7	7th adopted child listed in hhld grid (derived)	NUMASLVL	Total no. of A/S levels CM has (derived)
LINEGRI8	8th adopted child listed in hhld grid (derived)	NUMGCSAS	Total no. of GCE A level/S level quals CM has (derived)

FURTHER ASSESSMENT OF DATA QUALITY

In addition to the checks that were built into the CAPI code, or undertaken immediately after fieldwork, members of the CLS NCDS/BCS70 team have carried out further checks. In this task they have been greatly assisted by a number of researchers who are members of the NCDS/BCS70 Advisory Groups and/or their associates and, as such, are experienced in the use of data from previous NCDS/BCS70 follow-ups.

It can take some time to become familiar with a large and complex dataset, such as this, and it is important to stress that the checking continues. However, in assessing the quality of the new data, the CLS team and advisors were asked to concentrate on the aspects listed below.

Guide for Quality Assessment

- **Variable labels** – Check that these are present for all variables, comprehensible and accurate.
- **Value labels** – Check that these are present where appropriate, comprehensive, comprehensible and accurate.
- **Values** – Report all variables for which the values appear unusual/wrong in any way.
- **Consistency** – Report all instances of apparent inconsistency, *eg*:
 - Where the responses to a primary (filter) question (*eg*: “Do you have any of the qualifications on this card?”) and supplementary (filtered) question (*eg*: “IF YES, Which?”) are not consistent.
 - Where the respondent’s circumstances (*eg*: marital status, economic status) are not consistent throughout the dataset.
- **Missing values** – Report all instances where:
 - There are many missing cases.
 - Missing values are present but not declared or labelled.
- **Variable order** – Report all instances where confusion occurs because variables appear out of sequence.
- **Other problems** – Report any and all other problems encountered in using the dataset.
- **Derived variables** – Provide details of any derived variables developed which may be of value to other users and which may be considered for deposit with the UK Data Archive.

Some of the information presented above has drawn on lessons learned as a result of this quality assessment. Additional, and more detailed information and guidance is given below, organised by survey topic.

A list of the advisors and their associates who contributed to the assessment of the quality of the data, and whose work is drawn on below is given in Appendix 3. The contribution of individual advisors is acknowledged below. Wherever appropriate, the detailed documentation provided by advisors is included in Appendices 4-8.

Survey topics

Household
Housing
Relationships
Children
Family Social Relationships & Support
Family Income
Employment
Lifelong Learning
Health
Citizenship and Values
Self-completion (covering attitudes/especially sensitive topics)

HOUSEHOLD

Household Grid	
Household grid	Fairly standard grid. Gathers details of sex, age, relationship to respondent, marital status
Ethnicity	Based on the new question being developed for the 2001 Census
Language spoken in the home	Language usually spoken

Review of these data by the CLS NCDS/BCS70 Team and by advisors has, to date, revealed no particular problems, nor have any derived variables been developed. (A number of CAPI derived variables are available – see above).

However, it should be noted that the household grid includes information about the current de-facto marital status for all persons in the household. For a small number of cohort members, this information has been shown to be inconsistent with information about their legal marital status recorded later in the interview. Further details are reported below under 'Relationships'.

The data will continue to be investigated and, if necessary, detailed updates will be made available, and the dataset will be updated.

Queries about problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

REVISED DATA FOR THE HOUSEHOLD GRID (December 2002)

Following the initial data deposit further checking revealed a number of errors in the NCDS and BCS70 household grid data. These data have been corrected in the revised deposit. Further details are available in:

CLS Cohort Studies Data Note 3: Household Grid Variables in the Combined NCDS/BCS70 1999/2000 Data

HOUSING

Housing	
Current address	Rooms, heating, housing benefit, tenure
Intentions to move	Moving, future tenure, why moving
Property inheritance	Ever inherited house/flat, when
Homelessness	Experience of, number of times, date of last, why, where stayed, duration, applied to LA
Housing history	Works backwards from last address. Seeks details of: dates, tenure, living arrangements, reasons for moving.

Review of these data by the CLS NCDS/BCS70 Team and by advisors has, to date, revealed no particular problems, nor have any derived variables been developed.

The data will continue to be investigated and, if necessary, detailed updates will be made available, and the dataset will be updated.

Queries about problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

RELATIONSHIPS

Relationships	
Marital status	Legal status, prior cohabitation, dates of marriage/cohabitation, age, marital status of partner
Relationship history	Marriage/cohabiting. Works backwards from last relationship. Seeks details of: dates, sex, age, marital status, marriage, separation/divorce

Partnership status

(Based on work undertaken by Dr Ann Berrington, University of Southampton)

The following is based on notes made by Dr Berrington whilst exploring the marital status and partnership history data. A number of problems are identified and solution suggested. The data have not been modified in line with these suggested solutions. Users may make such changes if they wish.

1. Proxy interview – IF PROXY=1

For these 66 individuals no information about marital status is available.

However, information is available in the variables PROXYAC1, PROXYAC2, PROXYAC3 and PROXYAC4 on whether these cohort members had ever been married, had children, a job or education. Just 5 are reported as having been married.

2. Inconsistencies between legal and de-facto marital status

The documentation suggests that at the start of the interview current de-facto marital status is asked for all persons in the household. Later on in the questionnaire, cohort members are asked for their legal marital status. Cross-tabulation of these two variables (MS and MARSTAT2) suggests a number of inconsistencies.

- a) 140+1 =141 cases give a valid legal marital status but have missing information for de-facto marital status.

Use the information from the household grid to identify the de-facto marital status

(i) MARSTAT2=1, MS=missing (n=76)

Most of these people do not have a current partner in the household. i.e. NPART=0, NSPOUSE=0. Code these as MS=3. 4 cases have a current partner – code these as MS=2. However, only one of these four goes on to provide any information about their current partner

(ii) MARSTAT2=2 or 3, MS=missing (n=42)

Classify all those living with a spouse (NSPOUSE=1) as de-facto currently married (n=12). The remainder can be classified as de-facto separated (n=30). NOTE: This assumes that all there is no spouse who was missed on the household grid – see comment later on.

(iii) MARSTAT2=4 or 5 or 6, MS=missing

3 cases report a current partner in the household grid. 2 of the 3 provide details of this current partner. These 3 can be coded as cohabiting, the remainder can be coded as MS=MARSTAT2.

- b) 12 people gave a valid code for de-facto marital status but no information for legal marital status.

Code all of those who have a de-facto status of “single never married” as legal marital status “never married” (n=3).

Code those whose de-facto marital status is "divorced" as legal marital status "divorced" (n=1).

Code those whose de-facto marital status is "separated" as legal marital status "separated" (n=1). Use information from the partnership history data to identify the legal marital status of the 7 cases who are currently married or cohabiting. I.e. identify whether these individuals have been previously married.

- c) 24 cases report themselves as legally separated but de-facto married. 12 appear to have a partner in the household. We can't tell for definite whether these are married to their partner as CURPARTB is missing for all: code either as MS=1 or 2. The remaining 12 do not appear to have a partner: code as MS=4.
- d) 2 cases say that they are legally separated and say that they are single never married. In fact they were both previously legally married and experienced marital dissolution. One went on to have a current cohabiting partnership and should be coded as MS=2, whilst the other should be coded MS=4.
- e) 10 cases say that they are legally divorced but currently married. 7 of these appear to be currently cohabiting and could be coded MS=2. The remaining 3 do not have a current partner and so should be coded MS=5.
- f) 2 cases say that they are legally widowed and currently married. Both appear to be cohabiting and so should be coded MS=2.
- g) 3 cases say that they are legally divorced but currently widowed. One of these is cohabiting and should be coded MS=2. The other has no partner and should be coded as MS=5.
- h) 1 case says that they are legally widowed and their de-facto status is divorced. This person appears to be cohabiting and could be coded MS=2.

3. Inconsistencies between reported current de-facto marital status (MS) and the derived variable indicating whether a spouse or partner is present.

There are 316 cases who are either currently married or cohabiting (MS=1 or 2) where NPART=0 and NSPOUSE=0. It may be that in all of these cases there is a partner living elsewhere. For the analyses discussed below these 316 cases are deleted.

A total of 500 of those who report de-facto status separated, divorced or widowed (MS=4-6), also report a current partner in the household. All 500 provide a valid code for the year and month of their current partnership. These cases can be recoded as cohabiting MS=2.

4. Inconsistencies between CURPART and CURPARTC for those who did not live together before marriage.

There appear to be two start dates for the start of the current partnership among those who said that they did not live together before marriage. Was this on purpose – to act as a check? For 203 cases the year as reported in CURPART and CURPARTC was not the same. For some of these cases CURPARTC>CURPART suggesting perhaps a period of co-residence which was not considered "living together"?? In other cases CURPARTC<CURPART which though possible seems unlikely.

5. Years and months of start and end dates of partnerships.

The data appear to be very complete with very few missing cases for months of marriage and cohabitation. Given the obvious heaping of the dates, especially of cohabitation, on June (and September to a lesser extent) it would be helpful to clarify whether these months were estimated in the field by the respondent (as per the questionnaire instructions), whether they were estimated in the field by the interviewer, or whether there has been some post-field imputation.

As a result of this estimation there are inevitably some inconsistencies between the start of date of cohabitation and marriage. For example among those respondents who are currently married and who lived with their spouse prior to marriage:

88 cases give the same month and year for the start of cohabitation and marriage

30 individuals have a negative period of premarital cohabitation.

Queries about these problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Partnership Histories (BCS70 cohort members only)

(Based on work undertaken by Dr Ann Berrington, University of Southampton)

The following problem cases have been identified by Dr Berrington for BCS70 cases:

NB: The identifiers listed below are for the variable BSERIAL – the unique identifier used for BCS70 cohort members during the 1999-2000 follow-up

1. Union histories incomplete - don't tell us whether they had a previous partnership

04230096 06467062 06499041 10652074 10921001 11033011 15357009

2. Missing all partnership information

01962229

3. Separated/Divorced/Widowed no current partner and no information on previous partnership

00725069 01245013 01695003 03228043 05794018 06129076 07393091 08042081
09855079 11343019 12294077

4. Separated/Divorced/Widowed, MS=2,4,5 No information on current partnership.

00437064 01239038 01563047 01881229 02013058 02178069 02215063 02218089
03310042 03404044 03685084 03940034 04012039 04318022 04356077 04521026
04540003 05152000 05373082 06338031 07019000 07429086 08007052 08734096
09203084 09353042 10126060 10388022 11144040 11847006 13600077 14438031
14507004 14858093 15467063 16349062 17093038 17173063 17537068 17800044

5. Other Inconsistencies in partnership history which could not be resolved

00509063 01949055 02688031 03203014 04203093 04361077 05458083 06631036
07000047 08656021 09784055 10728074 10745000 11528098 13938039 14692092
16486018 18006056

Queries about these problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

CHILDREN

Children	
Pregnancy history	Works backwards from last pregnancy conceived/ fathered. Seeks details of: outcome, name, sex, birth weight, DOB, problems, smoking in pregnancy, other parent, whereabouts, support and circumstances of absent children
Lone parenthood	Periods of ≥ 1 month. Works forward for up to 4 periods. Seeks details of dates and numbers of children
Infertility	Ability to have and plans for (more) children
Adopted children	For up to 4 children. Seeks details of age of child on adoption and nature of adoption
Partner's children from a previous relationship	Number, whether seen or visit
Children over 16 (NCDS only)	Name, age, economic, marital and parent status
Absent formerly resident children (NCDS only)	Name, age, whereabouts, economic, marital and parent status; CM contact and payments for children absent from the household, but who were living with CM in 1991
Family activities	Things done as a family
Demands of parenting	Physical/emotional/time demands, worries, closeness of family

Pregnancy Histories

(Based on work undertaken by Dr Ann Berrington, University of Southampton and subsequently by the CLS NCDS/BCS70 Team)

In analysing the fertility histories, Dr Berrington noted that the data seems generally to be in good shape, if rather cumbersome to work with. However, she did note "...a significant minority who have reported consecutive pregnancies in one confinement...".

As noted above, during the interview, information was gathered for each child conceived in each pregnancy, starting with the most recent and working back through each earlier pregnancy. This means that the birth dates recorded in the most recent pregnancy must be later than those recorded in the second most recent, etc. It also means that within any one pregnancy, the dates recorded for each child conceived should be similar if not the same.

Further investigation by the CLS NCDS/BCS70 Team suggests that Dr Berrington was indeed correct to be concerned. For some cohort members, the dates of successive pregnancies appear to be recorded in the variables reserved for multiple outcomes of a single pregnancy. This appears to be true for some 261 cases (154 BCS70 and 107 NCDS), although in some instances the apparent error may reflect the timing of a mix of outcomes, including live births, stillbirths and miscarriages. For some, there may also be a problem with the order of pregnancies which do not appear to start with the most recent and work backward.

Following further investigation, a list of detailed updates will be made available, and the dataset will be updated. In the meantime, a list of the cases which currently appear to be effected by this problem is given below.

NB: The identifiers listed below are for the variables NSERIAL and BSERIAL

NCDS

044014L	045002J	045011K	050059U	052008M	053003F	056052M	085010J	087029S
091017J	092206S	093039E	093052W	094028E	099010M	099050Z	100009C	110213J
110244V	120017Q	120080X	120141R	180007A	181038T	183016T	188030P	200075Z
218005B	235011T	280005D	280064W	282113T	285061S	286044X	380041R	380090E
385005P	400075Q	440006V	481001R	500125L	500376N	504034D	506014J	510097Q
512002L	512012P	513033C	513073R	513143L	514012Z	515008Q	517045H	520056F
528014H	528021D	550188U	560008V	610087U	630006D	650034X	650212V	710034D
720027P	750043F	823501Q	825073F	850027R	850032J	860001C	880017H	882021J
882051T	933019C	937020J	950123U	950168T	950290Q	980004E	982053E	984024J
985067J	986100J	986205Y	986250D	986431J	X10020X	X25101F	X32005A	X41020X
X60006M	X66030R	X71009E	X77014E	X77016K	X77021B	X78086N	X82347P	Y00087K
Y00293N	Y01189Y	Y20134D	Y20154L	Y30150J	Y30168D	Y30169F	Y31043N	

BCS70

00279065	00387068	00530014	00607090	00700040	01065011	01236012	01251089	01254014
01266041	01618019	01763001	02294098	02346044	02441045	02454047	02880084	02914028
02934081	02939057	03085020	03264047	03366075	03371075	03575030	04046069	04239073
04394031	05014068	05036071	05289032	05399086	05408250	05521004	05533031	05542032
05585063	05794018	06031073	06072053	06092005	06123024	06145027	06183082	06266032
06305077	06466087	06481063	06501030	06570089	06641012	06666041	06707036	06797022
06853094	06959022	07052079	07244007	07342034	07416084	07469091	07494043	08052057
08078061	08239086	08263063	08283015	08302007	08409011	08445015	08460092	08667073
08716094	08787002	08818021	08930049	08962028	08995083	09054086	09145062	09204059
09261091	09262066	09310011	09318013	09324089	09490099	09495075	09573049	09591051
09716072	09803047	09840026	09891083	10047010	10215086	10455045	10477048	10527044
10582025	10657050	10675052	10732099	10937029	11032036	11047089	11162042	11212038
11487002	11636000	11706049	12091097	12155070	12434074	12609076	12958013	13189078
13561032	13930037	13945090	14137024	14293058	14538008	14671064	14687092	14858093
14869044	14978022	15084081	15093082	15440084	15455036	15511007	15527035	15567040
15571065	15737066	15810064	16102076	16152057	16224056	16354062	16568094	16630040
16717092	16822069	17386019	18013006	18046061	20043000	20091000	20212000	20305000
21667000								

Queries about these problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

REVISED DATA FOR THE PREGNANCY HISTORIES (December 2002)

Following the initial data deposit the problems with the pregnancy histories identified above, and others revealed by further checking have been corrected. The updates are incorporated into the revised deposit. Further details are available in:

CLS Cohort Studies Data Note 2: Pregnancy Histories in the Combined NCDS/BCS70 1999/2000 Data

FAMILY SOCIAL RELATIONSHIPS & SUPPORT

Family, Social Relationships & Support	
Contact with family	Parents alive/contact/close/divorce/worries/help. Contact with siblings/in-laws/grandparents
Emotional support	Sources/nature of emotional support

Review of these data by the CLS NCDS/BCS70 Team and by advisors has, to date, revealed no particular problems, nor have any derived variables been developed.

The data will continue to be investigated and, if necessary, detailed updates will be made available, and the dataset will be updated.

Queries about problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

FAMILY INCOME

Family Income	
Other Income	Income from benefits and regular income from other sources received by respondent and/or partner
Financial situation	Organisation of household money and money problems

(Based on work undertaken by Dr Gaëlle Pierre, Institute for Employment Research, University of Warwick)

In reviewing the data on 'Family income', Dr Pierre has reported that, for some BCS70 cases, values for amount of benefit received (BENAMT- BENAMT25) and the amount of other income (INCAMT- INCAMT12) appear quite high.

This will be subject to further investigation by the CLS NCDS/BCS70 Team, following which a list of necessary updates will be made available, and the dataset will be updated. In the meantime, users should use these data with some caution.

Queries about these problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

EMPLOYMENT

Employment	
Economic activity	Fairly standard question
Current job	Seeks details of: earnings, hours, fringe and Other benefits, pensions, prospects
Other paid work	Weekly earnings from odd/casual jobs
Currently unemployed	How became unemployed and job search
Labour market histories	Periods in a job or not in a job lasting ≥ 1 month. Works backwards. Seeks details of: circumstances and dates. If in job gathers details of job, employer, responsibilities (for SOC and SIC coding)
Partner's job	Age left full-time education, economic status, job, employer, responsibilities (for SOC and SIC coding), earnings

Earnings and hours worked

(Based on work undertaken by Dr Gaëlle Pierre, Institute for Employment Research, University of Warwick)

In reviewing the data on employment, Dr Pierre has noted the following:

1. Some respondents report numbers of hours worked below 30 while reporting f/t employment (for both employees and self-employed).
2. Some of the observations on pay appear very low/high.

Further information about earnings are given below.

3. Some of the Standard Industrial Classification (SIC) codes are not in the SIC classification, and some are missing.

Further details of the SIC coding are to be found in the following which also accompanies the data deposit:

National Centre for Social Research and Centre for Longitudinal Studies
NCDS/BCS70 Teams (2001) *National Child Development Study, British Cohort Study 1970: Technical Report*

4. Some of the previous activities have the same start dates.
5. Some of the start years for previous activities are equal to 9999 or 9998. As noted above, the following missing values will be found within the data:

Missing values (unless otherwise labelled)

7, 97, 997, 9997, 99997, 999997	=	Refused
8, 98, 998, 9998, 99998, 999998	=	Don't know
9, 99, 999, 9999, 99999, 999999	=	Not answered
. (sysmis)	=	Not applicable

6. Some of the start years for the current activity go back to when the respondent was younger than 16. One observation reports a start year of 1958 for a BCS70 respondent.

These problems will be subject to further investigation by the CLS NCDS/BCS70 Team, following which a list of necessary updates will be made available, and the dataset will be updated. In the meantime, users should use these data with some caution.

Queries about these problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Earnings

(Based on work undertaken by Lorraine Dearden and Alissa Goodman, Institute for Fiscal Studies)

In reviewing the data on earnings, Lorraine Dearden and Alissa Goodman have identified a number of problems and produced a number of very valuable revised and derived variables. These are briefly described below. Further details of all variables are to be found in the detailed documentation supplied by Lorraine Dearden and Alissa Goodman, which is included as Appendix 4.

Section 1. Interview date

1. Corrected interview date

Name: lintdate
Description: Corrected interview date
Purpose: To correct interview dates that were originally coded as outside the interviewing period. In the original data there were 23 individuals whose interview dates were not between 10/99 and 9/00.

Section 2. Deflators

2. RPI at interview date

Name: rpi_int
Description: Retail price index (RPI) for month of interview, where Jan 2001 = 1.000.
Purpose: To allow wages, incomes etc. to be converted into other prices. To uprate wages to Jan 2001 prices, divide by this deflator.

Section 3. Employment variables

3. Total usual weekly hours

Name: Hours
Description: Total usual weekly hours, including paid and unpaid overtime, not including meal breaks.
Purpose: To derive a measure of total hours usually worked, in order to convert gross and net pay information available into an hourly pay measure.

4. Imputed/Corrected net last pay

Name: lnetpay
Description: Imputed or corrected net last pay variable.
Purpose: Clean last net pay variable, corrected for implausible values.

5. Imputed/Corrected net period

Name: lnetprd
Description: Imputed or corrected net last pay PERIOD variable.
Purpose: Clean last net pay PERIOD variable, corrected for implausible values and imputed from gross where net unavailable

6. Imputed/Corrected gross pay

Name: lgropay
Description: Imputed or corrected gross last pay variable.
Purpose: Clean last gross pay variable, corrected for implausible values and imputed from gross where net unavailable

7. Imputed/Corrected gross period

Name: lgroprd
Description: Imputed or corrected gross last pay PERIOD variable.

Purpose: Clean last gross pay PERIOD variable, corrected for implausible values and imputed from gross where net unavailable

8. Hourly net last pay

Name: hr_net

Description: Hourly net last pay.

Purpose: Clean hourly equivalent last net pay. Where net pay variables are missing but gross pay variables are available, net pay has been imputed from gross using known parameters of the tax system in the relevant year.

9. Annual net last pay

Name: ann_net

Description: Clean annual equivalent net last pay. Where net pay variables are missing but gross pay variables are available, net pay has been imputed from gross using known parameters of the tax system in the relevant year.

Purpose: Clean annual equivalent last net pay, with missing values imputed wherever possible.

10. Hourly gross last pay

Name: hr_gro

Description: Clean hourly equivalent last gross pay. Where gross pay variables are missing but net pay variables are available, gross pay has been imputed from net using known parameters of the tax system in the relevant year.

Purpose: Clean hourly equivalent gross last pay, with missing values imputed wherever possible.

11. Annual gross last pay

Name: ann_gro

Description: Clean annual equivalent last gross pay. Where gross pay variables are missing but net pay variables are available, gross pay has been imputed from net using known parameters of the tax system in the relevant year.

Purpose: Clean annual equivalent gross last pay with missing values imputed wherever possible.

Section 4. Partner's earnings variables

12. Partner's Imputed/Corrected last net pay

Name: lpnetpay

Description: Partner's last net pay variable, implausible values corrected.

Purpose: Clean partner's last net pay.

13. Partner's Imputed/Corrected last net pay PERIOD

Name: lpnetprd

Description: Partner's last net pay PERIOD variable, implausible values corrected.

Purpose: Clean partner's last net pay PERIOD.

14. Partner's Imputed/Corrected last net pay PERIOD

Name: lpnpred

Description: Partner's last net pay PERIOD variable, implausible values corrected.

Purpose: Clean partner's last net pay PERIOD.

15. Partner's annual last net pay

Name: ann_pnet

Description: Annual equivalent of partner's last net pay, implausible values corrected.

Purpose: Clean annual equivalent partner's last net pay

16. Partner's IMPUTED annual last gross pay

Name: ann_pgro

Description: Annual equivalent of partner's last gross pay, imputed from partner's last net pay using the known parameters of the tax system in the relevant year.

Purpose: Clean annual equivalent partner's last gross pay - imputed from net.

Section 5. Intermediate variables – used during cleaning

17. Type of fix to pay variable and period codes

Name: fixtype

Description: Categorical variable containing the type of coding error (if any) which we have identified in the pay information of each employee, and the fix which has been implemented.

Purpose: To identify and record corrections to last net and gross pay variables.

18. Indicator of odd proportion of gross to net pay

Name: oddprop

Description: Indicator of whether ratio of gross to net pay suggests coding error in some of the pay variables which REMAINS UNCORRECTED.

Purpose: To identify odd wage information

19. Indicator of wage variable probably too high or too low

Name: oddwage

Description: Indicator of whether pay variables remain apparently implausibly high or low AFTER CORRECTIONS have been implemented.

Purpose: To identify odd wage information

20. Missing hours variable

Name: zhours.

Description: Indicator of Usual weekly hours variable (hours) missing.

21. Missing hourly net pay indicator

Name: zhr_net.

Description: Indicator of if/ why net hourly pay variable is missing.

22. Missing hourly gross pay indicator

Name: zhr_gro.

Description: Indicator of if/ why gross hourly pay variable is missing.

23. Odd partner's wage indicator

Name: podd.

Description: Indicator of apparently implausibly high or low partner's net wage, remaining AFTER CORRECTIONS have been implemented.

Queries about these problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

LIFELONG LEARNING

Lifelong Learning	
Qualifications	Age left full-time education and educational and vocational qualifications held (subjects, grades, dates awarded, where studied)
Current course for qualification	Qualification, subject(s), date started, where studied
Assessment of current/most recent course	Why taken and expected/experienced benefits
Other courses and training	Number, why taken and expected/experienced benefits
No formal learning	Reasons why no learning
Learning overview	Useful and enjoyable periods of learning
Contact with information technology	Use of computers at home and at work
Literacy and numeracy	Problems with reading, writing and maths, implications and courses to improve

Age left education

(Based on work undertaken by:

- *Dr Gaëlle Pierre, Institute for Employment Research, University of Warwick)*
- *Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London)*

In reviewing the data Lifelong Learning, Dr Pierre has noted problems with the data on the age of leaving full-time education: "...several answers are >30 for NCDS and some NCDS respondents say "still in f/t education...".

Samantha Parsons has reviewed these and related data and produced the revised and derived variables are identified below. Further details of all variables are to be found in the detailed documentation supplied by Samantha Parsons, which is included as Appendix 5.

Details of further problems identified by Samantha in relation to other aspects of Lifelong Learning and her suggested solutions are summarised below. Details of problems in other areas of the data are given in the sections dealing with 'Health' (Diet, food and exercise; Eating problems; Smoking; Drinking; and Accidents), and with the 'Self-completion' (Skills and Illegal drug use).

Age left full-time education

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	N	%	N	%	N	%	
actage12	607	2.7			^1	0	22072
<i>actage12 details age CM left full-time education</i>							
agelfte2	66	0.3	6	0	36	0.2	22572
<i>agelfte2 captures CMs who were in full-time education at time of interview</i>							
furthered2	607	2.7					22073
<i>furthered2 details CMs who returned to full-time education less than 3 years after initially leaving full-time education</i>							
lftmore2	20522	90.5			^2		2156
<i>lftmore2 details age CM left second period of full-time education</i>							
plefted	5981	26.4	^12	0.1	^347	2.1	16340
<i>plefted details age CM partner left full-time education</i>							

Derived AGE LEFT FULL-TIME EDUCATION Variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	N	%	N	%	N	%	
Ageftedu	109	0.5					22571
<i>6 category variables combining information from actagel2 and agelfte2. Values: 1'before 16' 2'at 16' 3'post 16' 4'post 18' 5'post 21' 6'still in ft edu'.</i>							
Agelfted	109	0.5					22571
<i>continuous variable combining information from actagel2, agelfte2, furthered2 and lftmore2. Value range: 14 to 42, 99=still in education</i>							
Ageledu	109	0.5					22571
<i>6 category variable collapsing information in agelfted. Values: 1'before 16' 2'at 16' 3'post 16' 4'post 18' 5'post 21' 6'still in ft edu'</i>							
plefted1	6340	28.0					16340
<i>5 category variable collapsing information in plefted. Values: 1'before 16' 2'at 16' 3'post 16' 4'post 18' 5'post 21'</i>							

Queries about these problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Qualifications

Problems with reporting of GCSEs

(Based on work undertaken by Andrew Jenkins, Institute of Education, University of London)

In reviewing the data on qualifications, Andrew Jenkins identified problems with the reporting of GCSE qualifications. The nature of the problem is described below in a summary of the report prepared by Andrew. The full report is included as Appendix 6.

1. The Extent of the Problem

The earliest date at which it is possible to have obtained a GCSE qualification is Summer 1988, yet some 94 NCDS cohort members and over 1,000 BCS70 cohort members (CMs) report having obtained one or more GCSEs before 1988. To try to resolve this apparent inconsistency, the data for all of the 94 NCDS CMs and the first 200 BCS CMs have been checked against earlier sweeps of the NCDS/BCS surveys where information on qualifications is also available.

2. GCSE Date Errors for the NCDS

In the 2000 sweep of the data, those NCDS CMs who had participated in the previous sweep in 1991 were asked to supply information on all qualifications which they had obtained since 1991; those who were absent from the previous sweep were asked about all qualifications which they had obtained since 1974, or school leaving age. Asking for information over such a long spell of time is very likely to introduce inaccuracies due to imperfect recall. However, for NCDS, the 2000 sweep represents the sixth follow-up on the lives of cohort members, which means that there is potential information in earlier sweeps which can be used for checking. In 1978, schools and colleges attended by cohort members were contacted and asked to supply details of the qualifications which NCDS CMs had obtained. This is a valuable resource, since it represents an objective source of information on qualifications obtained. The main drawback is that it only covers the period up to 1978. In what follows it is referred to as the EXAMS file. It contains data on the number of qualifications obtained, and also some data on the subjects of qualifications. NCDS4 (1981) has some details of qualifications obtained since 1974, but NCDS5 (1991) is more informative. In NCDS5 CMs were asked (a) about qualifications ever obtained and (b) about qualifications obtained since 1981. Note that the NCDS questionnaire did ask for GCSEs, as well as O levels and CSEs. The main drawback of NCDS5 is that it does not tell us about the number of qualifications. For example, CMs were asked if they had CSEs at grades 2-5, but were not asked to report how many they had obtained. It also, of course, has problems of recall errors. The EXAMS file, supplemented by NCDS5, then, are the main sources which we can use to check the accuracy of the data in NCDS6 (2000).

A detailed case-by-case listing for each of the 94 NCDS records in question is contained in Appendix 7, and this should be referred to when considering the summary given here.

Fortunately, most of the 94 cases were present in the EXAMS file, and since the majority of erroneous cases were reporting that they got their GCSEs in the 1973 to 1976 period, the data contained in the EXAMS file is relevant. In more detail, all but ten of the 94 cases were present in the EXAMS file. Out of these, five were present in NCDS5. There is no earlier data for the remaining five.

For some of the remaining cases, EXAMS file data was found to be not relevant because they were reporting GCSE qualifications after 1978, but fortunately this applied to 7 cases. Three of these were covered by NCDS5, but the remaining 4 were not.

In sum, then, there were only 9 cases where data could not be obtained from earlier sweeps to check the GCSE errors in NCDS6.

The majority of those for whom data is available would appear to have CSEs rather than the GCSEs reported in NCDS6. However, some have O levels, and some a mixture of O levels and CSEs. Although a good many cases appear to show confusion between GCSEs and either O levels or CSEs, some cases are less simple, showing differences in the number of O levels recorded in the

EXAMS file and the number of GCSEs listed in NCDS6. In other cases, there is little evidence in the EXAMS file that the qualifications reported in NCDS6 were actually obtained. Other cases show duplication, and can be easily resolved out in the absence of EXAM data.

3. GCSE date errors among a sample of BCS70 cases

Because there are such a large number of GCSE errors among the BCS70 cohort, it has only been possible to look at the first 200 of these (rather less than a fifth of the total).

The BCS70 CMs were one of the last cohorts to sit O levels and CSEs as their school-leaving exams before they were replaced by GCSEs. BCS CMs should only have GCSEs if they stayed in education to at least age 18, or subsequently returned to study for a qualification.

The only source of data for checking the responses in BCS 2000 is the earlier BCS postal survey conducted in 1996. Questions about qualifications in the BCS 1996 survey were asked in the order: CSE, O level, GCSE, rather than the order GCSE, O level, CSE in BCS 2000. Thus checking BCS70 2000 against BCS70 1996 should help to root out errors which may have arisen from the order in which the questions were asked in BCS70 2000. However, BCS70 1996 had a relatively small sample size (n = 9,003), so many of those in BCS70 2000 may not be present in BCS 1996.

Of the 200 BCS70 CMs checked, some 134 (67%) were also present in BCS70 1996. Out of these, 2 also reported GCSEs in BCS70 1996 (although not necessarily the same number as in BCS70 2000); 17 provided no information about qualifications in BCS70 96, and 75 did not report any GCSEs in BCS70 96. Detailed information on each case is presented in Appendix ?.

Among the 134 for whom 1996 data are available it is usually possible to make some reasonable guess as to what should be done about the erroneous GCSE data in BCS 2000. For many, it is clear that the GCSE data should be recoded to O levels, while in others the GCSEs should be recoded as CSEs. Other cases are not quite so obvious, and some, where there are large discrepancies between 1996 and 2000, or where errors have been repeated. The real difficulty is, however, that we have a substantial number of cases not present in 1996, so there is no data to check the BCS 2000 results against.

4. O levels and CSEs reported as gained after 1988

Among all the NCDS CMs, some 59 had reported obtaining one or more O levels after 1988 - after they were (presumably) phased out in 1988.

The dates for the reported O levels were all in the 1990s. The source of the error is likely to be either that (a) they have actually got GCSEs rather than O levels or (b) a mistake was made inputting the data, so that the O level was actually obtained in, say, 1974 but this was coded as 1994.

Some 47 of the 59 cases (79%) were also present in NCDS5, which means that they were (or should have been) only asked questions about qualifications obtained since 1991. This would make it unlikely that error (b) occurred. The data do indeed show that most of these 47 were only reporting O levels obtained in the 1990s: 28 reported a single O level obtained in the 1990s and no earlier O levels; a further 13 reported two O levels obtained in the 1990s and no O levels. However, the remaining six of the 47, despite being present in NCDS5 in 1991, also reported O levels obtained in both the 1970s and the 1990s.

For the 12 not present in 1991, it is difficult to discriminate between the (a) and (b) errors listed above.

So it is not possible to resolve all the cases, but for many recoding to GCSE would be a sensible solution.

Only two NCDS CMs reported obtaining CSEs after 1988.

For BCS, there appear to be 133 cases in which CMs report O levels obtained after 1988, and 32 cases report CSEs obtained after 1988, but these have not been studied in any detail.

Queries about these problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Highest qualification

(Based on work undertaken by:

- **Andrew Jenkins, Institute of Education, University of London and Professor Gerald Makepeace, University of Cardiff**
- **Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London)**

(Based on work undertaken by Andrew Jenkins, Institute of Education, University of London and Professor Gerald Makepeace, University of Cardiff

In reviewing the data relating to qualifications, Andrew Jenkins and Gerald makepeace have developed a number of derived variables relating to qualifications. In particular, they provide dereived variables for the highest academic qualification, the highest vocational qualification and the highest NVQ level recorded in the 2000 survey, as shown below.

HIACA00 'Highest academic qualification recorded in 2000 survey'.

- 9 Missing
- 0 'None '
- 1 'Bad O-levels'
- 2 'CSE 2-5, other Scottish school qualification'
- 3 'Good O levels'
- 4 '1 A level or more than 1 AS level at grade A-C'
- 5 '2 or more A-levels'
- 6 'Diploma'
- 7 'Degree, PGCE, other degree level'
- 8 'Higher degree ' .

NVQACA00 'Highest NVQ level from an academic qualification in 2000 survey'.

GHMACA00 'Highest academic qualification in 2000 survey ghm measure see also IFS'.

- 0 'None'
- 1 'Bad O levels, CSE 2-5'
- 2 'Good O-levels, 1 A-level'
- 3 '2 or more A-levels'
- 4 'Sub-degree'
- 5 'Degree' 6 'Higher Degree' .

HIVOC00 'Highest vocational qualification recorded in 2000 survey'.

- 9 Missing
- 0 'None '
- 1 'C&G Opo & Other, RSA 1, HGV, Other tech or bus qualification'
- 2 'C&G Craft-Inter-Ord-Part1-Cant say, RSA2, Insig,JIB NJC etc'
- 3 'C&G Advanced-Final-Part II or III, ONC-OND, RSA 3 '
- 4 'Nursing, BTEC HNC-HND, HNC-HND, C&G Full Tech (FTC)'
- 5 'NVQ awarded at level 5'.

HiNVQ00 'Highest NVQ level (whether academic or vocational)'

It is important to stress that:

- These variables are based only on the data gathered in the most recent survey.
- For the majority of NCDS cohort members, the variables may have only limited value as they only relate to qualifications gained since the age of 33 years. This is because for most NCDS cohort members, information was only gathered on qualifications gained since the last follow-up in 1991. Only if they did not take part in the 1991 survey was information gathered about qualifications gained since 1981 – the date of the previous follow-up. In this case, the variables relate to qualifications gained since the age of 23. Comprehensive information about qualifications gained since the age of 16 can only be obtained by linking the data gathered in the new survey with that gathered in earlier follow-ups.
- For BCS70 cohort members, the variables relate to qualifications gained since the age of 16.
- Previous published work has used slightly different definitions of these variables so some users may need to adapt the code they provide. The table below shows how the educational qualifications obtained were mapped into the national qualifications framework.

Level of Qualification	General (Academic)	Vocationally-related (Applied)	Occupational (Vocational)
5	Higher Degree		NVQ level 5 PGCE
4	Degree HE Diploma	BTEC Higher Certificate/Diploma HNC/HND	NVQ level 4 Professional degree level qualifications Nursing/paramedic Other teacher training qualification City & Guilds Part 4/Career Ext/Full Tech RSA Higher Diploma
3	A level AS levels Scottish Highers Scottish Cert of 6 th Year Studies	Advanced GNVQ BTEC National Diploma ONC/OND	NVQ level 3 City & Guilds Part 3/Final/Advanced Craft RSA Advanced Diploma Pitmans level 3
2	GCSE grade A*-C O levels grade A-C O levels grade D-E CSE grade 1 Scottish standard grades 1-3 Scottish lower or ordinary grades	Intermediate GNVQ BTEC First Certificate BTEC First Diploma	NVQ level 2 Apprenticeships City & Guilds Part 2/Craft/Intermediate City & Guilds Part 1/Other RSA First Diploma Pitmans level 2
1	GCSE grade D-G CSEs grades 2-5 Scottish standard grades 4-5 Other Scottish school qualification	Foundation GNVQ Other GNVQ	NVQ level 1 Other NVQ Units towards NVQ RSA Cert/Other Pitmans level 1 Other vocational qualifications HGV

Jenkins and Makepeace also define a set of binary variables showing whether an individual has a particular qualification. These variables may be used to devise various typologies for training. They are listed below.

A copy of the SPSS code and documentation supplied by Jenkins and Makepeace is included as Appendix 7.

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Qualification variables derived by Jenkins and Makepeace

Variable	Label
GCSEHI00	Has CM obtained one or more GCSEs at grades A to C ?
GCSELO00	Has CM obtained one or more GCSEs at grades D to E ?
OLEVHI00	Has CM obtained one or more O levels at grades A to C?
OLEVLO00	Has CM obtained one or more O levels at grades D to E?
CSEHI00	Has CM obtained one or more CSEs at grade 1?
CSELO00	Has CM obtained one or more CSEs at grades 2 to 5?
ASLHI00	Has CM obtained one or more AS levels at grades A to C?
ASLLO00	Has CM obtained one or more AS levels at grades D to G?
NASLEV00	Number of AS level passes at grade A-C
ALEVHI00	Has CM obtained one or more A levels at grades A to C?
ALEVLO00	Has CM obtained one or more A levels at grades D to E?
ALEV00	Has CM obtained any A level passes at any grade?
NALEV00	Number of A level passes at any grade
SCOTLO00	Has CM obtained Scottish SCE standard grades 4-5 or equivalent?
SCOT200	Has CM obtained Scottish SCE standard grades 1-3 or equivalent?
SCOT300	Has CM obtained Scottish SUPE-SLC lower or ordinary grade?
SCOT400	Has CM obtained Scottish SCE-SUPE-SLC higher grade or equivalent?
SCOT500	Has CM obtained Scottish Certificate of 6th Year Studies?
SCOT600	Has CM obtained other Scottish school qualification?
SCOTLO00	Has CM obtained Scottish SCE standard grades 4-5 or equivalent?
SCOT200	Has CM obtained Scottish SCE standard grades 1-3 or equivalent?
SCOT300	Has CM obtained Scottish SUPE-SLC lower or ordinary grade?
SCOT400	Has CM obtained Scottish SCE-SUPE-SLC higher grade or equivalent?
SCOT500	Has CM obtained Scottish Certificate of 6th Year Studies?
SCOT600	Has CM obtained other Scottish school qualification?
HDEG00	Has CM obtained a higher degree?
DIP00	Diploma-Certificate-Teacher Training Qual (not pgce) in Higher Ed
NURSE00	Has CM obtained any nursing or paramedic qualifications?
PGCE00	Has CM obtained a PGCE or other postgraduate certificate in HE?
ONCOND00	Does CM have ONC or OND?.
HNCHND00	Does CM have HNC or HND?.
BTEC100	Does CM have a BTEC etc First-General Certificate?
BTEC200	Does CM have a BTEC etc First-General Diploma?
BTEC300	Does CM have a BTEC etc National Certificate Diploma?
BTEC400	Does CM have a BTEC etc Higher Certificate Diploma?
BTEC500	Does CM have other BTEC qualification?
CITY100	Does CM have a City and Guilds Part 1
CITY200	Does CM have a City and Guilds Part2 or Craft or Intermediate
CITY300	Does CM have a City and Guilds Part3 or Final or Advanced Craft
CITY400	Does CM have a City and Guilds Part4 or Career Extension or Full tec
CITY500	Does CM have a City and Guilds Other C&G qualification
RSA100	Does CM have RSA certificate?
RSA200	Does CM have RSA First Diploma?
RSA300	Does CM have RSA Advanced Diploma or Certificate?
RSA400	Does CM have RSA Higher Diploma
RSA500	Does CM have other RSA qualification?
PIT100	Does CM have a Pitmans level 1?
PIT200	Does CM have a Pitmans level 2?

Variable	Label
PIT300	Does CM have a Pitmans level 3?
PIT400	Does CM have otherPitmans qualification?
NVQ100	Does CM have an NVQ at Level 1?
NVQ200	Does CM have an NVQ at Level 2?
NVQ300	Does CM have an NVQ at Level 3?
NVQ400	Does CM have an NVQ at Level 4?
NVQ500	Does CM have an NVQ at Level 5?
NVQ600	Does CM have an NVQ at Level 6?
NVQ700	Does CM have Trusts towards NVQ-SVQ?
NVQ800	Does CM have other NVQ?.
GNVQ100	Does CM have GNVQ Foundation?
GNVQ200	Does CM have GNVQ Intermediate?
GNVQ300	Does CM have GNVQ Advanced?
GNVQ400	Does CM have other GNVQ qualification?
APPREN00	Has CM obtained any recognised trade apprentice qualifications?
HGV00	Has CM obtained at least one hgv qualification?
OTHVOC00	Has CM obtained at least one other vocational qualification?
ACAMISS	Dummy for missing value for do you have an educational qualification
VOCMISS	Dummy for missing value for do you have a vocational qualification
ALLMISS	Missing values, Do you have any of these qualifications?
PROXY	Was interview completed by a proxy
YTS00	Has CM done a YTS course?.
GOV00	Has CM done another gov course (including New Deal)?
MODAPP00	Has CM done a modern apprenticeship?
NFAIL00	Number of courses started where CM did not obtain the qualification
ACCESS00	Has CM done an access course?
WORKC00	Number of other work related courses
LEIS00	Number of courses for interest or leisure
READ00	Number of courses to improve reading
WRITE00	Number of courses to improve writing
MATHS00	Number of courses to improve maths
ACA000	1 if CM has no academic qualifications
ACA100	1 if CM has bad O-levels .
ACA200	1 if CM has CSE 2-5, Other Scots
ACA300	1 if CM has good O-levels .
ACA400	1 if CM has 1 A-level or more than 1 AS at grade A-C
ACA500	1 if CM has 2 or more A-levels
ACA600	1 if CM has diploma .
ACA700	1 if CM has degree, PGCE, other degree level
ACA800	1 if CM has higher degree.
VOC000	1 if CM has no vocational qualifications
VOC100	1 if CM has vocational qualification equivalent to NVQ1
VOC200	1 if CM has vocational qualification equivalent to NVQ2
VOC300	1 if CM has vocational qualification equivalent to NVQ3
VOC400	1 if CM has vocational qualification equivalent to NVQ4
VOC500	1 if CM has vocational qualification equivalent to NVQ5
HIACA00	Highest academic qualification recorded in 2000 survey
NVQACA00	Highest NVQ level from an academic qualification in 2000 survey
GHMACA00	Highest academic qualification in 2000 survey ghm measures see also IFS
HIVOC00	Highest vocational qualification recorded in 2000 survey

Variable	Label
HINVQ00	Highest NVQ level (whether academic or vocational)
GROUP	3 categories - BCS70 and NCDS by whether present in 1991 survey

(Based on work undertaken by Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London)

Building on the work of Jenkins and Makepeace, Samantha Parsons, has developed derived variables which match the measures of highest academic and vocational qualifications, but which incorporate information gathered during the NCDS5 follow-up in 1991. This ensures that they provide information based on all qualifications gained since the age of 16. The variables are listed below.

It is important to stress that in these variables:

- Uses the NCDS5 variables: N501441 to N501469
- Degree is now coded at NVQ4 level.
- At NCDS5 it is impossible to tell how many 'A' levels cohort members have. Hence, this code assumes NCDS5 'A' levels are 2 or more. The Jenkins and Makepeace variables differentiate between those with 1 'A' level, and those with 2 or more.
- City & Guilds Full Technological is allocated to NVQ4. The Jenkins and Makepeace variables allocate this qualification to NVQ3.

HQUAI33 "Highest qual gained at age 33 – based on NCDS5 variables".

- 0 'No qualification'
- 1 'CSE 2-5/equiv NVQ1'
- 2 'O Level/equiv NVQ2'
- 3 'A Level/equiv NVQ3'
- 4 'Higher qual NVQ4'
- 5 'Degree/higher NVQ5/6'
- 1 'No information'.

HQ332K 'highest NVQ qualification - academic or vocational - recoding 2000 to match NCDS5'.

HACAD33B 'ncds33: highest academic qualification to match HIACA00'.

- 0'no quals'
- 2'cse2-5'
- 3'good o levels'
- 5'a levels'
- 6'diploma'
- 7'degree PGCE'
- 8'higher degree'.

HACNVQ33 'ncds33: highest NVQ level qual from academic qualification to match NVQACA00'.

HACGHM33 'ncds33: highest academic qualification - ghm measure to match GHMACA00'.
0'no quals'
1'cse2-5'
2'o levels'
3'a levels'
4'sub degree'
5'degree'
6'higher degree'.

HVOC33 'ncds33: highest vocational qualification - to match HIVOC00'.

HQ33A 'ncds33: highest NVQ qualification - academic or vocational - to match HINVQ00'.

HQACA00 'highest academic qualification - combining hacad33b and hiaca00'.

0'no quals'
1'bad o levels'
2'cse2-5'
3'good o levels'
4'as levels or 1 a level'
5'a levels' 6'diploma'
7'degree PGCE'
8'higher degree'.

HQANVQ00 'highest NVQ level qual from academic qualification - combining hacnvq33 and nvqaca00'.

HQAGHM00 'highest academic qualification - ghm measure - combining hacghm33 and ghmaca00'.

0'no quals'
1'cse2-5'
2'o levels'
3'a levels'
4'sub degree'
5'degree'
6'higher degree'.

HQVOC00 'highest vocational qualification - combining hvoc33 and hivoc00'.

HQNVQ00 'highest NVQ qualification - academic or vocational - combining hq33a and hinvq00'.

A copy of the SPSS code supplied by Samantha Parsons is also included as Appendix 7.

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Basic skills

(Based on work undertaken by Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London)

Whilst reviewing the data on basic skills, Samantha Parsons identified a number of problems and produced a number of very valuable revised and derived variables. These relate to the following:

1. Maths difficulties
2. Reading difficulties
3. Writing difficulties
4. Difficulties resulting from having 3R problem
5. Course attendance
6. Computer use at home and work

The problems and revised and derived variables are identified below. Further details of all variables are to be found in the detailed documentation supplied by Samantha Parsons, which is included as Appendix 5.

Details of further problems identified by Samantha in other areas, and her suggested solutions are given in the sections dealing with 'Health' (Diet, food and exercise; Eating problems; Smoking; Drinking; and Accidents), and with the 'Self-completion' (Skills and Illegal drug use).

1. Maths difficulties

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	N	%	n	%	
mathsprb	66	0.3	4	0	21	0.1	22589
<i>549 (2.4%) had difficulty/were not able to work out change from £10. However, only 426 (78% of 549) asked additional 5 questions on type of maths difficulties CM had:</i>							
mprbtype							426
addup							426
subtract							426
multiply					1	0	425
subtract					1	0	425
mathcors	66	0.3	4	0	20	0.1	22590
<i>429 (1.9%) had done a course to improve their maths. 429 asked if doing course now or previously</i>							
mathsnow	22251	98.1					429
mathimp	139	0.6			3	0	22538
<i>6234 (27.7%) ever wanted to improve their maths. 6234 asked multi-coded questions i) why they wanted to improve their maths and ii) where/how they would improve them</i>							
mthlike1 – mthlike5	16446	72.5			2		6232
mthplac1- mthplac9	16446	72.5			^4		6230
datesprb	1377	6.1	1	0	15	0.1	21271
<i>696 (3.3%) had difficulty/were not able to work out dates/use a calendar</i>							
mathconf	11421	50.4	3	0	20	0.1	11236
<i>1536 (13.7%) not confident helping their child(ren) with maths</i>							
mathskid	11421	50.4	3	0	16	0.1	11240
<i>1093 (9.7%) did not help their child(ren) with maths</i>							
partmath	12733	56.1	4	0	22	0.1	9921
<i>1439 (14.5%) of CM partner did not help their child(ren) with maths</i>							

Derived MATHS DIFFICULTIES variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
bsamd	1394	6.1					21284
<p>4 category variable combining information given in mathsprb and datesprb. Only for CMs who answered both questions. CM reporting 'difficulties' or 'no, can't do' treated as one category. Values: 1'no difficulties' 2'maths difficulties' 3'date/calendar difficulties' 4'maths and 3'date/calendar difficulties'.</p>							
bsamp1	22254	98.1					426
<p>6 category variable counting the specific maths problems of CMs who had reported maths difficulties in mathsprb. Included mprbtype (recognising numbers), addup, subtract, multiply, and divide. Values: 0'not these problems' 1'one problem' 2'two problems' 3'three problems' 4'four problems' 5'all five problems'.</p>							
bsamp2	22254	98.1					426
<p>4 category variable counting the specific maths problems of CMs who had reported maths difficulties in mathsprb. Included addup, subtract, multiply, and divide. Values: 0'not these problems' 1'one problem' 2'two problems' 3'three problems' 4'all four problems'.</p>							
<p>Derived variables from multi-coded mthlike1-mthlike5: reasons why CM wants to improve their maths: for their children, to get a job, to get a promotion, get a better job, for self satisfaction.</p>							
impmkid	16448	72.5					6232
impmgjob							
impmprom							
impmbjob							
impmself							
<p>Derived variables from multi-coded mthplac1-mthplac9: means by which CM would improve their maths: on a day college course, on an evening college course, on a weekend college course, on ANY college course, at a community centre, by using local library resources, by using PC packages, by using programmes on TV, by using programmes on the radio, by using books at home.</p>							
domcolld	16450	72.5					6230
domcolle							
domcollw							
domcoll							
domcomc							
domllib							
dompc							
domtv							
domradio							
dombook							

2. Reading difficulties

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
readprb1 <i>679 (3.2%) had difficulty/were not able read/understand magazine/newspaper text. 680 asked additional 2 questions on other reading difficulties (includes CM with value of 9 in readprb1)</i>	1377	6.1			1	0	21302
readprb2	22000	97.0					680
readprb3	22003	97.0			3	0	677
readprb4 <i>Asks if CM feels ability to read paperwork has improved over the last decade. 789 (3.7%) feel it has got worse.</i>	1377	6.1			3	0	21300
Readcors <i>197 (0.9%) had done a course to improve their reading. 197 asked if doing course now or previously</i>	66	0.3	4	0	20	0.1	22590
Readnow	22483	99.1					197
Readimp <i>1856 (8.2%) ever wanted to improve their reading. 1856 asked multi-coded questions i) why they wanted to improve their reading and ii) where/how they would improve them</i>	128	0.6			2	0	22550
redlike1 – redlike5	20824	91.8					1856
redplac1- redplac9 <i>(redplac8 and redplac9 hold no information)</i>	20824	91.8			^2		1854
Kidrdcnf <i>561 (5.0%) not confident helping their child(ren) with reading</i>	11421	50.4	3	0	21	0.1	11235
Readkid <i>822 (7.3%) did not read to their child(ren)</i>	11421	50.4	3	0	15	0.1	11241
Partread <i>1271 (12.8%) of CM partner did not read to their child(ren)</i>	12733	56.1	4	0	21	0.1	9922

Derived READING DIFFICULTIES variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
bsaread1	1378	6.1					21302
<p><i>4 category variable combining information given in readprb1, readprb2 and readprb3. CM reporting 'difficulties' or 'no, can't do' treated as one category. Values: 1'no reading difficulties' 2'1 type of reading difficulty' 3'2 types of reading difficulties' 4'3 types of reading difficulties'.</i></p>							
bsaread2	22003	97.0					677
<p><i>3 category variable combining information given in readprb1, readprb2 and readprb3 for CMs who reported 'difficulties' or 'no, can't do' in readprb1 and answered readprb1, readprb2 and readprb3. CM reporting 'difficulties' or 'no, can't do' treated as one category. Values: 1'no reading difficulties' 2'1 type of reading difficulty' 3'2 types of reading difficulties' 4'3 types of reading difficulties'.</i></p>							
<p>Derived variables from multi-coded redlike1-redlike5: reasons why CM wants to improve their reading: for their children, to get a job, to get a promotion, get a better job, for self satisfaction.</p>							
imprkid	20824	91.8					1856
imprgjob							
imprprom							
imprbjob							
imprself							
<p>Derived variables from multi-coded redplac1-redplac9: means by which CM would improve their reading: on a day college course, on an evening college course, on a weekend college course, on ANY college course, at a community centre, by using local library resources, by using PC packages, by using programmes on TV, by using programmes on the radio, by using books at home.</p>							
dorcolld	20826	91.8					1854
dorcolle							
dorcollw							
dorcoll							
dorcomc dorllib							
dorpc							
dortv							
dorradio							
dorbook							

3. Writing difficulties

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
writeprb <i>1087 (3.2%) had difficulty/were not able to write a thank-you letter to a friend. 1087 asked additional question on spelling difficulties (includes CM with value of 9 in writeprb)</i>	1377	6.1			1	0	21302
wprbtype <i>70 (6.4%) never tried to write a thank-you letter to a friend. 1017 asked additional questions on handwriting legibility and difficulties putting down in words what want to say</i>	21593	95.2			1	0.1	1086
hwritprb	21663	95.5			1	0.1	1016
wordsprb	21663	95.5					1017
writcors <i>315 (1.4%) had done a course to improve their writing. 315 asked if doing course now or previously</i>	66	0.3	4	0	21	0.1	22589
writenow	22365	98.6					315
writimp <i>2636 (11.7%) ever wanted to improve their writing. 2636 asked multi-coded questions i) why they wanted to improve their reading and ii) where/how they would improve them</i>	138	0.6			3	0	22539
wrilike1 – wrilike5	20044	88.4			1	0	2635
wriplac1- wriplac9	20044	88.4			^5		2631
writconf <i>806 (7.2%) not confident helping their child(ren) with writing</i>	11421	50.4	4	0	26	0.1	11229
writekid <i>1440 (12.8%) did not help their child(ren) with writing</i>	11421	50.4	3	0	14	0.1	11242
partwrit <i>1687 (17.0%) of CM partner did help their child(ren) with writing</i>	12733	56.1	4	0	21	0.1	9922

Derived WRITING DIFFICULTIES variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
bsawrit1	1379	6.1					21301
<p>8 category variable combining information given in writeprb, hwritprb, wprbtype and wordsprb. CM reporting 'difficulties' or 'no, can't do' treated as one category. Values: 0'no probs' 1'writing spelling handwriting articulating' 2'writing spelling handwriting' 3'writing spelling articulating' 4'writing handwriting articulating' 5'writing spelling' 6'writing articulating' 7'writing handwriting' 8'writing only'.</p>							
bsawrit2	21594	95.2					1086
<p>4 category variable counting the specific writing problems of CMs who had reported 'difficulties' or 'no, can't do' in writeprb. Included writeprb, wprbtype (spelling), hwritprb (handwriting), and wordsprb (articulating). CM reporting 'difficulties' or 'no, can't do' in writeprb treated as one category. Values: 1'1 difficulty' 2'2 types of writing difficulties' 3'3 types of writing difficulties' 4'4 types of writing difficulties'.</p>							
<p>Derived variables from multi-coded redlike1-redlike5: reasons why CM wants to improve their reading: for their children, to get a job, to get a promotion, get a better job, for self satisfaction. Values: 0'not this reason' 1'yes, for this reason'.</p>							
impwkid	20045	88.4					2635
impwgjob							
impwprom							
impwbjob							
impwself							
<p>Derived variables from multi-coded redplac1-redplac9: means by which CM would improve their reading: on a day college course, on an evening college course, on a weekend college course, on ANY college course, at a community centre, by using local library resources, by using PC packages, by using programmes on TV, by using programmes on the radio, by using books at home. Values: 0'not using this method' 1'this method'.</p>							
dowcolld	20049	88.4					2631
dowcolle							
dowcollw							
dowcoll							
dowcomc							
dowllib							
dowpc							
dowtv							
dowradio							
dowbook							

4. Difficulties resulting from having basic skills problems

Variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
Getjob 620 (38.5%) CMs with 3R problem felt affected them getting new job	21047	92.8			21	0.1	1612
Copejob 391 (24.2%) CMs with 3R problem felt affected them coping in current job	21047	92.8			15	0.1	1618
Gtpromot 570 (35.7%) CMs with 3R problem felt affected them getting a promotion	21047	92.8			38	0.2	1595
Copehmb 253 (15.6%) CMs with 3R problem felt affected them in managing household business	21047	92.8			6	0	1627
Helpkids 419 (25.7%) CMs with 3R problem felt affected them being able to help their children lear. In addition, 409 (25.1%) never helped their children to read or learn	21047	92.8			3	0	1630
Copeleis 341 (20.9%) CMs with 3R problem felt affected them pursuing other interests	21047	92.8			4	0	1629

5. Course attendance

Variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	N	%	n	%	n	%	
Failqual	66	0.3	3		34	0.2	22577
<i>3404 (8.8%) failed/cot completed course. 3404 asked how many courses they had failed/not completed</i>							
Numfqual	19276	84.9			^3		3401
yts (bcs70 only)	11454	50.5	3		12	0.1	11211
<i>3303 (29.5%) CMs had done a yts course. 3303 asked how many they had done, and if they were on a course now</i>							
Numyts	19377	85.4	^1		^1		3301
Ytsnow	19377	85.4	1				3302
Othgov	66	0.3	4		22	0.1	22588
<i>550 (2.4%) CMs had done an 'other' type of Government course. 550 asked how many they had done, and if they were on a course now</i>							
Numgov	22130	97.6			^3	0.5	547
Govnow	22130	97.6					550
Aptrain	66	0.3	4		24	0.1	22586
<i>329 (1.5%) CMs had done a modern apprenticeship course. 329 asked how many they had done, and if they were on a course now</i>							
Numap	22351	98.5					329
Apnow	22351	98.5					329
Actrain	66	0.3	4		23	0.1	22587
<i>352 (1.6%) CMs had done an access course. 3303 asked how many they had done, and if they were on a course now</i>							
Numac	22328	98.4					352
Acnow	22328	98.4					352
Wrktrain	66	0.3	4		24	0.1	22586
<i>7871 (34.7%) CMs had done a work-related training course. 7871 asked how many they had done, and if they were on a course now</i>							
Numwrktr	14809	65.3			^66	0.8	7805
Wrktrnow	14809	65.3			1	0	7870
Leiscors	66	0.3	5		23	0.1	22586
<i>5528 (24.4%) CMs had done a course for interest or leisure. 5528 asked how many they had done, and if they were on a course now</i>							
Numleis	17152	75.6			^4	0.1	5524
Leisnow	17152	75.6					5528

Derived COURSE ATTENDANCE variables

Variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	N	%	n	%	n	%	
yts1 <i>3 category variable combining information from yts and ytsnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	11470	50.6					11210
Gov1 <i>3 category variable combining information from othgov and govnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	94	0.4					22588
apt1 <i>3 category variable combining information from aptrain and apnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	92	0.4					22586
access1 <i>3 category variable combining information from actrain and acnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	95	0.4					22587
wrt1 <i>3 category variable combining information from wrktrain and wrktrnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	93	0.4					22585
Leis1 <i>3 category variable combining information from leiscors and leisnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	94	0.4					22586
read1 <i>3 category variable combining information from readcors and readnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	90	0.4					22590
Writ1 <i>3 category variable combining information from writcors and writenow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	91	0.4					22589
math1 <i>3 category variable combining information from mathcors and mathsnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	90	0.4					22590

6. Computer use at Home and Work

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
pchome 13379 (59.0%) CM had a PC at home. 13379 asked how often they used the	66	0.3	4		20	0.1	22590
hpcuse 11289 (84.4%) CMs used the PC at home. 11289 asked about the ways they use a PC	9301	41.0			4	0	13375
howuse01-howuse33 (howuse18-howuse33 hold no information)	11391	50.2			^8	0.1	11281
pcwork 12026 (64.1%) CM use a PC at work. 12026 asked how often use a PC and the ways they use a PC	3915	17.3	2		9	0	18754
wpcuse	10654	47.0	1				12025
howuse34-howuse46	10654	47.0			^2	0	12024

Derived HOME COMPUTER variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
pchome1 6 category variable combining information from pchome and hpcuse . Values: 0'no' 1'yes, but do not use it' 2'yes, use it less than once a week' 3'yes, use it once a week' 4'yes, use it 2-4 times a week' 5'yes, use it daily'	94	0.4					22586
Derived variables from multi-coded howuse01-howuse33. For 11289 (50.0%) CMs who have and use a home computer. Ways CM uses home computer: for word processing, WWW, e-mail, data analysis, data base work, design packages, playing games, sending/receiving faxes, using CD ROM or Encyclopedia, composing music, listening to music, photography, programming, managing home finances, spreadsheets, webb design, scanning, other things, unspecified things. Values: 0'not use home PC this way' 1'yes, uses home PC this way'.							
hwp	11391	50.2					11289
hwww							
hemail							
hdatan							
hdatan							
hdesign							
hgames							
hfax							
hencyrom							
hmusicc							
hmusicl							
hphoto							
hprog							
hhomefin							
hspread							
hwebdes							
hscan							
hother							
hunspec							

Derived WORK COMPUTER variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	N	%	n	%	
pcwork1	3927	17.3					18753
<i>5 category variable combining information from pcwork and wpcuse. Values 0'no' 1'yes, use it less than once a week' 2'yes, use it once a week' 3'yes, use it 2-4 times a week' 4'yes, use it daily'.</i>							
Derived variables from multi-coded howuse34-howuse46. For 12025 (64.1%) CMs who use a PC at work. Ways CM uses work computer: for word processing, WWW, e-mail, data analysis, data base work, design packages, playing games, sending/receiving faxes, using CD ROM or Encyclopedia, composing music, listening to music, photography, programming, spreadsheets, webb design, scanning, other things, unspecified things. Values: 0'not use work PC this way' 1'yes, uses work PC this way'.							
wwp	10655	47.0					12025
www							
wemail							
wdata							
wdata							
wdesign							
wgames							
wfax							
wencyrom							
wmusic							
wmusic							
wphoto							
wprog							
wother							

Queries about these problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

HEALTH

Health	
General health	Self-assessed health and experience of list of conditions, including age at onset and contact with doctor
Long-term health conditions	Details of longstanding illnesses, etc (including limiting impact and age at onset), impact on employment, registered disabled.
Respiratory problems	Coughing, phlegm and shortness of breath
Mental health	Experience of mental health problems, including age at onset
Seeing and hearing	Problems with sight/eyes and with hearing
Other conditions	Details of other health conditions requiring regular medical supervision
Accidents/injuries	Works backwards. Details of accidents/ injuries/assaults (age, why admitted, out/in-patient, type of injury). Nature of any permanent disability resulting from any accident/etc
Hospital admissions	Works backwards. Age and why admitted
Smoking	Smoking habit of respondent and partner
Drinking	Alcohol consumption in last 7 days, other aspects of drinking behaviour
Diet	Frequency of consumption of types of food, vegetarian or other special diets
Exercise	Exercise at work and in daily life
Height and weight	Self-reported height and weight and assessment of weight

(Based on work undertaken by:

- *Barbara Jefferis, Institute of Child Health*
- *Scott Montgomery, Karolinska Sjukhuset, Stockholm, Sweden*
- *Richard Rowe and Barbara Maughan, Institute of Psychiatry*
- *Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London)*

Whilst reviewing the data on health, those named above identified a number of problems and produced a number of very valuable revised and derived variables. Details are reported below for each area of the areas identified above.

Additional information on health was included in the self-completion and further information is given on these data are given in the relevant section below.

General Health

(Based on work undertaken by:

- *Barbara Jefferis, Institute of Child Health*
- *Richard Rowe and Barbara Maughan, Institute of Psychiatry)*

Health Generally

1. How is your health generally? (HLTHGEN) -Rates similar in BCS70 and NCDS.

For NCDS: Reported as Excellent (31%men, 29%women), good (51%men, 52%women), fair (15%men, 14%women), poor (3%men, 4%women),

2. How is your health over the last 12 months? (HLTHYR) - Rates in poorer categories perhaps a little lower in BCS70 as may be expected from their younger age

For NCDS, as in the NCDS4 (age 23) and NCDS5 (age 33) surveys, more men than women report excellent health - Excellent (29%men, 26%women), good (47%men, 46%women), fairly good (16%men, 16%women), not so good (8%men, 12%women). Overall, 76% of NCDS men and 72% NCDS women report good or excellent health in the last 12 months, more women report poorer health. This follows the same patterns as in the NCDS4 and NCDS5 year surveys.

Specific Health Conditions

(Based on work undertaken by:

- *Barbara Jefferis, Institute of Child Health*
- *Scott Montgomery,??*
- *Richard Rowe and Barbara Maughan, Institute of Psychiatry*
- *Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London)*

Variables with no valid cases

The following variables are redundant, as they have no valid cases:

OTHSKIN SKINCOND SL1AGE, SL112M, SL1DOC SKINCON7 EATING4 HERNIWH4 GYNAEP09 GYNAEP10 KINDDIA3 (CANCTY04 TO CANCTY12) UGICOCR5

Eating problems

1. The "Some other eating problem" category is the most common positive response but it is not clear how this may be unpacked further, as EATOTH does not appear in the data set.

EATOTH is an open question which along with the answers to other open health questions, has recently been coded by CLS. These data will be made available shortly, along with the text of all open questions.

2. EATING4 variable included but all cases are missing (none in NCDS either).
3. Eating problems – some distributions and derived variables.

Variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	N	%	N	%	n	%	
eatprob	66	0.3	6		21	0.1	22587
<i>729 (3.2%) CMs reported an eating disorder. 729 asked name of eating problem(s) they have had in multi-coded variables</i>							
eating1-eating4 (eating4 holds no information)	21951	96.8	1				728
el1age	21951	96.8			^5	0.7	724
el112m	21951	96.8	1	0.1	1	0.1	727
<i>349 (47.9%) CMs had an eating disorder in last 12 months. 349 asked if they had seen a doctor re: eating disorder</i>							
el1doc	22331						349

Source: Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London.

Derived EATING PROBLEMS variables

Variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		Valid n
	N	%	n	%	n	%	
Eatprnow	95	0.4					22585
<i>3 category variable combining information given in eatprob and el112m. Values: 0'no eating problems' 1'previous eating problem' 2'current eating problem'</i>							
bulemia	93	0.4					22587
Anorexia	93						22587
Swallow	93						22587
Otheatpr	93						22587
<i>2 category variables. If CM gave valid answer to eatprob, used information in multi-coded eating1-eating4 variables. Values: 0'not this eating problem' 1'yes, this eating problem'</i>							

Source: Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London.

Further details of these derived variables are to be found in the detailed documentation supplied by Samantha Parsons, which is included as Appendix 5.

Blood Pressure

For NCDS, 10% (n=592) men and 12.5% (n=708) of women report ever having had high blood pressure (perhaps gestational blood pressure may explain the difference?). This compares with 256 men and 356 women reporting ever had high blood pressure at 33 years. Approximately 6% of the cohort have had high blood pressure in the last 12 months.

Cancers

For NCDS, more cancers are reported in women than in men: 3.5 % compared to 2%. At age 41, 61 men and 202 women have had cancers, compared to 33 men and 130 women at age 33. Women report mostly breast, cervix and skin cancers, for men the most common cancers are skin and testes. 11 men and 38 women have had cancer in the past 12 months, all but 3 have seen their doctors about it.

Diabetes

For NCDS, approximately 1.7% of the cohort report ever having had diabetes, 98 men and 99 women, this compares with 39 men and 32 women reporting 'ever had' diabetes at age 33. 1.2% have had it in the past 12 months, most consulting their doctors about it in this time period.

Allergy and headache

For NCDS, 14% men and 15% women have had hay fever in the last 12 months (compared to 16% at age 33), 3% men and 5% women have seen a doctor about it in the last 12 months. Average age of onset for men was 18 years and for women, 19 years.

Bronchitis

For NCDS, 2% men and 3% of women in the sample have had bronchitis in last 12 months, with slightly fewer consulting their doctor about it in the last 12 months. Average age at which men first had bronchitis was 18 for men and 21 for women. Similar incidence rates are reported in the 33 year sample.

Asthma

For NCDS, 5% and 7% women have had asthma in the last 12 months. Average age men first had asthma was 17 years, and 24 years for women. 2% men and 5% women have had allergic rhinitis in last 12 months. Average age at which cohort first had allergic rhinitis was 24 years.

20% (n=1146) men and 25% (n=1444) women report ever having skin problems at age 41, in comparison at 33 15% men (n=820) and 20% women (n=1114) report ever having eczema or skin problems.

Eczema or other skin problems

For NCDS, 932 men and 632 women have had eczema or other skin problems (mostly psoriasis, acne and contact dermatitis, mouth ulcers, fungal infections) in the last 12 months. This is higher than the reported 12 month incidence at 33 years, for men 10% (n=556) and 14% (n=798 for women).

Average age men and women first had eczema was around 20 years. Average age men first had acne was 14 years, for women it was 17 years. For psoriasis, it was 23.5 years for men and 22 years for women. The average age men and women first had cold sores was 21.

7% men and 19% of women have had severe headaches in the last 12 months. 2% men and 8% of women have seen a doctor about these in the last 12 months. Average age that cohort members first had severe headaches was 22.

ME, back pain, eating problems, hernia and fits

Less than 1% of the NCDS cohort report having suffered ME in the last 12 months. For NCDS, 22% of men and women have suffered persistent back pain, lumbago or sciatica. This compares with 51% men and 43% women reporting ever having lower back pain at age 33. 16% have had persistent back pain over the last 12 months, with 9% visiting their doctor about it.

For NCDS, 2% (n=97) men and 4% (n=225) women have ever had eating problems. 7% men and 3% women report having had hernias. 1.6% (88) men and 1% (64) women have had a hernia in the past 12 months.

For NCDS, 2% of the sample have ever had fits or convulsions, less than 1% in the past month. Average age men first had fits was 20, and women, 17 years.

Women's health

For NCDS, 26% women reported problems with their periods- mostly heavy or painful. 15% had problems with their periods over the last 12 months, and 11% had seen their doctor about it. This compares with 18% reporting ever having persistent trouble with periods at age 33 and 14% having persistent troubles over the last 12 months.

At age 41, 22% had other gynaecological problems, commonly ovarian cysts, endometriosis and fibroids. 6% had gynaecological problems over the last 12 months and 5% had seen their doctor about it. Average age women first had gynaecological problems was 31. At age 33, 16% women reported ever having other gynecological problems and 10% in the last 12 months.

9% of women are currently taking the pill and 79% have ever taken the pill.

26% men and 23% women have been sterilised or had vasectomy/ hysterectomy.

Kidney, bladder problems, etc

For NCDS:

- 6% men and 10% women have ever had kidney or bladder problems (similar to age 33, 4% and 9% for men and women). In the 41 year old men it was mostly kidney or bladder stones, in women, mostly kidney or bladder infections.
- 2.2% men and 4.1% women have had kidney or bladder problems in last 12 months, most have seen their doctor about it.
- 5% women and 12% men have had irritable bowel syndrome.

- 5% men and 3% women have had duodenal or peptic ulcers. <1% men or women have had ulcers in the last 12 months.
- 4% men and 8% women have had gall stones in the last 12 months.
- 0.3% of cohort had IBS in last 12 months and 0.2% had ulcerative colitis in last 12 months.

A summary of the more prevalent conditions in the NCDS cohort is given below.

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Most prevalent health conditions in the NCDS cohort

Ever had	Men (%)	Women (%)	In the last 12 months	Men (%)	Women (%)
Accidents	45	25			
Wheezing/ whistling chest	28	28	Back pain/ lumbago/ sciatica	22	22
Problems with periods		26	Persistent back pain	16	
Painful periods		15	Wheezing/ whistling chest	19	
Heavy periods		12	Skin problems	20	
Gynecological problems		22	Migrane / headache	7	18
Ovarian cyst		4			
Fibroids		3	Problems with periods		15
Endometriosis		3	Respiratory problems		
Hay fever	20	21	Winter cough (am)	14	13
Migrane / headache	13	27	Winter cough (am or pm)	13	14
Problems with sight or eyes	20	20	All year cough	10	10
Both eyes	15	16	Winter phlegm (am)	11	7
One eye	5	4	Winter phlegm (am or pm)	8	6
Eczema	10	15	All year phlegm	8	5
High blood pressure	10	12.5			
Asthma	10	12	Gynecological problems		6
Low/depressed/sad	8	15	High blood pressure	6	6
Bronchitis	8	13	Hay fever	6	6
Hearing problems	9	8	Gall stones	4	8
IBS	5	12	Bronchitis	7	3
Kidney / bladder problems	6	10	Asthma	5	5
Duodenal or peptic ulcers	6	3	Eating problems	2	4
Anxious/ jittery	3.5	6.5	Diabetes	1.2	
Psoriasis	4	4	Allergic rhinitis	0.8	2
Phobic	2	4	ME	<1	
Rhinitis	3	6	Duodenal or peptic ulcers	<1	
Cold sores	2	4	IBS	0.3	
Cancers	2	3.5	Ulcerative colitis	0.2	
Acne	3	2			
Fungus infections	3	2			
Contact dermatitis	2	3			
Gallstones	0.8	4			
Diabetes	1.7				
Fits/ convulsions	<2				
Mouth ulcers	0.8	1			

Source: Barbara Jefferis, Institute of Child Health

Long-term health conditions

The following variables are redundant, as they have no valid cases:

LSIIMWK4 LSIIMWK5 LSIIMWK6 LSIIMWK7 LSIIMWK8 LSIIMWK9
LSIAGE4 LSIAGE5 LSIAGE6 LSIAGE7 LSIAGE8 LSIAGE9

Respiratory problems

For NCDS:

- 14% men and 13% women cough first thing in the morning in winter.
- 11% men and 7% women bring up phlegm from chest in the morning in winter.
- 8% men and 6% women usually bring up phlegm from chest during day or night in winter.
- 12.6% men and 14.2% women usually cough day or night in winter.
- 10% men and women cough on most days for at least 3 months of the year.
- 8% men and 5% bring up phlegm from chest most days for at least 3 months of the year.
- These responses to the respiratory health questions are very similar at age 33.
- 2.3% men and 1.9% women cannot walk.
- 6% men and 11% women are ever out of breath walking on level ground (more than at 33 years).
- 28% men and women have ever had whistling or wheezing chest. 18.5% have had this in the last 12 months. Of these, 60% have had it less than 5 times and 40% five times or more.
- 8% men and 11% women in the cohort have used an inhaler or other prescribed asthma medicine in the last 12 months.

Mental health

Whilst reviewing the data on health, Richard Rowe and Barbara Maughan made a number of observations about the quality and utility of the data. These are summarised below.

1. MHPROBS1

- a) MHPROBS1 contain 98 and 99 codes; these need labelling.

As noted above, the following missing values will be found within the data:

Missing values (unless otherwise labelled)

7, 97, 997, 9997, 99997, 999997	=	Refused
8, 98, 998, 9998, 99998, 999998	=	Don't know
9, 99, 999, 9999, 99999, 999999	=	Not answered
. (sysmis)	=	Not applicable

- b) It would be much easier to code MHPROBS1-9 by disorder as the age of onset variables are coded. ie. MHPROBS1 would be a binary variable indicating the presence or absence of depression etc.

As noted above, variable names have been allocated through CAPI (Blaise). MHPROBS is the name of the CAPI variable and where the question has been repeated, Blaise places adds numerical suffix starting with 2, hence the sequence for MHPROBS (and all repeated variables) is MHPROBS, MHPROBS2,...,MHPROBSn.

- c) Prevalence by sex looks plausible (see table below).

2. Age of onset variables

- a) Depression age at onset: MHAGE more consistently called MHAGE1. Without the number, it looks like this may be a summary of the MHSPECS that include a numerical identifier. 99 codes need to be labelled as missing.

See comments above.

- b) All look consistent with mhprobs variables.

3. MHSPEC

- a) Similarly to age of onset, mhspec should be called mhspec1.
- b) MHSPEC labelled 'seen specialist/been to hosp' whereas mhprobs say 'seen specialist'. It would be helpful to make this consistent.
- c) A lot of missing data on these variables. Of those who report depression since refdate. 52% are missing on mhspec. Similar levels of missing regarding other disorders. It is not clear why this a problem here, particularly as it is not a problem for the mhstill variables.

4. MHSTILL

- a) Similarly to age of onset and MHSPEC, MHSTILL should be called MHSTILL1.
- b) Rates look plausible (see table below)

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Prevalence of mental health problems reported in NCDS/BCS70

<u>Disorder</u>		<u>In entire sample</u>		<u>In those having experienced problem since redate (males and females combined)</u>		
		<u>Male</u>	<u>Female</u>	<u>Still has problems most of the time</u>	<u>Still has problems occasionally</u>	<u>No longer has problems</u>
Depression	NCDS	14% (N=811)	29% (N=1655)	14%	41%	45%
	BCS-70	13% (N=702)	28% (N=1631)	10%	38%	52%
Anxiety	NCDS	6% (N=306)	10% (N=548)	18%	48%	33%
	BCS-70	5% (N=283)	8% (N=447)	17%	45%	38%
Phobias	NCDS	3% (N=165)	6% (N=332)	20%	49%	30%
	BCS-70	4% (N=196)	6% (N=317)	18%	43%	39%
Overexcited/over confident	NCDS	0.4% (N=22)	0.4% (N=24)	23%	34%	41%
	BCS-70	0.5% (N=26)	0.4% (N=21)	25%	47%	28%
Compelled to repeat activities	NCDS	1% (N=54)	1% (N=60)	33%	41%	25%
	BCS-70	1% (N=54)	1% (N=62)	33%	40%	28%
Hallucinations	NCDS	1% (N=40)	1% (N=40)	24%	48%	29%

	BCS-70	1% (N=50)	1% (N=49)	24%	34%	41%
Alcohol problems	NCDS	2% (N=93)	1% (N=64)	26%	25%	49%
	BCS-70	2% (N=135)	0.6% (N=37)	21%	17%	62%
Drug problems	NCDS	0.6% (N=33)	0.3% (N=18)	35%	20%	45%
	BCS-70	2% (N=114)	0.7% (N=41)	21%	14%	65%
Bulimia	NCDS	0.04% (N=2)	0.9% (N=51)	-	-	-
	BCS-70	0.2% (N=9)	1.7% (N=98)	-	-	-
Anorexia	NCDS	0.1% (N=4)	1.3% (N=73)	-	-	-
	BCS-70	0.1% (N=5)	2% (N=111)	-	-	-
Swallowing problems	NCDS	0.2% (N=9)	0.2% (N=9)	-	-	-
	BCS-70	0.1% (N=4)	0.1% (N=8)	-	-	-

Seeing and hearing

For NCDS:

- 20% of the cohort report problems with sight or eyes: 15% in both eyes and 5% in one eye.
- 9% have hearing problems: evenly distributed between problems with one ear or with both.
- 8% of the cohort are short sighted, 4% long sighted and 2.5% have other sight or eye conditions.
- 60% men and 40% women do not wear glasses or contact lenses. 35% men wear glasses and 5% wear contact lenses- women wear lenses more often than men do. 5% men and 7% women wear both contact lenses and glasses.
- 65% men and 58% women read books, watch television and see into the distance without problems.
- 90% of those reporting their everyday vision thought it was average or better than average and 7% reported slightly abnormal vision.
- 5% men and 3% women have some hearing problems (tinnitus, repeated ear infections and discharge)
- <0.5% wear hearing aids.

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Other conditions

1. The following variables are redundant, as they have no valid cases.

HOSUP33 HOSUP36 HOSUP38 HOSUP39 HOSUP41 HOSUP42 HOSUP43 HOSUP44
HOSUP45 HOMORE14 HOAGE15 HOMORE15 HOAGE16 HOSUP46 HOSUP47 HOSUP48
HOMORE16 HOAGE17 HOSUP49 HOSUP50 HOSUP51 HOMORE17 HOAGE18 HOSUP52
HOSUP53 HOSUP54 HOMORE18 HOAGE19 HOSUP55 HOSUP56 HOSUP57 HOMORE19
HOAGE20 HOSUP58 HOSUP59 HOSUP60 HOMORE20

2. For NCDS, 5% men and 7% women have another health condition. This is mostly monitored, either by hospital or clinic or by GP. 20% men and 105 women report that their conditions are not monitored. 0.2% men and 0.5% women are monitored for more than one other health condition.

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Accidents/injuries

1. The following variables are redundant, as they have no valid cases:

ACCIDAN8 ACCIDAN9 ACCINJ18ACCINJ24 ACCINJ30 ACCINJ36

2. 45% men have had accidents since “refdate”, this appears to be age 33. Of these, 13% have had accidents at work, 10% a sports accident and 8% a road accident. Of the 25% women cohort members who have had accidents, 8% have had road accidents, 7% have had an accident in the home and 4% have had an accident at work.

3. For NCDS, men and women differ in the number of accidents that they have seen a doctor for. 13% men have had a sports accident and 12% have had an accident at work. 6% have had an accident at home and 5% have been a driver or passenger in a road accident. 7% women have had an accident at home and 6% have been a driver or a passenger in a road accident. 4% have had an accident at work or an ‘other’ type of accident. 3% have had a sports accident. 2% have had a violent assault or mugging compared to 4% men.

4. Further details of aspects of the information about accidents, and a number of derived variables are given below.

variable	system-missing		user-missing: 98 (^ indicates 998)		user-missing: 99 (^ indicates 999)		valid n
	N	%	N	%	N	%	
<i>Multi-coded questions ask if CM had an accident, and if yes, what type of accident.</i>							
Accidan1	66	0.3	8	0	24	0.1	22582
accidan2							
accidan3							
accidan4							
accidan5							
accidan6							
accidan7							
accidan8							
accidan9							
<i>12458 (55%) reported no accidents with 10124 (45%) reporting they had some kind of accident. These CMs asked how many accidents they had had in total, the age and type of their most recent accident</i>							
Accidno	12524	55.2	^10	0.1	^42	0.4	10104
Accage	12576	55.4	^1	0	^18	0.2	10085
Accwhy	12576	55.4	1	0	2	0	10101

Source: Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London.

Derived ACCIDENT variables

variable	system-missing		user-missing: 98 (^ indicates 998)		user-missing: 99 (^ indicates 999)		valid n
	N	%	N	%	N	%	
2-category derived variables from multi-coded accidan1-accidan9. Each variable gives % experiencing type of accident: road accident (pedestrian), road accident (driver/passenger), at work, at home, at school/college, playing sport, other type, violent assault, sexual assault. Values: 0‘not experienced’ 1‘yes, experienced’.							
accroadp	98	0.4					22582
accroadd							
accwork							
accchome							
accschcl							
accsport							
accother							
vassault							
asssex							

Source: Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London.

Further details of these derived variables are to be found in the detailed documentation supplied by Samantha Parsons, which is included as Appendix 5.

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Smoking

1. Missing values:

There are occasional 8s and 9s on SMOKING, NOFCIGS, AGEQUIT, OTHSMOKE and PARTCIGS

As noted above, the following missing values will be found within the data:

Missing values (unless otherwise labelled)

7, 97, 997, 9997, 99997, 999997	=	Refused
8, 98, 998, 9998, 99998, 999998	=	Don't know
9, 99, 999, 9999, 99999, 999999	=	Not answered
. (sysmis)	=	Not applicable

2. Basic distributions:

	Men		Women	
	BCS70 (n=5438) %	NCDS (n=5600) %	BCS70 (n=5767) %	NCDS (n=5775) %
Never smoked cigarettes	41.8	43.7	46.2	45.5
Used to smoke, don't at all now	18.2	26.0*	19.7	24.5*
Smoke occasionally	8.1	4.6	7.3	4.0
Smoke every day	31.9	25.7	26.8**	26.0**

* Mean age of quitting is similar in men and in women, at around 30.5 years.

** Of the current smokers, consumption in men and women peaks at 20/day (31%) with only 5% smoking more than 30 /day. The other 2 peaks are at 10 and 15 per day. Average number of cigarettes smoked per day is 23 for men and 18 for women.

3. Further details of aspects of the information about smoking, and a derived variable giving categories for number of cigarettes smoked are given below.

Variable	system-missing		user-missing: 8 (^ indicates 98; ^^indicates 998)		user-missing: 9 (^ indicates 99; ^^indicates 999)		valid n
	n	%	n	%	N	%	
Smoking <i>6222 (27.4%) CMs smoke every day. 6222 sked how many they smoke a day</i>	66	0.3	8		26	0.1	22580
Nofcigs <i>6347 (28%) CMs ex-smokers or occasional smokers. 6347 asked if ever smoked cigarettes regularly</i>	16458	72.6			^^16	0.3	6206
exsmoker <i>4532 (71.4%) CMs had smoked regularly. 4532 asked age last smoked regularly</i>	16333	72.0					6347
Agequit	18148	80.0			^6	0.1	4526
othsmoke <i>5864 (25.9%) live with someone who smokes in CM home. 5864 asked whether this was spouse/partner or someone else</i>	2382	10.5	8		29	0.1	20261
whosmoke <i>4521 (77.1%) CMs said partner smoked. 4521 asked how many spouse/partner smoked a day</i>	16816	74.1					5864
Partcigs	18159	80.1	1		27	0.1	4493

Source: Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London.

Derived SMOKING variable

variable	system-missing		user-missing: 8 (^ indicates 98; ^^indicates 998)		user-missing: 9 (^ indicates 99; ^^indicates 999)		valid n
	N	%	n	%	n	%	
Smoke <i>6 category variables combining information from smoking and nofcigs. Values: 0'never smoked' 1'ex-smoker' 2'occasional smoker' 3'up to 10 a day' 4'11 to 20 a day' 5'more than 20 a day'.</i>	116	0.5					22564

Source: Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London.

Further details of these derived variables are to be found in the detailed documentation supplied by Samantha Parsons, which is included as Appendix 5.

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Drinking

1. Further details of aspects of the information about alcohol consumption, and a number of derived variables are given below.

variable	system-missing		user-missing: 8 (^ indicates 98; ^^indicates 998)		user-missing: 9 (^ indicates 99; ^^indicates 999)		valid n
	n	%	n	%	n	%	
Drinks	66	0.3	7		27	0.1	22580
<i>18260 (80.7%) CMs drank alcohol regularly (daily, weekly or monthly). 18260 asked how many of each type of alcoholic drink they had in last 7 days prior to interview</i>							
beer	4420	19.5			^^^16	0.1	18244
Spirits	4420	19.5	^^^1		^^^16	0.1	18243
Wine	4420	19.5	^^^1		^^^3		18256
Sherry	4420	19.5	^^^1		^^^1		18258
Pops	4420	19.5	^^^1		^^^1		18258
othdrink	4420	19.5	^^^1				18259
cage1	493	2.2			5		22182
<i>5521 (24.9%) CMs ever felt should cut down on their drinking. 5521 asked if felt this in last year</i>							
cage2	17159	75.7			1		5520
cage3	493	2.2					22187
<i>2193 (9.7%) CMs ever felt criticised about their drinking. 2193 asked if felt this in last year</i>							
cage4	20487	90.3					2193
cage5	493	2.2					22187
<i>2693 (11.9%) CMs ever felt bad or guilty about their drinking. 2693 asked if felt this in last year</i>							
cage6	19987	88.1					2693
cage7	493	2.2	1		1		22185
<i>874 (3.9%) CMs ever had a drink first thing in A.M. to steady hands. 874 asked if they had in last year</i>							
cage8	21806	96.1					874
cage9	4208	18.6			4	0	18468
<i>2539 (11.2%) CMs ever drink during breaks at work. 2539 asked how often they did this</i>							
cage10	20141	88.8			2		2537

Source: Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London.

Derived DRINKING variables

variable	system-missing		user-missing: 8 (^ indicates 98; ^^indicates 998)		user-missing: 9 (^ indicates 99; ^^indicates 999)		valid n
	n	%	n	%	n	%	
drkunits	44250						18230
<i>Continuous variable giving total number of alcohol units CM had in 7 days prior to interview. Combines information given in beer, spirits, wine, sherry and pops. Range from 0-427. Mean: 19.26, Median: 11.00, sd: 25.77</i>							
Drkcut	499	2.2					22181
<i>3 category variable combining cag1 and cage2. Values: 0'no' 1'yes, previously' 2'yes, last 12 months'.</i>							
Drkcrit	493	2.2					22187
<i>3 category variable combining cag3 and cage4. Values: 0'no' 1'yes, previously' 2'yes, last 12 months'.</i>							
Drkbad	493	2.2					22187
<i>3 category variable combining cag5 and cage6. Values: 0'no' 1'yes, previously' 2'yes, last 12 months'.</i>							
drkhand	495	2.2					22185
<i>3 category variable combining cag7 and cage8. Values: 0'no' 1'yes, previously' 2'yes, last 12 months'.</i>							
drkwork	4214	18.6					18466
<i>3 category variable combining cag9 and cage10. Values: 1'yes, special occasions' 2'yes, 2 or 3 times a month' 3'yes, 2 or 3 times a week' 4'yes, most days'.</i>							

Source: Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London.

Further details of these derived variables are to be found in the detailed documentation supplied by Samantha Parsons, which is included as Appendix 5.

2. The information gathered on drinking also includes 10 items from the CAGE scale, see below.

Cage variables

Variable	Label
CAGE1	Ever felt ought to cut down on drinking
CAGE2	Felt ought to cut down on drinking in last year?
CAGE3	Have people annoyed you by criticising your drinking
CAGE4	People annoyed CM re: drinking in last year
CAGE5	Ever felt bad or guilty about your drinking
CAGE6	Felt bad/guilty about drinking in past year?
CAGE7	Ever had a drink first thing in am to steady hands etc
CAGE8	Had drink first thing in am in past year?
CAGE9	Ever have an alcoholic drink during breaks in daily work
CAGE10	Freq of drinking during work breaks in past year

Items 1, 3, 5, 7 may be used to create a drinking problem scale; scores of 2 or more positively endorsed items on this scale have been shown to provide a good indicator of alcohol problems in community samples (King, M. (1986). At risk drinking among general practice attenders. Validation of the CAGE questionnaire. *Psychological Medicine* 16, 213-217; Mayfield, D., McCloed, G. & Hall, P. (1984). The CAGE questionnaire. Validation of a new alcohol screening instrument. *American Journal of Psychiatry* 131, 1121-1123).

SPSS syntax to generate CAGETOT (total score) and CAGED1 (low vs high) is given below..

```
*CAGE total score (count of positively endorsed items) and  
binary alcohol problem indicator (see Mayfield et al, 1984; King, 1986 for details).
```

```
*set no to 0 and yes to 1 and declare missing values.  
recode cage1 cage3 cage5 cage7 (2 = 0).  
mis val cage1 cage3 cage5 cage7 (8, 9).  
val lab cage1 cage3 cage5 cage7 0 'no' 1 'yes'.
```

```
*compute CAGE total score.  
compute cagetot = cage1 + cage3 + cage5 + cage7.  
recode cagetot (sysmis = 9).  
var lab cagetot 'Cage (ever), total score'.  
val lab cagetot 9 'not known'.
```

```
*compute binary indicator of alcohol problems.  
compute cagedi = cagetot.  
recode cagedi (2 through 4 = 1) (0, 1 = 0) (sysmis = 9).  
mis val cagetot cagedi (9).  
var lab cagedi 'Cage (any drinking problem ever), 2+ cutoff cagetot'.  
val lab cagedi 0 'no' 1 'yes' 9 'not known'.
```

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Diet

1. For NCDS:

- 45% men and 60% women eat fresh fruit once / day or more. 20% men and 14% women eat it less than once a week or occasionally.
- 11% men and 25% women eat salads or raw vegetables once a day or more. 25% men and 15% women eat salads or raw vegetables less than once a week.
- 35% men and 43% women eat cooked vegetables at least once per day.
- 66% men and 53% women eat food fried in vegetable oil less than once a week. 20 % men and 10% women eat food fried in vegetable oil 3-6 days a week.
- 8% men and 5% women eat food fried in hard fat 1-2 times a week. 60% men and 73% women never eat food fried in hard fat.
- 53% men and 35% women eat chips more than 1-2 days a week.
- 50% men and 40% women eat eggs less frequently than twice a week.
- 9% men and 21% women eat sweets and chocolates once a day or more. 25% men and women eat sweets and chocolates only occasionally (<1 day /week).
- 23% men and 21% women eat cakes or biscuits at least once per day. 23% men and 26% women eat cakes and biscuits only occasionally.
- 40% men and 47% women eat whole meal bread or rolls 3-6 days per week. 68% men and 60% women eat other bread or rolls 3-6 times a week.
- 32% men and 24% women eat red meat 3-6 days a week.
- 34% men and 40% women eat poultry 3-6 days a week.
- 70% men and women eat fresh fish less than one day a week.
- 76% men and 70% women eat pulses less frequently than one day per week.
- 2% men and 3% women are vegetarians.
- 3% men and 7% women are on a special diet- sometimes prescribed by doctor.

2. Further details of aspects of the information about diet are given below.

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	N	%	N	%	n	%	
Fruit	66	0.3	9		28	0.1	22577
Eggs	66	0.3	10		29	0.1	22575
Salads	66	0.3	10		28	0.1	22576
cookdveg	66	0.3	10		26	0.1	22578
Oilfried	66	0.3	11		37	0.2	22566
Chops	66	0.3	11		45	0.2	22558
Sweets	66	0.3	11		27	0.1	22576
Cakes	66	0.3	11		26	0.1	22577
whlbread	66	0.3	11		27	0.1	22576
Othbread	66	0.3	11		26	0.1	22577
Redmeat	66	0.3	11		26	0.1	22577
Poultry	66	0.3	11		26	0.1	22577
Fish	66	0.3	11		26	0.1	22577
Pulses	66	0.3	11		27	0.1	22576
Veggy	66	0.3	9		26	0.1	22579
<i>808 (3.6%) CMs vegetarian. 808 asked the type of vegetarian they were</i>							
Vegtype							808
Spshdiet	66	0.3	9		26	0.1	22579
<i>1061 (4.7%) CMs followed a special diet. 1061 asked the type of special diet they followed. Only 1009 (95% of 1061) asked if the diet had been recommended by a doctor</i>							
Diettype							1061
Dietdoc							1009
Exercise	66	0.3	10		27	0.1	22577
<i>17263 (76.3%) CMs do regular exercise. 17263 asked how often they take regular exercise and if they become breathless/sweaty during exercis</i>							
Breathle					2	0	17261
Sweat					3	0	17260

Source: Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London.

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

Exercise

For NCDS:

- 75% men and 72% women do regular exercise.
- 15% men and 20 % women exercise every day.
- 61% men and 67% women do regular exercise 2-3 days a week.
- 38% men and 23% women report getting out of breath or sweaty most times that they exercise.

Queries about these problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

CITIZENSHIP AND VALUES

Citizenship and Values	
	Involvement with organisations, voting behaviour and intentions, political alignment, trade union membership, religion, newspaper readership, car ownership, values, political activity

Review of these data by the CLS NCDS/BCS70 Team and by advisors has, to date, revealed no particular problems, nor have any derived variables been developed.

The data will continue to be investigated and, if necessary, detailed updates will be made available, and the dataset will be updated.

Queries about problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

SELF-COMPLETION

SELF-COMPLETION	
Your views	Attitude statements
How you get on with your husband, wife or partner	Includes Locke-Wallace
Some more of your views	Attitude statements
How you feel	Malaise Inventory
Your skills	How good at skill/is skill used at work
More of your views	Attitude statements
How you feel about your life so far	GHQ 12
More of your views	Attitude statements
School exclusion and truancy	Number of temporary/permanent suspensions/exclusions; frequency of truancy
Contact with the police and crime	Number of times moved on, questioned, warned, taken to police station, cautioned, found guilty by a court
Use of illegal drugs	Whether tried number of specific drugs ever/in last 12 months

Your views/Some more of your views/More of your views

(Based on work undertaken by John Preston, Centre for Longitudinal Studies, Institute of Education, University of London)

In reviewing the data for the questions on 'Views' in the self-completion (CASI), John Preston developed the following ten attitude scales:

Scale	Derived variable	Alpha coefficients	
		NCDS	BCS70
1 Anti-racism	TOLERACE	0.8000	0.8222
2 Left-Right	LEFTR	0.7313	0.6792
2 Support for Authority	AUTHORIT	0.6470	0.6176
4 Support for traditional marital values	PROFAM	0.6620	0.6345
5 Effect of children on quality of life	NOKIDS	0.6141	0.6232
6 Support for working Mothers	MUMWORK	0.6880	0.6782
7 Political cynicism	POLCYN	0.6742	0.6526
8 Technophobia	ANTITECH	0.5892	0.5732
9 Environmentalism	ENVIRON	0.4664	0.5070
10 Support for the work ethic	WORKETH	0.5388	0.5469

The scales were constructed using the 50 'Views' items of the self-completion questionnaire (see below), although a small number were not included in the scales, and did not appear to form any clear factors of their own. Data for NCDS and BCS70 was analysed separately, although the factors identified in each were, fortuitously, identical. Principal components analysis was employed using varimax rotation to orthogonal simple structure. The number of factors chosen was based upon those with Eigen values greater than one, followed by the 'scree' test to determine that point at which Eigen values leveled off. Factors were examined for interpretability and to ensure that at least three items had loadings on each of them. Cronbach's alpha was employed as a test of internal consistency of the factors. In all cases, alpha scores were above 0.5 which indicates that it is acceptable to consider the items as forming an identifiable factor.

For each factor, the derived variable identified above is the mean of the scores for each item within the component. Further details of the development of these derived variables and appropriate SPSS

code are to be found in the detailed documentation supplied by John Preston, which is included as Appendix 8.

Questions on 'Views' in the self-completion (CASI)

Variable Label		Variable Label	
LR1	(SC) Big business benefits owners at expense of workers	LR5	(SC) Ordinary people dont get fair share of nations wealth
AR1	(SC) Mixed race marriage is OK	AR4	(SC) Wouldnt mind working with people from other races
E1	(SC)Problems in the environment not that serious	IT2	(SC) Computers enrich the lives of users
A1	(SC) The law should be obeyed even if wrong	MOR4	(SC) Marriage is for life?
C1	(SC) Unless have kids will be lonely in old age	A4	(SC) Give law breakers stiffer sentences
PC1	(SC) No political party would benefit me	E3	(SC) The environment vs economic growth
L1*	(SC) More likely to get better job if do training/educ	WE2	(SC) If I didn t like a job Id pack it in
MOR1	(SC) Divorce is too easy to get these days	LR6	(SC) Government should redistribute income
LR2	(SC) Private schools should be abolished	WM4	(SC)A mother and family happier if she goes out to work
WM1	(SC) Pre-school kids suffer if mum works	L3	(SC) Learning new things boosts confidence
IT1	(SC) Computers at work destroying peoples skills	C3*	(SC) Having children interferes with parents freedom
WE1	(SC) Any job is better than being unemployed	A5	(SC) Young people dont have respect for trad values
MOR2	(SC) Married people happier than unmarried	WM3	(SC) Kids benefit if mum has job outside home?
A2	(SC) Death penalty for some crimes	LR7	(SC) One law for the rich and one for poor?
MOR3	(SC) Couples with kids should not separate	MOR5	(SC) Women should have the right to an abortion?
LR3	(SC) Management get the better of employees	PC3	(SC) Politicians in politics for own benefit?
AR2	(SC) Wouldnt mind if family of diff race moved next door	L4*	(SC) Effort of getting quals more trouble than worth?
C2	(SC) Can have fulfilling life with no kids	AR5	(SC) Not want another race person as my boss?
L2*	(SC) Knowing right people helps more than quals to get job	MOR6	(SC) Alright for unmarried people to have kids?
PC2	(SC) No difference which political party is in power in GB	A6	(SC) Schools teach children to obey authority?
LR4*	(SC) Take out own private health care, stop relying on NHS	IT4	(SC) Every family should have a computer?
E2	(SC) Preserving environment most important	WE3	(SC) Important to hang onto job even if unhappy?
A3	(SC) Censorship is needed to uphold morals	WM5	(SC) Dads job is to earn money;mums to stay home?
WM2	(SC) Family life suffers if mum working ft	C4	(SC) People with no kids are missing out?
AR3	(SC) Would mind kids going to school with diff races	IT5*	(SC) Learning to use a computer more trouble than worth?

* Variable not included in the scales and not appearing to form any clear factors of their own.

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

How you feel (Malaise Inventory)

(Based on work undertaken by Richard Rowe & Barbara Maughan, Institute of Psychiatry)

The Malaise Inventory provides a measure of for assessing psychiatric morbidity, developed by the Rutter and others at the Institute of Psychiatry from the Cornell Medical Index (Rutter M, Tizard J, and Whitmore K (1970) *Education, Health and Behaviour*. London). It is a 24-item self-completion scale and has been included in earlier NCDS and BCS70 follow-ups. The 24 variables included on the dataset are identified below.

Variable	Label
MAL01	(SC) Do you often have backache?
MAL02	(SC) Do you feel tired most of the time?
MAL03	(SC) Do you often feel miserable or depressed?
MAL04	(SC) Do you often have bad headaches?
MAL05	(SC) Do you often get worried about things?
MAL06	(SC) Usually have difficulty falling or staying asleep?
MAL07	(SC) Usually wake unnecessarily early in morning?
MAL08	(SC) Do you wear yourself out worrying about health?
MAL09	(SC) Do you often get into a violent rage?
MAL10	(SC) Do people often annoy and irritate you?
MAL11	(SC) Have you had twitching of face/neck/shoulders?
MAL12	(SC) Often suddenly become scared for no reason?
MAL13	(SC) Often scared to be alone without friends near?
MAL14	(SC) Are you easily upset or irritated?
MAL15	(SC) Are you frightened of going out alone?
MAL16	(SC) Are you constantly keyed up and jittery?
MAL17	(SC) Do you suffer from indigestion?
MAL18	(SC) Do you suffer from an upset stomach?
MAL19	(SC) Is your appetite poor?
MAL20	(SC) Does every little thing get on your nerves?
MAL21	(SC) Does your heart often race like mad?
MAL22	(SC) Often have bad pains in your eyes?
MAL23	(SC) Troubled with rheumatism or fibrositis?
MAL24	(SC) Have you ever had a nervous breakdown?

Typically, the Malaise inventory has been used to produce:

- **Total score:** The sum of the positive responses to the 24 items
- **Binary categorisation:** Scores 0-6 categorised as 'Normal' and 7 and higher as 'Depressed'

Recent analysis has shown evidence that the responses to the items of the Malaise Inventory may represent two separate psychological and somatic sub-scales rather than a single underlying factor:

- **Psychological subscale:** 15 Items - 2, 3, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 19, 20
- **Somatic subscale:** 8 items - 1, 4, 11, 17, 18, 21, 22, 23

SPSS code to generate all four Malaise measures is given below.

```
*Malaise Inventory – Total Score (items 1-24=1)
compute maltot=0.
do repeat x=mal01 to mal24.
if(x eq 1)maltot=maltot+1.
end repeat.
count malmiss=mal01 to mal24(sysmis,8,9).
compute malvalid=24-malmiss.
if(malvalid ge 19)maltot=maltot*24/malvalid.
if(malvalid lt 19)maltot=-1.
missing values maltot(-1).

*Malaise Inventory – Binary categorisation (1-6=1;7-24=2)
compute malaise=maltot.
recode malaise(0 thru 6=1)(6 thru 24=2).

* Malaise Inventory – Psychological subscale (items 2, 3, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 19,
20=1)
count malpsych=mal02 mal03 mal05 mal06 mal07 mal08 mal09 mal10 mal12 mal13 mal14 mal15
mal16 mal19 mal20 (1).

* Malaise Inventory – Somatic subscale (items 1, 4, 11, 17, 18, 21, 22, 23=1)
count malsom=mal01 mal04 mal11 mal17 mal18 mal21 mal22 mal23 (1).
if missing(maltot)malpsych=-1.
if missing(maltot)malsom=-1.
missing values malpsych malsom (-1).
var labels
maltot "Total Malaise score (sum of 1s)"/
malaise "Malaise categories"/
malpsych "Malaise: Psychological subscale"/
malsom "Malaise: Somatic subscale".
value labels malaise 1 'Normal' 2 'Depressed'.
fre maltot malaise malpsych malsom.
```

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

For more information see:

Rodgers B., Pickles, A., Power, C., Collishaw, S. & Maughan, B. (1999). Validity of the Malaise Inventory in general population samples. *Social Psychiatry and Psychiatric Epidemiology* **34**, 333-341)

Available at: <http://link.springer.de/link/service/journals/00127/bibs/9034006/90340333.htm>

Your skills

(Based on work undertaken by Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London)

After reviewing the self-completion data, Samantha Parsons produced the following information about responses to questions included on work-related skills

Further details of all variables are to be found in the detailed documentation supplied by Samantha Parsons, which is included as Appendix 5.

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
SKILL1A <i>22344 CM reported 'good' 'okay' or 'poor' communicating skills. 21544 (96%) reported if they used this skill at work.</i>	283	1.2	3		5		
SKILL1B	1136	5.0			1		21543
SKILL2A <i>22268 CM reported 'good' 'okay' or 'poor' number/calculation skills. 21477 (96%) reported if they used this skill at work.</i>	283	1.2	3		4		
SKILL2B	2103	5.3			1		21476
SKILL3A <i>19726 CM reported 'good' 'okay' or 'poor' computer/IT skills. 19173 (86%) reported if they used this skill at work.</i>	283	1.2	3		7		
SKILL3B	3507	15.5			1		19172
SKILL4A <i>22219 CM reported 'good' 'okay' or 'poor' team working skills. 21469 (96%) reported if they used this skill at work.</i>	283	1.2	3		7		
SKILL4B	1211	5.3			3		21466
SKILL5A <i>22297 CM reported if they were 'good' 'okay' or 'poor' at learning new skills. 21517 (96%) reported if they used this skill at work.</i>	283	1.2	3		6		
SKILL5B	1163	5.1			1		21516
SKILL6A <i>22308 CM reported 'good' 'okay' or 'poor' problem solving skills. 21515 (96%) reported if they used this skill at work.</i>	283	1.2	3		7		
SKILL6B	1165	5.1			1		21514
SKILL7A <i>21909 CM reported 'good' 'okay' or 'poor' tool use skills. 21149 (94%) reported if they used this skill at work.</i>	283	1.2	3		6		
SKILL7B	1531	6.8			1		21148
SKILL8A <i>20975 CM reported 'good' 'okay' or 'poor' caring skills. 20217 (90%) reported if they used this skill at work.</i>	283	1.2	3		19	0.1%	
SKILL8B	2463	10.9			1		21216
SKILL9A <i>21228 CM reported 'good' 'okay' or 'poor' finance/accounting skills. 20523 (92%) reported if they used this skill at work.</i>	283	1.2	3		5		
SKILL9B	2157	9.5			1		20522

Source: Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London.

Queries about problems with these variables should be directed to the User Support Group
(cohort@cls.ioe.ac.uk)

How you feel about your life so far (General Health Questionnaire)

The General Health Questionnaire (GHQ) is a self-administered screening test, designed to identify short-term changes in mental health (depression, anxiety, social dysfunction and somatic symptoms). It is a pure state measure, responding to how much a subject feels that their *present* state "over the past few weeks" is unlike their *usual* state. It does not make clinical diagnoses and should not be used to measure long-standing attributes. There are four different versions:

- GHQ12 - A quick screener for survey use.
- GHQ28 - Used to examine a profile of scores.
- GHQ30 - A screener with 'physical' element items removed.
- GHQ60 - Used to identify cases for more intensive examination.

The GHQ12 was included in the self-completion (CASI) as it is very quick to administer. Nonetheless, it is just as reliable, valid and sensitive as the other longer versions. The 12 variables included in the dataset are identified below.

Variable	Label
GHQ1	(SC)...can concentrate on what you are doing?
GHQ2	(SC)...lost much sleep over worry?
GHQ3	(SC)...felt you were playing a useful part in things?
GHQ4	(SC)...felt capable of making decisions?
GHQ5	(SC)...felt constantly under strain?
GHQ6	(SC)...felt could not overcome difficulties?
GHQ7	(SC)...been able to enjoy normal activities?
GHQ8	(SC)...been able to face up to your problems?
GHQ9	(SC)...been feeling unhappy and depressed?
GHQ10	(SC)...been losing confidence in yourself?
GHQ11	(SC)...been thinking yourself as worthless?
GHQ12	(SC)...been feeling reasonable happy?

The GHQ12 is also quick to score. There are two approaches:

- **Binary scoring** -This the usual way of scoring the GHQ when it is to be used for case identification. The advantage of using this method of scoring is that by weighting the answer options 0, 0, 1, 1, it avoids problems due to "middle users". To score you simply assign 0 for the first two answer options, 1 for the second two answer options and total the scores. (If used as a screener, rather than for survey work, this score can be compared to a threshold).
- **Likert scoring** - The authors suggest that if subscale scores are required for GHQ-28, there are marginal advantages in using the Likert method of scoring which assigns weights of 0, 1, 2 and 3 to each answer option. To score you simply assign 0 for the first answer option, 1 for the second answer option, 2 for the third answer option and 3 for the fourth answer option and total the scores. (If used as a screener, rather than for survey work, this score can also be compared to a threshold).

SPSS code to generate both Binary and Likert GHQ12 scores is given below.

```
* GHQ12 Binary Score (weighting the answer options 0, 0, 1, 1)
fre ghq1 ghq2 ghq3 ghq4 ghq5 ghq6 ghq7 ghq8 ghq9 ghq10 ghq11 ghq12.
missing values ghq1 ghq2 ghq3 ghq4 ghq5 ghq6 ghq7 ghq8 ghq9 ghq10 ghq11 ghq12 (8,9).
recode
  ghq1 ghq2 ghq3 ghq4 ghq5 ghq6 ghq7 ghq8 ghq9 ghq10 ghq11 ghq12
  (1 thru 2=0) (3 thru 4=1) (else=copy) into ghq1bi ghq2bi ghq3bi
  ghq4bi ghq5bi ghq6bi ghq7bi ghq8bi ghq9bi ghq10bi ghq11bi ghq12bi.
compute ghq12bi = ghq1bi+ghq2bi+ghq3bi+ghq4bi+ghq5bi+ghq6bi+ghq7bi+ghq8bi
+ghq9bi+ghq10bi+ghq11bi+ghq12bi.
execute.

* GHQ12 Likert Score (weighting the answer options 0, 1, 2 and 3)
recode
  ghq1 ghq2 ghq3 ghq4 ghq5 ghq6 ghq7 ghq8 ghq9 ghq10 ghq11 ghq12
  (1=0) (2=1) (3=2) (4=3) (else=copy) into ghq1lik ghq2lik ghq3lik ghq4lik
  ghq5lik ghq6lik ghq7lik ghq8lik ghq9lik ghq10lik ghq11lik ghq12lik.
compute ghq12lik = ghq1lik+ghq2lik+ghq3lik+ghq4lik+ghq5lik+ghq6lik+ghq7lik
+ghq8lik+ghq9lik+ghq10lik+ghq11lik+ghq12lik.
var labels
  ghq12bi "GHQ12 score"/
  ghq12lik "GHQ12 Likert score".
fre ghq12bi ghq12lik.
```

Queries about these variables should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

A useful website about the General Health Questionnaire including the scoring methods is:

www.nfer-nelson.co.uk/html/health/products/ghq.htm

Contact with the police and crime

(Based on work undertaken by Richard Rowe & Barbara Maughan, Institute of Psychiatry)

Whilst reviewing the data from the self-completion, Richard Rowe & Barbara Maughan made a number of observations about the quality and utility of the data. These are summarised below.

COURT, POLICE1 TO POLICE5:

- As expected more men, and more respondents in BCS70, report police contacts or being found guilty by a court (e.g. 23.5% of men in BCS70 have been found guilty by a court).
- 64.8% of men in BCS70 report being stopped and questioned. Are such high rates plausible?
- NB: 170 people who have been found guilty in a court have not reported any police contacts.

Queries about problems with these variables should be directed to the User Support Group
(cohort@cls.ioe.ac.uk)

Use of illegal drugs

(Based on work undertaken by:

- **Richard Rowe & Barbara Maughan, Institute of Psychiatry**
- **Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London)**

1. Frequencies for reported use of illegal drugs look plausible (see table below). As expected, men and BCS70 cohort members reported higher levels of drug use.

Type of drug	Taken drug, not in last year	Taken drug in last year
Cannabis	28.6	11.7
Ecstasy	6.7	2.4
Amphetamines	12.9	2.5
LSD	8.8	0.5
Popper	11.5	1.1
Magic mushrooms	9.7	0.5
Cocaine	5.7	3.4
Temazepan	3.5	0.9
Ketamine	1.1	0.2
Crack	0.9	0.3
Heroin	1.2	0.3
Methadone	0.7	0.2
"Semeron"	0.2	<0.1
Other drugs		2.7

Source: Richard Rowe & Barbara Maughan, Institute of Psychiatry

2. "Semeron" is a fictitious drug which was included among the illegal drugs listed in the self-completion. This same "drug" has been included in many similar instruments in a variety of surveys. Although relatively few cohort members report using semeron, it is clear that some caution should be used when analyzing the data for those who do so. If semeron-users are excluded from the above table, the prevalence of the use of other genuine illegal drugs will be marginally reduced.

3. Additional information about the data relating to illegal drug use is given below. A number of derived variables are also identified. The latter exclude data for cohort members who reported using semeron. Further details of these variables are to be found in the detailed documentation supplied by Samantha Parsons, which is included as Appendix 5.

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
CANNABIS	283	1.2	13	0.1	6	0.0	22378
ECSTASY	283	1.2	12	0.1	7	0.0	22378
AMPHET	283	1.2	11	0	6	0.0	22380
LSD	283	1.2	11	0	6	0.0	22380
POPPER	283	1.2	11	0	8	0.0	22378
MAGMUSH	283	1.2	11	0	7	0.0	22379
COCAINE	283	1.2	11	0	7	0.0	22379
TEMAZ	283	1.2	11	0	9	0.0	22377
SEMERON	283	1.2	11	0	10	0.0	22376
KETAMINE	283	1.2	11	0	8	0.0	22378
CRACK	283	1.2	11	0	6	0.0	22380
HEROIN	283	1.2	11	0	6	0.0	22380
METHAD	283	1.2	11	0	6	0.0	22380
OTHDRUG	283	1.2	11	0	7	0.0	22379

DRUG-DRUG6: multi-coded string variables for the 614 CMs who answered 'yes' to **othdrug**. 22378 are 'numeric missing' leaving just 302 individual answers in drug, 37 in drug2, 13 in drug3, 8 in drug4, 4 in drug5, 2 in drug6. Each answer needs to be converted from string to numeric.

Source: Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London.

Derived ILLEGAL DRUG variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	N	%	n	%	
<i>If CMs said they had taken SEMERON they were omitted from derived variables on illegal drug use. 43 (0.2%) said they had taken the drug SEMERON.</i>							
anydrug1	347	1.5					22333
<i>Takes information from the 12 individual illegal drug questions (information in othdrug not coded yet). Counts the number of illegal drugs CM has taken in the last 12 months. 13.5% of CMs have taken at least 1 illegal drug in last 12 months.</i>							
drugs1	347	1.5					22333
<i>Collapses anydrug1 variable into 0'not taken any drug' 1'yes, taken at least 1 illegal drug in last 12 months'.</i>							
anydrug2	347	1.5					22333
<i>Takes information from the 12 individual illegal drug questions (information in othdrug not coded yet). Counts the number of illegal drugs CM has ever taken. 43.5% of CMs have taken at least 1 illegal drug in their lifetime.</i>							
drugs2	347	1.5					22333
<i>Collapses anydrug2 variable into 0'never taken an illegal drug' 1'yes, taken at least 1 illegal drug in lifetime'.</i>							

Source: Samantha Parsons, Centre for Longitudinal Studies, Institute of Education, University of London.

4. The variables DRUG to DRUG6 hold the 'Other drugs' keyed into the laptops by cohort members as they answered the self-completion. Some entries may appear nonsensical, and many are misspelled. Other entries could, in principle be assigned to one of the above categories, (eg: variants on spelling of cannabis etc). A number of other drugs not listed in the self-completion are also identified (eg: opium and solvents). Other entries may well represent street-names of particular drugs. It is hoped that it may be possible to update the variables relating to illegal drugs in order to incorporate this information in the future.

Queries about these problems should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

APPENDICES

		<i>Page</i>
Appendix 1	Pregnancy History Variables	A2
Appendix 2	Proxy Interview Variables	A5
Appendix 3	NCDS/BCS70 Advisors contributing to quality assessment of NCDS/ BCS70 data	A10
Appendix 4	Earnings and related variables	A11
Appendix 5	Basic Skills and other variables	A57
Appendix 6	Problems with reporting of GCSEs	A96
Appendix 7	Highest Qualification	A138
Appendix 8	Attitude Scales	A163

APPENDIX 1: Pregnancy History Variables

Pregnancy	1					2					
	Baby	1	2	3	4	5	1	2	3	4	5
Has CM ever been or got anyone else pregnant	EVERPREG										
(Number of babies carried)	PREGNUM						PREGNUM2				
How long had CM & partner not been using birth control	PREGJ						PREGJ2				
Months not been using birth control	PREGK						PREGK2				
Weeks not been using birth control	PREGKW						PREGKW2				
Years not been using birth control	PREGL						PREGL2				
MC:Whether smoked during pregnancy	CGPREGA1						CGPREGA4				
MC:Whether smoked during pregnancy	CGPREGA2						CGPREGA5				
MC:Whether smoked during pregnancy	CGPREGA3						CGPREGA6				
Whether smoked more, less or same as before	CPREGB						CPREGB2				
Outcome of pregnancy for baby(n)	PREGA	PREGA2	PREGA3	PREGA4	PREGA5	PREGA6	PREGA7	PREGA8	PREGA9	PREGA10	
Name of baby(n)	PREGB	PREGB2	PREGB3	PREGB4	PREGB5	PREGB6	PREGB7	PREGB8	PREGB9	PREGB10	
Sex of baby(n)	PREGC	PREGC2	PREGC3	PREGC4	PREGC5	PREGC6	PREGC7	PREGC8	PREGC9	PREGC10	
Birth weight of baby(n) – units	PREGD	PREGD2	PREGD3	PREGD4	PREGD5	PREGD6	PREGD7	PREGD8	PREGD9	PREGD10	
Birth weight of baby(n) pounds	POUND	POUND2	POUND3	POUND4	POUND5	POUND6	POUND7	POUND8	POUND9	POUND10	
Birth weight of baby(n) ounces	OUNCE	OUNCE2	OUNCE3	OUNCE4	OUNCE5	OUNCE6	OUNCE7	OUNCE8	OUNCE9	OUNCE10	
Birth weight of baby(n) kilos	KILO	KILO2	KILO3	KILO4	KILO5	KILO6	KILO7	KILO8	KILO9	KILO10	
Birth weight of baby(n) grammes	GRAMM	GRAMM2	GRAMM3	GRAMM4	GRAMM5	GRAMM6	GRAMM7	GRAMM8	GRAMM9	GRAMM10	
baby(n) day of live or still birth	PREGED	PREGED2	PREGED3	PREGED4	PREGED5	PREGED6	PREGED7	PREGED8	PREGED9	PREGED10	
baby(n) mnth of live/still birth, miscarriage/termination	PREGEM	PREGEM2	PREGEM3	PREGEM4	PREGEM5	PREGEM6	PREGEM7	PREGEM8	PREGEM9	PREGEM10	
baby(n) year of live/still birth, miscarriage/termination	PREGY	PREGY2	PREGY3	PREGY4	PREGY5	PREGY6	PREGY7	PREGY8	PREGY9	PREGY10	
baby(n) date of birth	PREGF	PREGF2	PREGF3	PREGF4	PREGF5	PREGF6	PREGF7	PREGF8	PREGF9	PREGF10	
Was baby(n) early, late or on time	PREGF	PREGF2	PREGF3	PREGF4	PREGF5	PREGF6	PREGF7	PREGF8	PREGF9	PREGF10	
How many weeks early or late was baby(n)	PREGG	PREGG2	PREGG3	PREGG4	PREGG5	PREGG6	PREGG7	PREGG8	PREGG9	PREGG10	
Was anything wrong with baby(n) at birth	PREGH	PREGH2	PREGH3	PREGH4	PREGH5	PREGH6	PREGH7	PREGH8	PREGH9	PREGH10	
What was the problem with baby(n)	PREGI	PREGI2	PREGI3	PREGI4	PREGI5	PREGI6	PREGI7	PREGI8	PREGI9	PREGI10	
Is CMs current partner baby(n)s other parent	WHOPARA	WHOPARA2	WHOPARA3	WHOPARA4	WHOPARA5	WHOPARA6	WHOPARA7	WHOPARA8	WHOPARA9	WHOPARA10	
Who is baby(n)s other parent	WHOPARB	WHOPARB2	WHOPARB3	WHOPARB4	WHOPARB5	WHOPARB6	WHOPARB7	WHOPARB8	WHOPARB9	WHOPARB10	
Does baby(n) live with CM	WHERKID	WHERKID2	WHERKID3	WHERKID4	WHERKID5	WHERKID6	WHERKID7	WHERKID8	WHERKID9	WHERKID10	
What is baby(n) currently doing	WHATKID	WHATKID2	WHATKID3	WHATKID4	WHATKID5	WHATKID6	WHATKID7	WHATKID8	WHATKID9	WHATKID10	
Does baby(n) ever see other parent	ABSPARA	ABSPARA2	ABSPARA3	ABSPARA4	ABSPARA5	ABSPARA6	ABSPARA7	ABSPARA8	ABSPARA9	ABSPARA10	
How often does baby(n) see other parent	ABSPARB	ABSPARB2	ABSPARB3	ABSPARB4	ABSPARB5	ABSPARB6	ABSPARB7	ABSPARB8	ABSPARB9	ABSPARB10	
Does the other parent of baby(n) pay maintenance	ABSPARC	ABSPARC2	ABSPARC3	ABSPARC4	ABSPARC5	ABSPARC6	ABSPARC7	ABSPARC8	ABSPARC9	ABSPARC10	
Where is baby(n) living now	ABSKIDA	ABSKIDA2	ABSKIDA3	ABSKIDA4	ABSKIDA5	ABSKIDA6	ABSKIDA7	ABSKIDA8	ABSKIDA9	ABSKIDA10	
Has baby(n) ever lived with CM	ABSKIDB	ABSKIDB2	ABSKIDB3	ABSKIDB4	ABSKIDB5	ABSKIDB6	ABSKIDB7	ABSKIDB8	ABSKIDB9	ABSKIDB10	
Year baby(n) last lived with CM	ABSYR	ABSYR2	ABSYR3	ABSYR4	ABSYR5	ABSYR6	ABSYR7	ABSYR8	ABSYR9	ABSYR10	
Month baby(n) last lived with CM	ABSMON	ABSMON2	ABSMON3	ABSMON4	ABSMON5	ABSMON6	ABSMON7	ABSMON8	ABSMON9	ABSMON10	
Does CM ever see baby(n) now	ABSKIDC	ABSKIDC2	ABSKIDC3	ABSKIDC4	ABSKIDC5	ABSKIDC6	ABSKIDC7	ABSKIDC8	ABSKIDC9	ABSKIDC10	
How often does CM see baby(n)	ABSKIDD	ABSKIDD2	ABSKIDD3	ABSKIDD4	ABSKIDD5	ABSKIDD6	ABSKIDD7	ABSKIDD8	ABSKIDD9	ABSKIDD10	
Does CM pay maintenance for baby(n)	ABSKIDE	ABSKIDE2	ABSKIDE3	ABSKIDE4	ABSKIDE5	ABSKIDE6	ABSKIDE7	ABSKIDE8	ABSKIDE9	ABSKIDE10	
Another pregnancy before this one?						MOREPREG				MOREPRE2	

Pregnancy History Variables (continued)

Pregnancy Baby	3					4				
	1	2	3	4	5	1	2	3	4	5
Has CM ever been or got anyone else pregnant										
(Number of babies carried)	PREGNUM3					PREGNUM4				
How long had CM & partner not been using birth control	PREGJ3					PREGJ4				
Months not been using birth control	PREGK3					PREGK4				
Weeks not been using birth control	PREGKW3					PREGKW4				
Years not been using birth control	PREGL3					PREGL4				
MC:Whether smoked during pregnancy	CGPREGA7					CGPREG10				
MC:Whether smoked during pregnancy	CGPREGA8					CGPREG11				
MC:Whether smoked during pregnancy	CGPREGA9					CGPREG12				
Whether smoked more, less or same as before	CPREGB3					CPREGB4				
Outcome of pregnancy for baby(n)	PREGA11	PREGA12	PREGA13	PREGA14	PREGA15	PREGA16	PREGA17	PREGA18	PREGA19	PREGA20
Name of baby(n)	PREGB11	PREGB12	PREGB13	PREGB14	PREGB15	PREGB16	PREGB17	PREGB18	PREGB19	PREGB20
Sex of baby(n)	PREGC11	PREGC12	PREGC13	PREGC14	PREGC15	PREGC16	PREGC17	PREGC18	PREGC19	PREGC20
Birth weight of baby(n) – units	PREGD11	PREGD12	PREGD13	PREGD14	PREGD15	PREGD16	PREGD17	PREGD18	PREGD19	PREGD20
Birth weight of baby(n) pounds	POUND11	POUND12	POUND13	POUND14	POUND15	POUND16	POUND17	POUND18	POUND19	POUND20
Birth weight of baby(n) ounces	OUNCE11	OUNCE12	OUNCE13	OUNCE14	OUNCE15	OUNCE16	OUNCE17	OUNCE18	OUNCE19	OUNCE20
Birth weight of baby(n) kilos	KILO11	KILO12	KILO13	KILO14	KILO15	KILO16	KILO17	KILO18	KILO19	KILO20
Birth weight of baby(n) grammes	GRAMM11	GRAMM12	GRAMM13	GRAMM14	GRAMM15	GRAMM16	GRAMM17	GRAMM18	GRAMM19	GRAMM20
baby(n) day of live or still birth	PREGED11	PREGED12	PREGED13	PREGED14	PREGED15	PREGED16	PREGED17	PREGED18	PREGED19	PREGED20
baby(n) mnth of live/still birth, miscarriage/termination	PREGEM11	PREGEM12	PREGEM13	PREGEM14	PREGEM15	PREGEM16	PREGEM17	PREGEM18	PREGEM19	PREGEM20
baby(n) year of live/still birth, miscarriage/termination	PREGEY11	PREGEY12	PREGEY13	PREGEY14	PREGEY15	PREGEY16	PREGEY17	PREGEY18	PREGEY19	PREGEY20
baby(n) date of birth	PREG11	PREG12	PREG13	PREG14	PREG15	PREG16	PREG17	PREG18	PREG19	PREG20
Was baby(n) early, late or on time	PREGF11	PREGF12	PREGF13	PREGF14	PREGF15	PREGF16	PREGF17	PREGF18	PREGF19	PREGF20
How many weeks early or late was baby(n)	PREGG11	PREGG12	PREGG13	PREGG14	PREGG15	PREGG16	PREGG17	PREGG18	PREGG19	PREGG20
Was anything wrong with baby(n) at birth	PREGH11	PREGH12	PREGH13	PREGH14	PREGH15	PREGH16	PREGH17	PREGH18	PREGH19	PREGH20
What was the problem with baby(n)	PREGI11	PREGI12	PREGI13	PREGI14	PREGI15	PREGI16	PREGI17	PREGI18	PREGI19	PREGI20
Is CMs current partner baby(n)s other parent	WHOPAR12	WHOPAR14	WHOPAR16	WHOPAR18	WHOPAR20	WHOPAR22	WHOPAR24	WHOPAR26	WHOPAR28	WHOPAR30
Who is baby(n)s other parent	WHOPAR13	WHOPAR15	WHOPAR17	WHOPAR19	WHOPAR21	WHOPAR23	WHOPAR25	WHOPAR27	WHOPAR29	WHOPAR31
Does baby(n) live with CM	WHERKI11	WHERKI12	WHERKI13	WHERKI14	WHERKI15	WHERKI16	WHERKI17	WHERKI18	WHERKI19	WHERKI20
What is baby(n) currently doing	WHATKI11	WHATKI12	WHATKI13	WHATKI14	WHATKI15	WHATKI16	WHATKI17	WHATKI18	WHATKI19	WHATKI20
Does baby(n) ever see other parent	ABSPAR13	ABSPAR16	ABSPAR19	ABSPAR22	ABSPAR25	ABSPAR28	ABSPAR31	ABSPAR34	ABSPAR37	ABSPAR40
How often does baby(n) see other parent	ABSPAR14	ABSPAR17	ABSPAR20	ABSPAR23	ABSPAR26	ABSPAR29	ABSPAR32	ABSPAR35	ABSPAR38	ABSPAR41
Does the other parent of baby(n) pay maintenance	ABSPAR15	ABSPAR18	ABSPAR21	ABSPAR24	ABSPAR27	ABSPAR30	ABSPAR33	ABSPAR36	ABSPAR39	ABSPAR42
Where is baby(n) living now	ABSKID15	ABSKID20	ABSKID25	ABSKID30	ABSKID35	ABSKID40	ABSKID45	ABSKID50	ABSKID55	ABSKID60
Has baby(n) ever lived with CM	ABSKID16	ABSKID21	ABSKID26	ABSKID31	ABSKID36	ABSKID41	ABSKID46	ABSKID51	ABSKID56	ABSKID61
Year baby(n) last lived with CM	ABSYR11	ABSYR12	ABSYR13	ABSYR14	ABSYR15	ABSYR16	ABSYR17	ABSYR18	ABSYR19	ABSYR20
Month baby(n) last lived with CM	ABSMON11	ABSMON12	ABSMON13	ABSMON14	ABSMON15	ABSMON16	ABSMON17	ABSMON18	ABSMON19	ABSMON20
Does CM ever see baby(n) now	ABSKID17	ABSKID22	ABSKID27	ABSKID32	ABSKID37	ABSKID42	ABSKID47	ABSKID52	ABSKID57	ABSKID62
How often does CM see baby(n)	ABSKID18	ABSKID23	ABSKID28	ABSKID33	ABSKID38	ABSKID43	ABSKID48	ABSKID53	ABSKID58	ABSKID63
Does CM pay maintenance for baby(n)	ABSKID19	ABSKID24	ABSKID29	ABSKID34	ABSKID39	ABSKID44	ABSKID49	ABSKID54	ABSKID59	ABSKID64
Another pregnancy before this one?					MOREPRE3					MOREPRE4

Pregnancy History Variables (continued)

Pregnancy Baby	5					6				
	1	2	3	4	5	1	2	3	4	5
Has CM ever been or got anyone else pregnant										
(Number of babies carried)	PREGNUM5					PREGNUM6				
How long had CM & partner not been using birth control	PREGJ5					PREGJ6				
Months not been using birth control	PREGK5					PREGK6				
Weeks not been using birth control	PREGKW5					PREGKW6				
Years not been using birth control	PREGL5					PREGL6				
MC:Whether smoked during pregnancy	CGPREG13					CGPREG16				
MC:Whether smoked during pregnancy	CGPREG14					CGPREG17				
MC:Whether smoked during pregnancy	CGPREG15					CGPREG18				
Whether smoked more, less or same as before	CPREGB5					CPREGB6				
Outcome of pregnancy for baby(n)	PREGA21	PREGA22	PREGA23	PREGA24	PREGA25	PREGA26	PREGA27	PREGA28	PREGA29	PREGA30
Name of baby(n)	PREGB21	PREGB22	PREGB23	PREGB24	PREGB25	PREGB26	PREGB27	PREGB28	PREGB29	PREGB30
Sex of baby(n)	PREGC21	PREGC22	PREGC23	PREGC24	PREGC25	PREGC26	PREGC27	PREGC28	PREGC29	PREGC30
Birth weight of baby(n) – units	PREGD21	PREGD22	PREGD23	PREGD24	PREGD25	PREGD26	PREGD27	PREGD28	PREGD29	PREGD30
Birth weight of baby(n) pounds	POUND21	POUND22	POUND23	POUND24	POUND25	POUND26	POUND27	POUND28	POUND29	POUND30
Birth weight of baby(n) ounces	OUNCE21	OUNCE22	OUNCE23	OUNCE24	OUNCE25	OUNCE26	OUNCE27	OUNCE28	OUNCE29	OUNCE30
Birth weight of baby(n) kilos	KILO21	KILO22	KILO23	KILO24	KILO25	KILO26	KILO27	KILO28	KILO29	KILO30
Birth weight of baby(n) grammes	GRAMM21	GRAMM22	GRAMM23	GRAMM24	GRAMM25	GRAMM26	GRAMM27	GRAMM28	GRAMM29	GRAMM30
baby(n) day of live or still birth	PREGED21	PREGED22	PREGED23	PREGED24	PREGED25	PREGED26	PREGED27	PREGED28	PREGED29	PREGED30
baby(n) mnth of live/still birth, miscarriage/termination	PREGEM21	PREGEM22	PREGEM23	PREGEM24	PREGEM25	PREGEM26	PREGEM27	PREGEM28	PREGEM29	PREGEM30
baby(n) year of live/still birth, miscarriage/termination	PREGFY21	PREGFY22	PREGFY23	PREGFY24	PREGFY25	PREGFY26	PREGFY27	PREGFY28	PREGFY29	PREGFY30
baby(n) date of birth	PREGF21	PREGF22	PREGF23	PREGF24	PREGF25	PREGF26	PREGF27	PREGF28	PREGF29	PREGF30
Was baby(n) early, late or on time	PREGF21	PREGF22	PREGF23	PREGF24	PREGF25	PREGF26	PREGF27	PREGF28	PREGF29	PREGF30
How many weeks early or late was baby(n)	PREGG21	PREGG22	PREGG23	PREGG24	PREGG25	PREGG26	PREGG27	PREGG28	PREGG29	PREGG30
Was anything wrong with baby(n) at birth	PREGH21	PREGH22	PREGH23	PREGH24	PREGH25	PREGH26	PREGH27	PREGH28	PREGH29	PREGH30
What was the problem with baby(n)	PREGI21	PREGI22	PREGI23	PREGI24	PREGI25	PREGI26	PREGI27	PREGI28	PREGI29	PREGI30
Is CMs current partner baby(n)s other parent	WHOPAR32	WHOPAR34	WHOPAR36	WHOPAR38	WHOPAR40	WHOPAR42	WHOPAR44	WHOPAR46	WHOPAR48	WHOPAR50
Who is baby(n)s other parent	WHOPAR33	WHOPAR35	WHOPAR37	WHOPAR39	WHOPAR41	WHOPAR43	WHOPAR45	WHOPAR47	WHOPAR49	WHOPAR51
Does baby(n) live with CM	WHERKI21	WHERKI22	WHERKI23	WHERKI24	WHERKI25	WHERKI26	WHERKI27	WHERKI28	WHERKI29	WHERKI30
What is baby(n) currently doing	WHATKI21	WHATKI22	WHATKI23	WHATKI24	WHATKI25	WHATKI26	WHATKI27	WHATKI28	WHATKI29	WHATKI30
Does baby(n) ever see other parent	ABSPAR43	ABSPAR46	ABSPAR49	ABSPAR52	ABSPAR55	ABSPAR58	ABSPAR61	ABSPAR64	ABSPAR67	ABSPAR70
How often does baby(n) see other parent	ABSPAR44	ABSPAR47	ABSPAR50	ABSPAR53	ABSPAR56	ABSPAR59	ABSPAR62	ABSPAR65	ABSPAR68	ABSPAR71
Does the other parent of baby(n) pay maintenance	ABSPAR45	ABSPAR48	ABSPAR51	ABSPAR54	ABSPAR57	ABSPAR60	ABSPAR63	ABSPAR66	ABSPAR69	ABSPAR72
Where is baby(n) living now	ABSKID65	ABSKID70	ABSKID75	ABSKID80	ABSKID85	ABSKID90	ABSKID95	ABSKI100	ABSKI105	ABSKI110
Has baby(n) ever lived with CM	ABSKID66	ABSKID71	ABSKID76	ABSKID81	ABSKID86	ABSKID91	ABSKID96	ABSKI101	ABSKI106	ABSKI111
Year baby(n) last lived with CM	ABSYR21	ABSYR22	ABSYR23	ABSYR24	ABSYR25	ABSYR26	ABSYR27	ABSYR28	ABSYR29	ABSYR30
Month baby(n) last lived with CM	ABSMON21	ABSMON22	ABSMON23	ABSMON24	ABSMON25	ABSMON26	ABSMON27	ABSMON28	ABSMON29	ABSMON30
Does CM ever see baby(n) now	ABSKID67	ABSKID72	ABSKID77	ABSKID82	ABSKID87	ABSKID92	ABSKID97	ABSKI102	ABSKI107	ABSKI112
How often does CM see baby(n)	ABSKID68	ABSKID73	ABSKID78	ABSKID83	ABSKID88	ABSKID93	ABSKID98	ABSKI103	ABSKI108	ABSKI113
Does CM pay maintenance for baby(n)	ABSKID69	ABSKID74	ABSKID79	ABSKID84	ABSKID89	ABSKID94	ABSKID99	ABSKI104	ABSKI109	ABSKI114
Another pregnancy before this one?					MOREPRE5					MOREPRE6

Pregnancy History Variables (continued)

Pregnancy Baby	7					8				
	1	2	3	4	5	1	2	3	4	5
Has CM ever been or got anyone else pregnant										
(Number of babies carried)	PREGNUM7					PREGNUM8				
How long had CM & partner not been using birth control	PREGJ7					PREGJ8				
Months not been using birth control	PREGK7					PREGK8				
Weeks not been using birth control	PREGKW7					PREGKW8				
Years not been using birth control	PREGL7					PREGL8				
MC:Whether smoked during pregnancy	CGPREG19					CGPREG22				
MC:Whether smoked during pregnancy	CGPREG20					CGPREG23				
MC:Whether smoked during pregnancy	CGPREG21					CGPREG24				
Whether smoked more, less or same as before	CPREGB7					CPREGB8				
Outcome of pregnancy for baby(n)	PREGA31	PREGA32	PREGA33	PREGA34	PREGA35	PREGA36	PREGA37	PREGA38	PREGA39	PREGA40
Name of baby(n)	PREGB31	PREGB32	PREGB33	PREGB34	PREGB35	PREGB36	PREGB37	PREGB38	PREGB39	PREGB40
Sex of baby(n)	PREGC31	PREGC32	PREGC33	PREGC34	PREGC35	PREGC36	PREGC37	PREGC38	PREGC39	PREGC40
Birth weight of baby(n) - units	PREGD31	PREGD32	PREGD33	PREGD34	PREGD35	PREGD36	PREGD37	PREGD38	PREGD39	PREGD40
Birth weight of baby(n) pounds	POUND31	POUND32	POUND33	POUND34	POUND35	POUND36	POUND37	POUND38	POUND39	POUND40
Birth weight of baby(n) ounces	OUNCE31	OUNCE32	OUNCE33	OUNCE34	OUNCE35	OUNCE36	OUNCE37	OUNCE38	OUNCE39	OUNCE40
Birth weight of baby(n) kilos	KILO31	KILO32	KILO33	KILO34	KILO35	KILO36	KILO37	KILO38	KILO39	KILO40
Birth weight of baby(n) grammes	GRAMM31	GRAMM32	GRAMM33	GRAMM34	GRAMM35	GRAMM36	GRAMM37	GRAMM38	GRAMM39	GRAMM40
baby(n) day of live or still birth	PREGED31	PREGED32	PREGED33	PREGED34	PREGED35	PREGED36	PREGED37	PREGED38	PREGED39	PREGED40
baby(n) mnth of live/still birth, miscarriage/termination	PREGEM31	PREGEM32	PREGEM33	PREGEM34	PREGEM35	PREGEM36	PREGEM37	PREGEM38	PREGEM39	PREGEM40
baby(n) year of live/still birth, miscarriage/termination	PREGFY31	PREGFY32	PREGFY33	PREGFY34	PREGFY35	PREGFY36	PREGFY37	PREGFY38	PREGFY39	PREGFY40
baby(n) date of birth	PREGF31	PREGF32	PREGF33	PREGF34	PREGF35	PREGF36	PREGF37	PREGF38	PREGF39	PREGF40
Was baby(n) early, late or on time	PREGG31	PREGG32	PREGG33	PREGG34	PREGG35	PREGG36	PREGG37	PREGG38	PREGG39	PREGG40
How many weeks early or late was baby(n)	PREGH31	PREGH32	PREGH33	PREGH34	PREGH35	PREGH36	PREGH37	PREGH38	PREGH39	PREGH40
Was anything wrong with baby(n) at birth	PREGI31	PREGI32	PREGI33	PREGI34	PREGI35	PREGI36	PREGI37	PREGI38	PREGI39	PREGI40
What was the problem with baby(n)	WHOPAR52	WHOPAR54	WHOPAR56	WHOPAR58	WHOPAR60	WHOPAR62	WHOPAR64	WHOPAR66	WHOPAR68	WHOPAR70
Is CMs current partner baby(n)s other parent	WHOPAR53	WHOPAR55	WHOPAR57	WHOPAR59	WHOPAR61	WHOPAR63	WHOPAR65	WHOPAR67	WHOPAR69	WHOPAR71
Who is baby(n)s other parent	WHERKI31	WHERKI32	WHERKI33	WHERKI34	WHERKI35	WHERKI36	WHERKI37	WHERKI38	WHERKI39	WHERKI40
Does baby(n) live with CM	WHATKI31	WHATKI32	WHATKI33	WHATKI34	WHATKI35	WHATKI36	WHATKI37	WHATKI38	WHATKI39	WHATKI40
What is baby(n) currently doing	ABSPAR73	ABSPAR76	ABSPAR79	ABSPAR82	ABSPAR85	ABSPAR88	ABSPAR91	ABSPAR94	ABSPAR97	ABSPA100
Does baby(n) ever see other parent	ABSPAR74	ABSPAR77	ABSPAR80	ABSPAR83	ABSPAR86	ABSPAR89	ABSPAR92	ABSPAR95	ABSPAR98	ABSPA101
How often does baby(n) see other parent	ABSPAR75	ABSPAR78	ABSPAR81	ABSPAR84	ABSPAR87	ABSPAR90	ABSPAR93	ABSPAR96	ABSPAR99	ABSPA102
Does the other parent of baby(n) pay maintenance	ABSKI115	ABSKI120	ABSKI125	ABSKI130	ABSKI135	ABSKI140	ABSKI145	ABSKI150	ABSKI155	ABSKI160
Where is baby(n) living now	ABSKI116	ABSKI121	ABSKI126	ABSKI131	ABSKI136	ABSKI141	ABSKI146	ABSKI151	ABSKI156	ABSKI161
Has baby(n) ever lived with CM	ABSYR31	ABSYR32	ABSYR33	ABSYR34	ABSYR35	ABSYR36	ABSYR37	ABSYR38	ABSYR39	ABSYR40
Year baby(n) last lived with CM	ABSMON31	ABSMON32	ABSMON33	ABSMON34	ABSMON35	ABSMON36	ABSMON37	ABSMON38	ABSMON39	ABSMON40
Month baby(n) last lived with CM	ABSKI117	ABSKI122	ABSKI127	ABSKI132	ABSKI137	ABSKI142	ABSKI147	ABSKI152	ABSKI157	ABSKI162
Does CM ever see baby(n) now	ABSKI118	ABSKI123	ABSKI128	ABSKI133	ABSKI138	ABSKI143	ABSKI148	ABSKI153	ABSKI158	ABSKI163
How often does CM see baby(n)	ABSKI119	ABSKI124	ABSKI129	ABSKI134	ABSKI139	ABSKI144	ABSKI149	ABSKI154	ABSKI159	ABSKI164
Does CM pay maintenance for baby(n)					MOREPRE7					MOREPRE8
Another pregnancy before this one?										

APPENDIX 2: Proxy Interview Variables

Where the cohort member was unable to understand or respond to questions put by the interviewer or to the self-completion, short proxy interviews were undertaken with a family member or carer. The variables which hold the data for the proxy interview are identifiable through their labels – all begin with the endorsement “(Proxy)”. These variables and associated labels are listed below.

Variable	Label	Sequential position
PROXTHK	(Proxy) Thank you for taking part in this interview	31
PROXYWHO	(Proxy) What is your relationship to CM	50
PROXYAC1	(Proxy) MC: CM been married, had children, a job or Ed?	51
PROXYAC2	(Proxy) MC: CM been married, had children, a job or Ed?	52
PROXYAC3	(Proxy) MC: CM been married, had children, a job or Ed?	53
PROXYAC4	(Proxy) MC CM been married, had children, a job or Ed?	54
LSIIMWK	(Proxy) (LSI1) Illness/disability that limits work CM can do	186
LSIAGE	(Proxy) (LSI1) Age when condition started	187
LSILIM	(Proxy) (LSI1) Does illness limit activities?	188
MORE11	(Proxy) (LSI1) Other illness, disability or infirmity	189
LSIIMWK2	(Proxy) (LSI2) Illness/disability that limits work CM can do	190
LSIAGE2	(Proxy) (LSI2) Age when condition started	191
LSILIM2	(Proxy) (LSI2) Does CM condition limit activities?	192
MORE12	(Proxy) (LSI2) Does CM have any other condition?	193
LSIIMWK3	(Proxy) (LSI3) Illness/disability that limits work CM can do	194
LSIAGE3	(Proxy) (LSI3) Age when condition started	195
LSILIM3	(Proxy) (LSI3) Does CM condition limit activities?	196
MORE13	(Proxy) (LSI3) Does CM have any other condition?	197
LSIIMWK4	(Proxy) (LSI4) Illness/disability that limits work CM can do	198
LSIAGE4	(Proxy) (LSI4) Age when condition started	199
LSILIM4	(Proxy) (LSI4) Does CM condition limit activities?	200
MORE14	(Proxy) (LSI4) Does CM have any other condition?	201
LSIIMWK5	(Proxy) (LSI5) Illness/disability that limits work CM can do	202
LSIAGE5	(Proxy) (LSI5) Age when condition started	203
LSILIM5	(Proxy) (LSI5) Does CM condition limit activities?	204
MORE15	(Proxy) (LSI5) Does CM have any other condition?	205
LSIIMWK6	(Proxy) (LSI6) Illness/disability that limits work CM can do	206
LSIAGE6	(Proxy) (LSI6) Age when condition started	207
LSILIM6	(Proxy) (LSI6) Does CM condition limit activities?	208
MORE16	(Proxy) (LSI6) Does CM have any other condition?	209
LSIIMWK7	(Proxy) (LSI7) Illness/disability that limits work CM can do	210
LSIAGE7	(Proxy) (LSI7) Age when condition started	211
LSILIM7	(Proxy) (LSI7) Does CM condition limit activities?	212
MORE17	(Proxy) (LSI7) Does CM have any other condition?	213
LSIIMWK8	(Proxy) (LSI8) Illness/disability that limits work CM can do	214
LSIAGE8	(Proxy) (LSI8) Age when condition started	215
LSILIM8	(Proxy) (LSI8) Does CM condition limit activities?	216
MORE18	(Proxy) (LSI8) Does CM have any other condition?	217
LSIIMWK9	(Proxy) (LSI9) Illness/disability that limits work CM can do	218
LSIAGE9	(Proxy) (LSI9) Age when condition started	219
LSILIM9	(Proxy) (LSI9) Does CMs condition limit activities?	220
MORE19	(Proxy) (LSI10) Does CM have any other condition?	221

Variable	Label	Sequential position
LSIIMW10	(Proxy) (LSI0) Illness/disability that limits work CM can do	222
LSIAGE10	(Proxy) (LSI10) Age when condition started	223
LSILIM10	(Proxy) (LSI10) Does CM condition limit activities?	224
MORE20	(Proxy) Does CM have any other condition?	225
HOAGE	(Proxy) (HC1) Age when CM first had condition?	226
HOSUP1	(Proxy)(HC1) MC Who monitors this condition	227
HOSUP2	(Proxy) (HC1) MC Who monitors this condition	228
HOSUP3	(Proxy) (HC1) MC Who monitors this condition	229
HOMORE	(Proxy) (HC1) Any other health condition	230
HOAGE2	(Proxy) (HC2) Age when CM first had condition?	231
HOSUP4	(Proxy) (HC2) MC Who monitors this condition	232
HOSUP5	(Proxy) (HC2) MC Who monitors this condition	233
HOSUP6	(Proxy) (HC2) MC Who monitors this condition	234
HOMORE2	(Proxy) (HC2) Any other health condition	235
HOAGE3	(Proxy) (HC3) Age when CM first had condition?	236
HOSUP7	(Proxy) (HC3) MC Who monitors this condition	237
HOSUP8	(Proxy) (HC3) MC Who monitors this condition	238
HOSUP9	(Proxy) (HC3) MC Who monitors this condition	239
HOMORE3	(Proxy) (HC3) Any other health condition	240
HOAGE4	(Proxy) (HC4) Age when CM first had condition?	241
HOSUP10	(Proxy) (HC4) MC Who monitors this condition	242
HOSUP11	(Proxy) (HC4) MC Who monitors this condition	243
HOSUP12	(Proxy) (HC4) MC Who monitors this condition	244
HOMORE4	(Proxy) (HC4) Any other health condition	245
HOAGE5	(Proxy) (HC5) Age when CM first had condition?	246
HOSUP13	(Proxy) (HC5) MC Who monitors this condition	247
HOSUP14	(Proxy) (HC5) MC Who monitors this condition	248
HOSUP15	(Proxy) (HC5) MC Who monitors this condition	249
HOMORE5	(Proxy) (HC5) Any other health condition	250
HOAGE6	(Proxy) (HC6) Age when CM first had condition?	251
HOSUP16	(Proxy) (HC6) MC Who monitors this condition	252
HOSUP17	(Proxy) (HC6) MC Who monitors this condition	253
HOSUP18	(Proxy) (HC6) MC Who monitors this condition	254
HOMORE6	(Proxy) (HC6) Any other health condition	255
HOAGE7	(Proxy) (HC7) Age when CM first had condition?	256
HOSUP19	(Proxy) (HC7) MC Who monitors this condition	257
HOSUP20	(Proxy) (HC7) MC Who monitors this condition	258
HOSUP21	(Proxy) (HC7) MC Who monitors this condition	259
HOMORE7	(Proxy) (HC7) Any other health condition	260
HOAGE8	(Proxy) (HC8) Age when CM first had condition?	261
HOSUP22	(Proxy) (HC8) MC Who monitors this condition	262
HOSUP23	(Proxy) (HC8) MC Who monitors this condition	263
HOSUP24	(Proxy) (HC8) MC Who monitors this condition	264
HOMORE8	(Proxy) (HC8) Any other health condition	265
HOAGE9	(Proxy) (HC9) Age when CM first had condition?	266
HOSUP25	(Proxy) (HC9) MC Who monitors this condition	267
HOSUP26	(Proxy) (HC9) MC Who monitors this condition	268
HOSUP27	(Proxy) (HC9) MC Who monitors this condition	269
HOMORE9	(Proxy) (HC9) Any other health condition	270

Variable	Label	Sequential position
HOAGE10	(Proxy) (HC10)Age when CM first had condition?	271
HOSUP28	(Proxy) (HC10) MC Who monitors this condition	272
HOSUP29	(Proxy) (HC10) MC Who monitors this condition	273
HOSUP30	(Proxy) (HC10) MC Who monitors this condition	274
HOMORE10	(Proxy) (HC10) Any other health condition	275
YEARIN	(Proxy) Year CM moved into current address	276
YEARM	(Proxy) Month CM moved into current address	277
TENURE	(Proxy) 'Does CM own or rent home'	278
MARSTAT	(Proxy) What is CMs legal marital status	279
PREGPROX	(Proxy) How many children does CM have	280
PROXYTYP	(Proxy)Whether CM was employed or self employed	281
PROXYECO	(Proxy) Was CM working FT or PT	282
CJSUP	(Proxy)Did CM have managerial duties	283
CJEMPS	(Proxy) How many employees were there	284
AGELFTED	(Proxy) CMs age when left FT ed	285
ACTAGEL	(Proxy) Whether still in FT education	286
FURTHED	(Proxy) Whether CM started FT education later	287
LFTMORED	(Proxy)Age CM left last period of education	288
EDQUALS	(Proxy) Any academic quals obtained since Ref date	289
EDQTYP01	(Proxy)MC: Quals obtained since Ref date	290
EDQTYP03	(Proxy)MC: Quals obtained since Ref date	292
EDQTYP04	(Proxy)MC: Quals obtained since Ref date	293
EDQTYP05	(Proxy)MC: Quals obtained since Ref date	294
EDQTYP06	(Proxy)MC: Quals obtained since Ref date	295
EDQTYP07	(Proxy)MC: Quals obtained since Ref date	296
EDQTYP08	(Proxy)MC: Quals obtained since Ref date	297
EDQTYP09	(Proxy)MC: Quals obtained since Ref date	298
EDQTYP10	(Proxy)MC: Quals obtained since Ref date	299
EDQTYP11	(Proxy)MC: Quals obtained since Ref date	300
EDQTYP12	(Proxy)MC: Quals obtained since Ref date	301
EDQTYP13	(Proxy)MC: Quals obtained since Ref date	302
VOCQUALS	(Proxy) Any vocational quals obtained since Refdate	303
VOCTYP01	(Proxy)MC: Voc. quals obtained since Ref date	304
VOCTYP02	(Proxy)MC: Voc. quals obtained since Ref date	305
VOCTYP03	(Proxy)MC: Voc. quals obtained since Ref date	306
VOCTYP04	(Proxy)MC: Voc. quals obtained since Ref date	307
VOCTYP05	(Proxy)MC: Voc. quals obtained since Ref date	308
VOCTYP06	(Proxy)MC: Voc. quals obtained since Ref date	309
VOCTYP07	(Proxy)MC: Voc. quals obtained since Ref date	310
VOCTYP08	(Proxy)MC: Voc. quals obtained since Ref date	311
VOCTYP09	(Proxy)MC: Voc. quals obtained since Ref date	312
VOCTYP10	(Proxy)MC: Voc. quals obtained since Ref date	313
VOCTYP11	(Proxy)MC: Voc. quals obtained since Ref date	314
LSIANY	(Proxy) Does CM have long-standing illness/disability?	315
LSIREG	(Proxy) Is CM a registered disabled person	316
HEIGHT	(Proxy) CM height without shoes - units	318
HTMETRES	(Proxy) CM height without shoes - metres	319
HTCMS	(Proxy) CM height without shoes - centimetres	320
HTFEET	(Proxy) CM height without shoes - feet	321

Variable	Label	Sequential position
HTINCHES	(Proxy) CM height without shoes - inches	322
WEIGHT	(Proxy) CM current weight without clothes - units	323
WTKILOS	(Proxy) CM current weight - kilos	324
WTSTONES	(Proxy) CM current weight - stones	325
WTPOUNDS	(Proxy) CM current weight - pounds	326
IFPROXY	Proxy interview indicator	6661
NMPRX	(Proxy) CM name	6662
NMPRX2	(Proxy) Person2 name	6664
NMPRX3	(Proxy) Person3 name	6666
NMPRX4	(Proxy) Person4 name	6668
NMPRX5	(Proxy) Person5 name	6670
NMPRX6	(Proxy) Person6 name	6672
NMPRX7	(Proxy) Person7 name	6674
NMPRX8	(Proxy) Person8 name	6676
NMPRX9	(Proxy) Person9 name	6678
NMPRX10	(Proxy) Person10 name	6680
SEXPRX	(Proxy) CM gender	6682
SEXPRX2	(Proxy) Person2 gender	6684
SEXPRX3	(Proxy) Person3 gender	6686
SEXPRX4	(Proxy) Person4 gender	6688
SEXPRX5	(Proxy) Person5 gender	6690
SEXPRX6	(Proxy) Person6 gender	6692
SEXPRX7	(Proxy) Person7 gender	6694
SEXPRX8	(Proxy) Person8 gender	6696
SEXPRX9	(Proxy) Person9 gender	6698
SEXPRX10	(Proxy) Person10 gender	6700
AGEPRX	(Proxy) CM age last birthday	6702
AGEPRX2	(Proxy) Person2 age last birthday	6704
AGEPRX3	(Proxy) Person3 age last birthday	6706
AGEPRX4	(Proxy) Person4 age last birthday	6708
AGEPRX5	(Proxy) Person5 age last birthday	6710
AGEPRX6	(Proxy) Person6 age last birthday	6712
AGEPRX7	(Proxy) Person7 age last birthday	6714
AGEPRX8	(Proxy) Person8 age last birthday	6716
AGEPRX9	(Proxy) Person9 age last birthday	6718
AGEPRX10	(Proxy) Person10 age last birthday	6720
MSPRX	(Proxy) CM Current partnership status	6722
MSPRX2	(Proxy) Person2 Current partnership status	6724
MSPRX3	(Proxy) Person 3 Current partnership status	6726
MSPRX4	(Proxy) Person4 Current partnership status	6728
MSPRX5	(Proxy) Person5 Current partnership status	6730
MSPRX6	(Proxy) Person6 Current partnership status	6732
MSPRX7	(Proxy) Person7 Current partnership status	6734
MSPRX8	(Proxy) Person8 Current partnership status	6736
MSPRX9	(Proxy) Person9 Current partnership status	6738
MSPRX10	(Proxy) Person10 Current partnership status	6740
RELPRX	(Proxy) Relationship to CM (always blank for CM)	6742
RELPRX2	(Proxy) Person2 Relationship to CM	6744
RELPRX3	(Proxy) Person3 Relationship to CM	6746

Variable	Label	Sequential position
RELPRX4	(Proxy) Person4 Relationship to CM	6748
RELPRX5	(Proxy) Person5 Relationship to CM	6750
RELPRX6	(Proxy) Person6 Relationship to CM	6752
RELPRX7	(Proxy) Person7 Relationship to CM	6754
RELPRX8	(Proxy) Person8 Relationship to CM	6756
RELPRX9	(Proxy) Person9 Relationship to CM	6758
RELPRX10	(Proxy) Person10 Relationship to CM	6760
CMSEXPX	(Proxy) Interviewer check on CM gender	6763
BDAT1PX	(Proxy) Check: what date in April is CMs birthday	6765
BDAT2PX	(Proxy) Check: what date in Match is CMs birthday	6767
CMNAMPRX	(Proxy) CM name from sample file	6769
RENAMPRX	(Proxy) Interviewer: record name change on ARF	6771
RESIDPRX	(Proxy) Type of accommodation	6774
NORMPRX	(Proxy) Does CM usually live at this address	6775

APPENDIX 3: NCDS/BCS70 Advisors contributing to quality assessment of NCDS/BCS70 data

A list of the advisors and their associates who contributed to the assessment of the quality of the data, and whose work is drawn on above is given in below. The contribution of individual advisors is acknowledged below. Wherever appropriate, the detailed documentation provided by advisors is included in Appendices 4-8.

Dr Ann Berrington	University of Southampton
Louise Blackwell	Centre For Longitudinal Studies, Institute of Education, University of London
Lorraine Dearden	Institute for Fiscal Studies
Alissa Goodman	Institute for Fiscal Studies
Barbara Jefferis	Institute of Child Health
Andrew Jenkins	Institute of Education, University of London
Gerald Makepeace	University of Cardiff
Scott Montgomery	Karolinska Sjukhuset, Stockholm, Sweden
Barbara Maughan	Institute of Psychiatry
Samantha Parsons	Centre For Longitudinal Studies, Institute of Education, University of London
Dr Gaëlle Pierre	Institute for Employment Research
John Preston	Centre For Longitudinal Studies, Institute of Education, University of London
Richard Rowe	Institute of Psychiatry

APPENDIX 4: Earnings and related variables

Documentation produced by Alissa Goodman and Lorraine Dearden of the Institute for Fiscal Studies relating to earnings and related variables is reproduced below. (**NB:** References to Appendix 1 and Appendix 2 in the original documentation have been changed below to Annex 1 and Annex 2 to avoid confusion with Appendices 1 and 2 above)

Queries about the contents of this appendix should be directed to the User Support Group
(cohort@cls.ioe.ac.uk)

Derived Variables.

Section 1. Interview date

1. Corrected interview date

Name: lintdate

Description

Corrected interview date

Purpose

To correct interview dates that were originally coded as outside the interviewing period. In the original data there were 23 individuals whose interview dates were not between 10/99 and 9/00.

Source variables

Intdate

Code

See intdat.do

Section 2. Deflators

2. RPI at interview date

Name: rpi_int

Description

Retail price index (RPI) for month of interview, where Jan 2001 = 1.000.

Purpose

To allow wages, incomes etc. to be converted into other prices. To uprate wages to Jan 2001 prices, divide by this deflator.

Source variables

I_intdate, Retail price index (RPI).

Frequency distribution

See Annex 1, Table 2.

Section 3. Employment variables

3. Total usual weekly hours

Name: Hours

Description

Total usual weekly hours, including paid and unpaid overtime, not including meal breaks.

Purpose

To derive a measure of total hours usually worked, in order to convert gross and net pay information available into an hourly pay measure.

Source variables

Chours1 Chours2 Chours3 Chours4, Otimeany.

This is calculated as: Chours1 + Chours2 + Chours3 + Chours4, where missing values of each of these four variables are set to zero. This ensures that these variables do not include missing values (ie so that no 999s are added in to the total). Notice that individuals have either responses at Chours1 (where Otimeany=2), or responses at Chours2, Chours3 or Chours4 (where Otimeany=1), but never both.

Note this variable is set to missing (.) for all those who a) are not employees b) have missing values in all of the above source variables.

Code

(Stata do-file format): see checkwage.do

Frequency distribution

See Annex, Table 3.

4. Imputed/Corrected net last pay

Name: lnetpay

Description

Imputed or corrected net last pay variable.

Purpose

Clean last net pay variable, corrected for implausible values.

Source variables

Cnetpay, Cgropay, cnetprd, cgroprd, cnetpred, cgorpred, hours.

Source code

See checkwage.do

Frequency distribution

See Annex 1, Table 4.

Further information about corrections made: see Annex 2

5. Imputed/Corrected net period

Name: lnetprd

Description

Imputed or corrected net last pay PERIOD variable.

Purpose

Clean last net pay PERIOD variable, corrected for implausible values and imputed from gross where net unavailable

Source variables

Cnetpay, Cgropay, cnetprd, cgroprd, cnetpred, cgorpred, hours.

Frequency distribution

See Annex 1, Table 5.

Further information about corrections made: see Annex 2

6. Imputed/Corrected gross pay

Name: lgropay

Description

Imputed or corrected gross last pay variable.

Purpose

Clean last gross pay variable, corrected for implausible values and imputed from gross where net unavailable

Source variables

Cnetpay, Cgropay, cnetprd, cgroprd, cnetpred, cgorpred, hours,

Source code

See checkwage.do

Frequency distribution

See Annex 1, Table 6.

Further information about corrections made: see Annex 2.

7. Imputed/Corrected gross period

Name: lgroprd

Description

Imputed or corrected gross last pay PERIOD variable.

Purpose

Clean last gross pay PERIOD variable, corrected for implausible values and imputed from gross where net unavailable

Source variables

Cnetpay, Cgropay, cnetprd, cgroprd, cnetpred, cgorpred, hours,

Further information about corrections made: see Annex 2.

Frequency distribution

See Annex 1, Table 7.

8. Hourly net last pay

Name: hr_net

Description:

Hourly net last pay.

Purpose

Clean hourly equivalent last net pay. Where net pay variables are missing but gross pay variables are available, net pay has been imputed from gross using known parameters of the tax system in the relevant year.

Source Variables

InetPay, InetPrd, Inetpred, IgroPay, Igorprd, IgroPred, Hours, ms, age2-age10

Code

Stata do file format: see checkwage.do, impute.do

Frequency distribution

See Annex 1, Table 8.

9. Annual net last pay

Name: ann_net

Description:

Clean annual equivalent net last pay. Where net pay variables are missing but gross pay variables are available, net pay has been imputed from gross using known parameters of the tax system in the relevant year.

Purpose

Clean annual equivalent last net pay, with missing values imputed wherever possible.

Source Variables

InetPay, InetPrd, Inetpred, IgroPay, Igorprd, IgroPred, Hours, ms, age2-age10

Code

Stata do file format: checkwage.do, impute.do

Frequency distribution

See Annex 1, Table 9.

10. Hourly gross last pay

Name: hr_gro

Description:

Clean hourly equivalent last gross pay. Where gross pay variables are missing but net pay variables are available, gross pay has been imputed from net using known parameters of the tax system in the relevant year.

Purpose

Clean hourly equivalent gross last pay, with missing values imputed wherever possible.

Source Variables

IgroPay, IgroPrd, IgroPred, Inetpay, Inetprd, Inetpred Hours, ms, age2-age10

Code

Stata do file format: see checkwage.do, impute.do.

Frequency distribution

See Annex 1, Table 10.

11. Annual gross last pay

Name: ann_gro

Description:

Clean annual equivalent last gross pay. Where gross pay variables are missing but net pay variables are available, gross pay has been imputed from net using known parameters of the tax system in the relevant year.

Purpose

Clean annual equivalent gross last pay with missing values imputed wherever possible.

Source Variables

I_groPay, I_groPrd, I_groPred, Inetpay, Inetprd, Inetpred, Hours, ms, age2-age10

Code

Stata do file format: checkwage.do, impute.do

Frequency distribution

See Annex 1, Table 11.

Section 4. Partner's earnings variables.

12. Partner's Imputed/Corrected last net pay

Name: lpNetpay

Description:

Partner's last net pay variable, implausible values corrected.

Purpose

Clean partner's last net pay.

Source Variables

Pnetpay, pnetprd, pnetpred

Code

Stata do file format: partner.do

Frequency distribution

See Annex 1, Table 12.

13. Partner's Imputed/Corrected last net pay PERIOD

Name: lpNetprd

Description:

Partner's last net pay PERIOD variable, implausible values corrected.

Purpose

Clean partner's last net pay PERIOD.

Source Variables

Pnetpay, pnetprd, pnetpred

Code

Stata do file format: partner.do

Frequency distribution

See Annex 1, Table 13.

14. Partner's Imputed/Corrected last net pay PERIOD

Name: lpnpred

Description:

Partner's last net pay PERIOD variable, implausible values corrected.

Purpose

Clean partner's last net pay PERIOD.

Source Variables

Pnetpay, pnetprd, pnetpred

Code

Stata do file format: partner.do

Frequency distribution

See Annex 1, Table 14.

15. Partner's annual last net pay

Name: ann_pnet

Description:

Annual equivalent of partner's last net pay, implausible values corrected.

Purpose

Clean annual equivalent partner's last net pay

Source Variables

lpNetpay, lpNetprd, lpNetpred

Code

Stata do file format: partner.do

Frequency distribution

See Annex 1, Table 15.

16. Partner's IMPUTED annual last gross pay

Name: ann_pgro

Description:

Annual equivalent of partner's last gross pay, imputed from partner's last net pay using the known parameters of the tax system in the relevant year.

Purpose

Clean annual equivalent partner's last gross pay - imputed from net.

Source Variables

lpnetpay, lpnetprd, lpnetpred, lintdate, ms

Code

Stata do file format: partner.do

Frequency distribution

See Annex 1, Table 16.

Section 5. Intermediate variables – used during cleaning.

17. Type of fix to pay variable and period codes

Name: fixtype

Description

Categorical variable containing the type of coding error (if any) which we have identified in the pay information of each employee, and the fix which has been implemented.

Purpose

To identify and record corrections to last net and gross pay variables.

Code

See checkwage.do

Frequency distribution

See Annex 1, Table 17.

Further information See Annex 2.

18. Indicator of odd proportion of gross to net pay

Name: oddprop

Description

Indicator of whether ratio of gross to net pay suggests coding error in some of the pay variables which REMAINS UNCORRECTED.

Purpose

To identify odd wage information

Code

Oddprop = 1 if (Hr_net/hr_gro) <0.4 or >4

See checkwage.do.

See Annex 1, Table 18.

Further information, including description and frequency distributions: See Annex 2

19. Indicator of wage variable probably too high or too low

Name: oddwage

Description

Indicator of whether pay variables remain apparently implausibly high or low AFTER CORRECTIONS have been implemented.

Purpose

To identify odd wage information

Code

Oddwage =1 id (Hr_gro<2) or (Hr_gro>100)

See checkwage.do

See Annex 1, Table 19.

Further information, including description and frequency distributions: See Annex 2

20. Missing hours variable

Name: zhours.

Description

Indicator of Usual weekly hours variable (hours) missing.

Purpose

Code

See checkwage.do

Frequency distribution

See Annex 1, Table 20.

21. Missing hourly net pay indicator

Name: zhr_net.

Description

Indicator of if/ why net hourly pay variable is missing.

Purpose

Code

See checkwage.do

Frequency distribution

See Annex 1, Table 21.

22. Missing hourly gross pay indicator

Name: zhr_gro.

Description

Indicator of if/ why gross hourly pay variable is missing.

Purpose

Code

See checkwage.do

Frequency distribution

See Annex 1, Table 22.

23. Odd partner's wage indicator

Name: podd.

Description

Indicator of apparently implausibly high or low partner's net wage, remaining AFTER CORRECTIONS have been implemented.

Purpose

Code

Podd=1 if (assuming fixed number of hours for part-time and full-time work) implied hourly net pay of partner is <2 or >100.

See partner.do

Frequency distribution

See Annex 1, Table 23.

ANNEX 1 FREQUENCY DISTRIBUTIONS

The frequency distributions are available on request from the User Support Group
(cohort@cls.ioe.ac.uk)

ANNEX 2: Explaining IFS suggested fixes to CnetPay, CnetPrd, CgroPay, CgroPrd.

Derived variables used for identifying coding errors in the pay information:

- 1) **Fixtype.** This is a derived categorical variable taking values 0-20. It contains the type of coding error (if any) which we have identified in the pay information of each employee, and the suggested fix.
- 2) **Oddprop.** This is a (0,1) dummy variable, which is set to 1 if AFTER ALL SUGGESTED CORRECTIONS, the ratio of net to gross hourly pay is still less than 0.4 or greater than 4.
- 3) **Oddwage** This is a (0,1) dummy variable, which is set to 1 if AFTER ALL SUGGESTED CORRECTIONS, gross hourly pay is still less than £2 per hour or more than £100 per hour.

How coding errors have been identified

In order to identify possible coding errors, we have converted all gross and net pay given to its hourly equivalent (wherever possible – there are a few cases of missing hours variables). We have done this using CnetPay, CnetPrd, Cnetpred, CgroPay, CgroPrd, CgroPred, Chours1-Chours4 (and in a small number of cases where pay is given daily, we also use EconAct). [All our derived variables and code will be deposited in due course.]

Some of the coding errors have been spotted because the ratio of net hourly pay to gross hourly pay is clearly wrong. We have considered just those whose ratio of net to gross hourly pay is over 4, or below 0.4. By looking at the data for these observations it has in many cases been fairly obvious what the coding error has been and how it should be corrected. Those with these coding errors (and their suggested solutions) are contained in fixtype 1-12. Those whose ratio of gross to net is wrong, but no obvious fix can be seen are coded as Fixtype=19.

Other coding errors have been identified because both gross and net hourly pay are either implausibly high or implausibly low. We have considered those whose implied gross hourly pay is over £100 per hour, or less than £2 per hour. All the coding errors for those whose implied hourly pay is too high or too low are contained in fixtype 13-18. Where it is not clear that a coding error has occurred, or the fix required is not obvious, fixtype is set to 20.

Note that these are not necessarily the only coding errors contained in the data. Those considered are those where the ratio of net to gross hourly pay is >4 or <0.4 , and those whose implied gross hourly pay is less than £2 or more than £100 per hour. These are arbitrary cut-offs, and a wider net could be considered if needs be.

Notice also that a judgement has to be made in the absence of any further information, for those coded in fixtype 8, 14 and fixtype 17 (and possibly others) where a number of different corrections could be made to the data to make it more plausible, and we have just chosen one possible solution. In particular in these cases it is not clear whether the problem is too many (or too few) zeros in the pay amount variables, or whether the period codes are wrong. We have suggested that it is the period codes are wrong, but CLS may take a different view.

For those who are coded in fixtype= 19 or fixtype=20, or for whom oddprop=1 or oddwage=1 despite the suggested fixes already put in place, we will have to make some decisions on an individual case-by-case basis whether to leave the information as it is, correct it in some way, or discard it.

The coding frame for Fixtype

The table below sets out the coding frame for the fixtype variable: it describes what the miscoding appears to be, and what the suggested correction is. The corrections are simply suggestions; in many cases other possible also corrections exist: see below.

Coding frame for variable fixtype

	FIXTYPE
0	No fixes made.
1	Gross pay given annually needs to be multiplied by 10.
2	Net pay given monthly needs to be divided by 100.
3	Net pay given monthly needs to be divided by 10.
4	Net pay given monthly or four weekly needs to be multiplied by 100.
5	Net pay given monthly or four weekly needs to be multiplied by 10.
6	Net monthly pay is coded as yearly: CnetPrd needs to be changed to 4.
7	Gross yearly pay is coded as monthly or 4-weekly: CGroPrd needs to be changed to 5.
8*	Gross yearly pay coded as weekly: period code needs to be changed to 5.
9	Gross pay given yearly needs to be divided by 10.
10	Gross monthly pay is coded as weekly: CGroPrd needs to be changed to 4.
11	Individual fixes: see code provided below.
12	Gross pay set to missing.
13	Net and Gross pay need to be divided by 100.
14*	Net and Gross monthly pay is coded as yearly: CnetPrd and CGroPrd need to be changed to 4.
15	Net and Gross weekly pay is coded as yearly: CnetPrd and CgroPrd need to be changed to 4.
16	Total hours variable (derived by IFS) is too low and incompatible with self-reported economic status. Set hours to missing.
17*	Net and Gross yearly pay is coded as monthly: CnetPrd and CgroPrd need to be changed to 5.
18	Net and Gross pay have both been coded as 1: set to missing.
19	No fix yet: proportion of net to gross pay is odd, but fix needed is not clear.
20	No fix yet: Hourly pay appears very high or low, but fix needed is not clear.

*Sometimes there are a number of possible fixes which would make the pay data look more sensible, and it is not entirely clear what the right one is.

In particular:

Fixtype 8: it has been assumed that the mistake made is that gross annual pay has been coded as weekly. In many cases it is also plausible that it could be gross weekly pay which needs to be divided by 100.

Fixtype 14: it has been assumed that the mistake made is that monthly pay has been coded as annual. In fact it could also be the case that it is in fact annual pay which needs to be multiplied by 10 (re-coding to monthly effectively multiplies by 12).

Fixtype 17: it has been assumed that the mistake made is that annual pay has been coded as monthly. Equally plausible is that monthly pay has been provided which needs to be divided by 10 (re-coding the period code to annual effectively multiplies it by 12).

Details of individual fixes already made.

Those coded fixtype=11 require individual fixes, set out below.

Individual fixes

Serial	Fix
Changes to Cgroprd	
112006	Cgroprd recoded to 1
101716	Cgroprd recoded to 4
213363	Cgroprd recoded to 4
215256	Cgroprd recoded to 6 Cgropred recoded to 15
Changes to Cnetprd	
226017	Cnetprd recoded to 4
229217	Cnetprd recoded to 4
200306	Cnetprd recoded to 1 (although net becomes a bit too high compared to gross)
117261	Cnetprd recoded to 1 (although net becomes a bit too high compared to gross)
219899	Cnetprd recoded to 1 (although net becomes a bit too high compared to gross)

208399	Cnetprd recoded to 1 (although net becomes a bit too high compared to gross)
104694	Cnetprd recoded to 1 (although net becomes a bit too high compared to gross)
105206	Cnetprd recoded to 1 (although net becomes a bit too high compared to gross)
109860	Cnetprd recoded to 2 (gross pay is given fortnightly and this looks more plausible)
Changes to cnetpay	
119813	Cnetpay should be equal to £100
124299	Cnetpay should be divided by 10
213092	Cnetpay should be multiplied by 10
229425	Cnetpay should be multiplied by 10
Changes to cgropay	
121730	Cgropay should be multiplied by 1000
210049	Cgropay should be divided by 10
219518	Cgropay should be divided by 10
228991	Cgropay should be divided by 10 {based on fact that real net weekly pay in NCDS 5 was £605.30 ie. In the region of £600, not £6000 as implied here}

Frequency distribution for fixtype

type of fix to net and/ or gross pay	Freq.	Percent	Cum.
no fix	15443	95.20	95.20
gro*10	134	0.83	96.03
net/100	4	0.02	96.05
net/10	17	0.10	96.16
net*100	3	0.02	96.18
net*10	9	0.06	96.23
net year to month	167	1.03	97.26
gro month to year	97	0.60	97.86
gro week to year	10	0.06	97.92
gro/10	15	0.09	98.01
gro week to month	15	0.09	98.11
individual fix	22	0.14	98.24
gro to missing	3	0.02	98.26
both/100	14	0.09	98.35
both year to month	135	0.83	99.18
both year to week	9	0.06	99.24
hours to missing	14	0.09	99.32
both month to year	11	0.07	99.39
no fix- prop still odd	11	0.07	99.46
no fix- wages still od	88	0.54	100.00
Total	16221	100.00	

Further serial numbers under question: oddprop and oddwage

After these "corrections" set out above have been carried out, there are still a number of individuals whose wages or ratio of net to gross pay are still odd. We will have to consider these on a case-by-case basis and decide whether to leave the information as it is, correct it in some way, or discard the information all together.

The frequency distributions for oddprop and oddwage are as follows.

ODDPROP (ratio net: gross hourly pay <0.4 or >4)

odd ratio of net to gross hourly pay	Freq.	Percent	Cum.
0	16464	99.91	99.91
1	15	0.09	100.00
Total	16479	100.00	

ODDWAGE (gross hourly wage >£100 or < £2)

Strangely high or low gross hrly wage	Freq.	Percent	Cum.
0	16379	99.39	99.39
1	100	0.61	100.00
Total	16479	100.00	

Source code

For reference, the exact stata code which has been used to derive the fixtype variable is also attached as a stata do-file, checkwage.do.

STATA CODE:

- CHECKWAGE.DO
 - IMPUTE.DO
 - PARTNER.DO
-

CHECKWAGE.DO

```
/* WAGES DATA */
#delimit;
set more 1;

/*cd "j:\2000cohorts\derived data";
/* TO MAKE SURE THIS DATASET IS CREATED, RUN GETDATA.DO */
u wagedata,clear; */

/* SET ORIGINAL VARIABLES cnetpay, cnetprd, cgropay, cgroprd IN STONE IN
onetpay, ogropay etc.
corrections and imputations are then added to cnetpay... */

cap drop onetpay;
gen onetpay=cnetpay;
lab var onetpay "original net pay variable";

cap drop ogropay;
gen ogropay=cgropay;
lab var ogropay "original gross pay variable";

cap drop onetprd;
gen onetprd=cnetprd;
lab var onetprd "original net period variable";

cap drop ogroprd;
gen ogroprd=cgroprd;
lab var ogroprd "original gross period variable";

/**** EMPLOYMENT STATUS : will be useful later *****/
cap drop employee;
gen employee =1 if econact<=2;
replace employee =0 if employee ~1;
lab var employee "in paid employment";

/***** INITIAL MISSING GROSS AND NET PAY VARIABLES
*****/
/* NOTE these will change later, as some nonsense values get set to missing
but we need to use them in the interim */
/* missing cnetpay */
cap drop znetpay;
gen znetpay =1 if cnetpay==9999998|cnetpay==9999999;
replace znetpay =0 if znetpay~1;
replace znetpay =. if cnetpay==.;
lab var znetpay "missing cnetpay";
```

```

/* missing cgropay */
cap drop zgropay;
gen zgropay      =1          if cgropay==9999998|cgropay==9999999;
replace zgropay  =0          if zgropay~=1;
replace zgropay  =.          if cgropay==.;
lab var zgropay  "missing cgropay";

/***** HOURS *****/
/*****/

/* have to set up USUAL WEEKLY HOURS before period codes, to be able to
convert pay into hourly */

cap pr drop dohours;
pr def dohours;

cap drop hours;

cap drop zhours;
gen      zhours      = 1          if chours1==. & chours2==. & chours3==.
& chours4==. & employee==1;
replace  zhours      = 0          if zhours~=1;
replace  zhours      = .          if employee==0;
lab var  zhours      "missing hours";

/* People either get coded at chours1 or at chours2-chours4. This makes
it safe to add them all up to get a total, once missing values
are taken care of */
/* For chours1, value 99 appears to be missing value. For chours2-
chours4, value 999 is missing */

/* Note: a few of these hours variables are set to missing later if they
are
implausibly low and so mess up the hourly pay variables: see FIX 14
below */

cap drop hours;
cap drop hrs1;
cap drop hrs2;
cap drop hrs3;
cap drop hrs4;
cap drop paidhrs;

gen hrs1= chours1 if chours1<99;
gen hrs2= chours2 if chours2<998;
gen hrs3= chours3 if chours3<998;
gen hrs4= chours4 if chours4<998;

replace hrs1 = 0 if hrs1==.;
replace hrs2 = 0 if hrs2==.;
replace hrs3 = 0 if hrs3==.;
replace hrs4 = 0 if hrs4==.;

gen      hours      =          hrs1 + hrs2 + hrs3 + hrs4 if zhours==0
& employee==1;
gen      paidhrs    =          hrs1 + hrs2 + hrs3          if zhours==0
& employee==1;

```

```

/* some have total hours recorded as zero */
replace    zhours      = 1                                if hours ==0
& employee==1;
replace    hours       = .                                if zhours ==1
& employee==1;

lab var hours "total usual weekly hours";
lab var paidhrs "total paid hours";

gr hours, bin(50) xlab(10,20,30,40,50,60,70,80,90,100,120,140,170)
                ylab(0,0.05,0.1,0.15,0.2)
                t1("distribution of hours");

cap log close;
log using hours, replace;
    tab hours;
    tab zhours;
log close;
end;
dohours;

/***** PAY PERIOD CODES *****/

/* the coding frame for CNETPROD and CGROPRD is:
    1 one week
    2 a fortnight
    3 four weeks
    4 a calendar month
    5 a year
    6 other period

the coding frame for CNETPRED and CGROPRED is:
    1 one week
    2 a fortnight/2 weeks
    3 four weeks
    4 a calendar month
    5 a year (12 months/52 weeks)
    6 three weeks
    7 five weeks
    8 six weeks
    9 seven weeks
    10 eight weeks
    11 two calendar months
    12 eight times a year
    13 nine times a year
    14 ten times a year
    15 three months/13 weeks
    16 six months/26 weeks
    17 hourly
    18 daily
    19 one off lump sum
    20 some other period
    21 varies
    22 refused
    23 some other comment
    24 irrelevant/unspecific response */

```

```

/***** NET PAY VARIABLES
*****/

/* First make sure not to start manipulating cnetpay and cgropay missing
values 9999997 and 9999998 ! */

replace cnetpay =. if znetpay==1;
replace cgropay =. if zgropay==1;

/* need to guess how many days per week someone works to convert any daily
pay into weekly
not too important as it only affects 7 individuals, with a max of 50
hours per week */
/* use paid hours, as we want to sum daily pay over the number of days
they get paid for
to get total weekly pay */

cap drop days;
gen days = paidhrs/8;
replace days = 7 if days>=7;
lab var days "days per week (estimated)";

/* this program converts net pay into different periods depending on the
period code */
cap pr drop donet;
pr def donet;

cap drop ann_net;
cap drop wk_net;
cap drop hr_net;
cap drop mo_net;
cap drop znetprd;

gen wk_net = (cnetpay) if cnetprd==1
& znetpay==0;
replace wk_net = (cnetpay / 2) if cnetprd==2
& znetpay==0;
replace wk_net = (cnetpay / 4) if cnetprd==3
& znetpay==0;
replace wk_net = (cnetpay / (13/3)) if cnetprd==4
& znetpay==0;
replace wk_net = (cnetpay / 52) if cnetprd==5
& znetpay==0;
replace wk_net = 9999997 if cnetprd==.
& employee==1
& znetpay==0;

replace wk_net = (cnetpay) if cnetprd==6 &
cnetpred==1
& znetpay==0;
replace wk_net = (cnetpay / 2) if cnetprd==6 &
cnetpred==2
& znetpay==0;
replace wk_net = (cnetpay / 3) if cnetprd==6 &
cnetpred==3
& znetpay==0;
replace wk_net = (cnetpay / (13/3)) if cnetprd==6 &
cnetpred==4
& znetpay==0;

```

```

replace wk_net      =      (cnetpay / 52)                if cnetprd==6 &
cnetprd==5        & znetpay==0;
replace wk_net      =      (cnetpay / 3)                if cnetprd==6 &
cnetprd==6        & znetpay==0;
replace wk_net      =      (cnetpay / 5)                if cnetprd==6 &
cnetprd==7        & znetpay==0;
replace wk_net      =      (cnetpay / 6)                if cnetprd==6 &
cnetprd==8        & znetpay==0;
replace wk_net      =      (cnetpay / 7)                if cnetprd==6 &
cnetprd==9        & znetpay==0;
replace wk_net      =      (cnetpay / 8)                if cnetprd==6 &
cnetprd==10       & znetpay==0;
replace wk_net      =      (cnetpay / (26/3))           if cnetprd==6 &
cnetprd==11       & znetpay==0;
replace wk_net      =      (cnetpay / (52/8))           if cnetprd==6 &
cnetprd==12       & znetpay==0;
replace wk_net      =      (cnetpay / (52/9))           if cnetprd==6 &
cnetprd==13       & znetpay==0;
replace wk_net      =      (cnetpay / (52/10))          if cnetprd==6 &
cnetprd==14       & znetpay==0;
replace wk_net      =      (cnetpay / 13)               if cnetprd==6 &
cnetprd==15       & znetpay==0;
replace wk_net      =      (cnetpay / 26)               if cnetprd==6 &
cnetprd==16       & znetpay==0;
replace wk_net      =      (cnetpay * paidhrs)           if cnetprd==6 &
cnetprd==17 & zhours==0 & znetpay==0;
replace wk_net      =      (cnetpay * days)             if cnetprd==6 &
cnetprd==18 & zhours==0 & znetpay==0;
replace wk_net      =      9999997                     if cnetprd==6 &
cnetprd>=19 & cnetprd~=. & znetpay==0;
replace wk_net      =      9999997                     if cnetprd==6 &
cnetprd==17 & zhours==1 & znetpay==0;
replace wk_net      =      9999997                     if cnetprd==. &
employee==1;
replace wk_net      =      .                             if employee==0;

```

```
lab var wk_net      "weekly net pay";
```

```
/* missing net period */
```

```
gen znetprd        =      1 if wk_net==9999997;
replace znetprd    =      0 if employee==1 & znetprd~=1;
replace znetprd    =      . if employee==0;
lab var znetprd    "missing net period code";
```

```
replace wk_net     =      . if znetprd==1;
```

```
/* annual net pay */
```

```
gen ann_net        =      (wk_net*52)                  if znetprd==0;
replace ann_net=.  if znetprd==1;
lab var ann_net    "annual net pay";
```

```
/* hourly net pay */
```

```
gen hr_net         =      (wk_net/ hours)              if zhours==0 &
znetprd==0;
```

```
/* can get hourly pay even if hours missing if cnetpay is given hourly */
```

```

replace hr_net      =          cnetpay if cnetprd==6 & cnetpred==17 &
zhours==1;
lab var hr_net      "hourly net pay";

/* monthly net pay */
gen mo_net          =          (wk_net * (13/3))          if znetprd==0;
lab var mo_net      "monthly net pay" ;

end;
donet;

/***** GROSS PAY VARIABLES *****/

/* this program converts gross pay into different periods depending on the
period code */

cap pr drop dogross;
pr def dogross;

cap drop ann_gro;
cap drop wk_gro;
cap drop hr_gro;
cap drop mo_gro;
cap drop zgroprd;

gen wk_gro          =          (cgropay)                  if cgroprd==1
& zgropay==0;
replace wk_gro      =          (cgropay / 2)              if cgroprd==2
& zgropay==0;
replace wk_gro      =          (cgropay / 4)              if cgroprd==3
& zgropay==0;
replace wk_gro      =          (cgropay / (13/3))         if cgroprd==4
& zgropay==0;
replace wk_gro      =          (cgropay / 52)             if cgroprd==5
& zgropay==0;
replace wk_gro      =          9999997                   if cgroprd==. &
employee==1
& zgropay==0;
replace wk_gro      =          (cgropay)                  if cgroprd==6 &
cgropred==1
& zgropay==0;
replace wk_gro      =          (cgropay / 2)              if cgroprd==6 &
cgropred==2
& zgropay==0;
replace wk_gro      =          (cgropay / 3)              if cgroprd==6 &
cgropred==3
& zgropay==0;
replace wk_gro      =          (cgropay / (13/3))         if cgroprd==6 &
cgropred==4
& zgropay==0;
replace wk_gro      =          (cgropay / 52)             if cgroprd==6 &
cgropred==5
& zgropay==0;
replace wk_gro      =          (cgropay / 3)              if cgroprd==6 &
cgropred==6
& zgropay==0;
replace wk_gro      =          (cgropay / 5)              if cgroprd==6 &
cgropred==7
& zgropay==0;
replace wk_gro      =          (cgropay / 6)              if cgroprd==6 &
cgropred==8
& zgropay==0;
replace wk_gro      =          (cgropay / 7)              if cgroprd==6 &
cgropred==9
& zgropay==0;
replace wk_gro      =          (cgropay / 8)              if cgroprd==6 &
cgropred==10
& zgropay==0;

```



```

replace wk_gro      =      (cgropay / (26/3))      if cgroprd==6 &
cgropred==11      & zgropay==0;
replace wk_gro      =      (cgropay / (52/8))      if cgroprd==6 &
cgropred==12      & zgropay==0;
replace wk_gro      =      (cgropay / (52/9))      if cgroprd==6 &
cgropred==13      & zgropay==0;
replace wk_gro      =      (cgropay / (52/10))     if cgroprd==6 &
cgropred==14      & zgropay==0;
replace wk_gro      =      (cgropay / 13)         if cgroprd==6 &
cgropred==15      & zgropay==0;
replace wk_gro      =      (cgropay / 26)         if cgroprd==6 &
cgropred==16      & zgropay==0;
replace wk_gro      =      (cgropay * paidhrs)     if cgroprd==6 &
cgropred==17 & zhours==0 & zgropay==0;
replace wk_gro      =      (cgropay * days)       if cgroprd==6 &
cgropred==18 & zhours==0 & zgropay==0;
replace wk_gro      =      9999997               if cgroprd==6 &
cgropred>=19 & cgropred~=. & zgropay==0;
replace wk_gro      =      9999997               if cgroprd==9; /*
one odd case here */
replace wk_gro      =      9999997               if cgroprd==. &
employee==1;
replace wk_gro      =      .                     if employee==0;

lab var wk_gro      "weekly gross pay";

gen zgroprd         =      1 if (wk_gro==9999997);
replace zgroprd     =      0 if (employee<=2 & zgroprd~=1);
replace zgroprd     =      . if employee==0;
lab var zgroprd     "missing gross period code";

replace wk_gro      =      . if zgroprd==1;

gen ann_gro         =      (wk_gro*52);
lab var ann_gro     "annual gross pay";

gen hr_gro          =      (wk_gro/ hours) if zhours==0;
/* can get hourly pay even if hours missing if cgropay is given hourly */
replace hr_gro      =      cgropay if cgroprd==6 & cgropred==17 &
zhours==1;
lab var hr_gro      "hourly gross pay";

gen mo_gro          =      (wk_gro * (13/3));
lab var mo_gro      "monthly gross pay";

end;
dogross;

```

```

/*****FIXES TO PAY INFORMATION
*****/

```

```

/* store type of fix for reference */

```

```

/* Some are fixed because the ratio of net pay to gross pay is clearly
wrong, and looking at the data
it is obvious what the coding error has been.

```

Others are fixed because both gross and net pay are both clearly miscoded.

Also some hours are miscoded */

```
cap drop fixtype;
gen fixtype=0 if employee==1;
lab var fixtype "type of fix to net and/ or gross pay";
```

```
lab def fixtype
    0 "no fix"
    1 "gro*10"
    2 "net/100"
    3 "net/10"
    4 "net*100"
    5 "net*10"
    6 "net year to month"
    7 "gro month to year"
    8 "gro week to year"
    9 "gro/10"
    10 "gro week to month"
    11 "individual fix"
    12 "gro to missing"
    13 "both/100"
    14 "both year to month"
    15 "both year to week"
    16 "hours to missing"
    17 "both month to year"
    18 "both to missing"
    19 "no fix- prop still odd"
    20 "no fix- wages still odd", modify;
lab val fixtype fixtype;
```

```
/****** CHECK NET AGAINST GROSS FOR ODD AMOUNTS OR PERIOD CODES *****/
```

```
cap pr drop doprop;
pr def doprop;

cap drop hr_prop;
gen hr_prop = hr_net/hr_gro;
lab var hr_prop "ratio of net to gross hourly pay";
```

```
cap drop oddprop;
gen oddprop =1 if (hr_prop<0.4|hr_prop>4) & hr_prop~=. ;
lab var oddprop "odd ratio of net to gross hourly pay";
cap log close;
```

```
end;
doprop;
```

```
log using oddprop, replace;
di "no fixes yet";
tab oddprop;
log off;
```

```
/******
*   FIX 1   *
*****/
```

```
/* lots of gross pay given annually needs to be multiplied by 10! */
```

```

replace fixtype=1          if (cgroprd==5|cgropred==5) & hr_prop>=5 &
hr_prop~= . & cgropay<=6000;
replace cgropay= cgropay*10  if fixtype==1;

donet;
dogross;
doprop;
log using oddprop, append;
    di "after fix 1:";
    tab oddprop;
log off;

/*****
*     FIX 2     *
*****/
/* there are a couple of monthly net pay which need to be divided by 100 */

replace fixtype=2          if (cnetprd==4|cnetpred==4) & hr_prop>=50 &
hr_prop~= . & cnetpay>=80000;
replace cnetpay = cnetpay/100 if fixtype==2;
donet;
dogross;
doprop;
log using oddprop, append;
    di "after fix 2:";
    tab oddprop;
log off;

/*****
*     FIX 3     *
*****/
/* there are some monthly net pay which need to be divided by 10 */

replace fixtype=3          if (cnetprd==4|cnetpred==4) & hr_prop>=5 &
hr_prop~= . & cnetpay>=5000;
replace cnetpay= cnetpay/10  if fixtype==3;
donet;
dogross;
doprop;
log using oddprop, append;
    di "after fix 3:";
    tab oddprop;
log off;

/*****
*     FIX 4     *
*****/
/* there are some net monthly and four weekly pay that need to be
multiplied by 100 */

replace fixtype=4          if
(cnetprd==4|cnetprd==3|cnetpred==4|cnetpred==3) & hr_prop<0.2 & cnetpay<50;

replace cnetpay = cnetpay *100 if fixtype==4;

donet;
dogross;

```

```

doprop;
log using oddprop, append;
    di "after fix 4:";
    tab oddprop;
log off;

/*****
*     FIX 5     *
*****/
/* there are some net monthly and four weekly pay that need multiplying
by 10 */

replace fixtype=5          if
(cnetprd==4|cnetprd==3|cnetpred==4|cnetpred==3) & hr_prop<0.2 &
cnetpay<500;
replace cnetpay = cnetpay *10 if fixtype==5 ;
donet;
dogross;
doprop;
log using oddprop, append;
    di "after fix 5:";
    tab oddprop;
log off;

/*****
*     FIX 6     *
*****/

/* lots of net monthly pay coded as yearly: period code needs to be changed
*/

replace fixtype=6          if (cnetprd==5|cnetpred==5) & cnetpay<=6000 &
hr_prop<=0.1;
replace cnetprd = 4          if fixtype==6;

donet;
dogross;
doprop;
log using oddprop, append;
    di "after fix 6:";
    tab oddprop;
log off;

/*****
*     FIX 7     *
*****/
/* there are some gross yearly pay coded as monthly and four weekly:
period code needs to be changed */

replace fixtype = 7          if (cgroprd==3| cgroprd==4| cgropred==3|
cgropred==4) & cgropay>=10000 & hr_prop<0.2;
replace cgroprd = 5          if fixtype==7;
donet;
dogross;
doprop;
log using oddprop, append;
    di "after fix 7:";
    tab oddprop;

```

```

log off;

/*****
*   FIX 8   *
*****/

/* there are some gross yearly pay coded as weekly: period code needs to be
changed */

replace fixtype = 8          if (cgroprd==1|cgropred==1) & cgropay>=6000 &
hr_prop<0.2;
replace cgroprd = 5          if fixtype==8;
donet;
dogross;
doprop;
log using oddprop, append;
  di "after fix 8:";
  tab oddprop;
log off;

/*****
*   FIX 9   *
*****/

/* there are some gross yearly pay which need to be divided by 10 */

replace fixtype= 9          if (cgroprd==5|cgropred==5) & cgropay>=100000 &
hr_prop<0.2;
replace cgropay= cgropay/10 if fixtype==9;

donet;
dogross;
doprop;
log using oddprop, append;
  di "after fix 9:";
  tab oddprop;
log off;

/*****
*   FIX 10  *
*****/

/* there are some gross monthly pay coded as weekly: period code needs to
be changed */

replace fixtype= 10 if (cgroprd==1 | cgropred==1) & cgropay>=350 &
hr_prop<0.2;
replace cgroprd= 4 if fixtype==10;

donet;
dogross;
doprop;
log using oddprop, append;
  di "after fix 10:";
  tab oddprop;

  log off;

/*****
*   FIX 11  *
*****/

/* INDIVIDUAL FIXES - these come from eyeballing individual cases */

```

```

/* fixes to cgroprd */
replace cgroprd=1 if serial==          112006;
replace cgroprd=4 if serial==          101716;
replace cgroprd=4 if serial==          213363;

replace cgroprd=6 if serial ==          215256; /* net pay is 3 monthly
and this looks more plausible */
replace cgroprd=15 if serial==          215256;

/* fixes to cnetprd */
replace cnetprd=4 if serial==          226017;
replace cnetprd=4 if serial==          229217;
replace cnetprd=1 if serial==          200306; /* from monthly to
weekly, although makes net a bit high compared to gross */
replace cnetprd=1 if serial==          117261; /* from monthly to
weekly, although makes net a bit high compared to gross */
replace cnetprd=1 if serial==          219899; /* from monthly to
weekly, although makes net a bit high compared to gross */
replace cnetprd=1 if serial==          208399; /* from monthly to
weekly, although makes net a bit high compared to gross */
replace cnetprd=1 if serial==          104694; /* from monthly to
weekly, although makes net a bit high compared to gross */
replace cnetprd=1 if serial==          105206; /* from monthly to
weekly, although makes net a bit high compared to gross */
replace cnetprd=2 if serial==          109860; /* gross pay is given
fortnightly, and this looks more plausible */

/* fixes to cnetpay: make sure changes only get made to original miscoded
value: ie. don't keep dividing
or multiplying after the problem is already fixed */

replace cnetpay= onetpay/10 if serial==  124299;
replace cnetpay= onetpay*10 if serial==  213092;
replace cnetpay= onetpay*10 if serial==  229425;
replace cnetpay=100 if serial== 119813; /* this one gets changed
around by my different fixes, but eyeballing shows that it should be £100
*/
replace cnetpay= onetpay*10 if serial==  111346;
/* fixes to cgropay */
replace cgropay= ogropay*1000 if serial== 121730;
replace cgropay= ogropay/10 if serial== 210049;
replace cgropay= ogropay/10 if serial== 219518;

donet;
dogross;
doprop;
log using oddprop, append;
di "after individual fixes ";
tab oddprop;
log off;

/* update fixtype variable */

replace fixtype = 11 if serial==          112006;
replace fixtype = 11 if serial==          101716;
replace fixtype = 11 if serial==          213363;
replace fixtype = 11 if serial==          215256;

```

```

replace fixtype = 11 if serial==          226017;
replace fixtype = 11 if serial==          229217;
replace fixtype = 11 if serial==          200306;
replace fixtype = 11 if serial==          117261;
replace fixtype = 11 if serial==          219899;
replace fixtype = 11 if serial==          208399;
replace fixtype = 11 if serial==          104694;
replace fixtype = 11 if serial==          105206;
replace fixtype = 11 if serial==          109860;

replace fixtype = 11 if serial==          124299;
replace fixtype = 11 if serial==          213092;
replace fixtype = 11 if serial==          229425;
replace fixtype = 11 if serial==          119813;
replace fixtype = 11 if serial==          121730;
replace fixtype = 11 if serial==          210049;
replace fixtype = 11 if serial==          219518;
replace fixtype = 11 if serial==          111346;

/*****
*   FIX 12   *
*****/
/* SET SOME GROSS TO MISSING
   if net looks sensisble, but gross looks rubbish, set gross to missing
and then impute from net later */

replace cgropay=. if serial==          219802;
replace cgropay=. if serial==          209108;
replace cgropay=. if serial==          100831;

/* update fixtype variable */
replace fixtype=12 if serial==          219802;
replace fixtype=12 if serial==          209108;
replace fixtype=12 if serial==          100831;

/* update zgropay variable */
replace zgropay=1 if serial==          219802;
replace zgropay=1 if serial==          209108;
replace zgropay=1 if serial==          100831;

donet;
dogross;
doprop;
log using oddprop, append;
   di "after some set to missing ";
   tab oddprop;
log off;

/***** ODD LOOKING WAGES
--- WHERE GROSS AND NET ARE IN LINE BUT BOTH ARE WRONGLY
CODED IN SOME WAY *****/

so hr_gro;

```

```

cap log close;
log using oddwage, replace;

cap pr drop oddwage;
pr def oddwage;
  cap drop oddwage;
  gen oddwage = 1 if      (hr_gro>100 & hr_gro~=. ) | hr_gro<2;
  replace oddwage= 0 if oddwage~1;
  replace oddwage=. if employee==0;
  lab var oddwage "Strangely high or low gross hrly wage ";
end;
oddwage;

/* record numbers that look too high or too low */
cap log close
log using oddwage, append;
  di "NO FIXES YET ";
  tab oddwage;
cap log close;

/*****
*   FIX 13   *
*****/
/*there are a number where the ratio of net to gross is fine, but both need
to be divided by 100 */

      /* Note these are guesses, but it would appear that full amounts
including pence have been provided,
      but coded without the decimal point. In practice this applies to
all gross hourly wages over £360 per hour ... */

replace fixtype =13                if hr_gro> 360 & hr_gro~=. ;
replace cnetpay=onetpay/100        if fixtype==13;
replace cgropay=ogropay/100       if fixtype==13;

donet;
dogross;
oddwage;

cap log close;
log using oddwage, append;
  di "after FIX 13 ";
  tab oddwage;
log off;

/*****
*   FIX 14   *
*****/
/* there are MANY where the ratio of net to gross is fine, but
monthly pay has been recorded as annual.

NOTE: A PLAUSIBLE ALTERNATIVE IS that both have lost a zero, and
need to be multiplied by 10 (rather than 12 which is what we are
implicitly doing here)

Change it to monthly for now - can reconsider

```


Note: only change those whose implied hourly pay is less than £2 per hour, although there are many more who also look suspect
 Also only change to monthly those whose implied hourly pay is more than 36.7p per hour ---
 this is because from looking at them, the ones below this should be coded at WEEKLY instead */

```
/* quick data description of who these people are */
cap log close;
log using oddwage, append;
      su hr_net hr_gro if cnetprd==5 & cgroprd==5 &
cnetpay<2500 & cgropay<3000 & econact==1;
      su hr_net hr_gro if cnetprd==5 & cgroprd==5 &
cnetpay<2500 & cgropay<3000 & econact==2;
cap log close;

      replace fixtype = 14                                if  cnetprd==5 & cgroprd==5
                                                    & cnetpay<2500 & cgropay<3000
                                                    & hr_gro>0.367 & hr_gro<2;

replace cnetprd=4 if fixtype ==14;
replace cgroprd=4 if fixtype ==14;

donet;
dogross;
oddwage;
cap log close;
log using oddwage, append;
  di "after FIX 14 ";
  tab oddwage;
log off;

/*****
*   FIX 15   *
*****/
/* there are SOME where weekly pay has been recorded as annual.

      From looking at them, those whose implied hourly pay is less
than 36.7p per hour ought to be
      coded as weekly */
```

```
replace fixtype = 15                                if  cnetprd==5 & cgroprd==5
                                                    & cnetpay<700 & cgropay<700
                                                    & hr_gro<0.367;

replace cnetprd=1 if fixtype ==15;
replace cgroprd=1 if fixtype ==15;

donet;
dogross;
oddwage;
cap log close;
log using oddwage, append;
  di "after FIX 15 ";
  tab oddwage;
log off;
```

```

/*****
*   FIX 16   *
*****/
/* there are a number with IMPLAUSIBLE, VERY low weekly hours, which are
not compatible
    with their self-reported work status, recorded in econact */
/* I'm not sure what to do with these? Set them to missing? Assume some
average hours for them?
    Set to missing for now */

replace fixtype=16 if hours<=10 & hr_gro>50 & hr_gro~=. & econact==1;
cap log close;
    log using oddwage, append;
    tab econact    if hours<=10 & hr_gro>50 & hr_gro<=.;
cap log close;

replace hours=0  if fixtype==16;
replace zhours=1 if fixtype==16;

dohours;
dogross;
donet;
oddwage;

cap log close;
log using oddwage, append;
    di "after FIX 16 ";
    tab oddwage;
log off;

/*****
*   FIX 17   *
*****/
/* there are some who have annual amounts recorded as monthly: change
period code to annual
    (* or else they could need to be divided by 10? see above? *)
*/

replace fixtype = 17 if (cnetprd==3| cnetprd==4) & cnetpay>10000 &
hr_net>100 & hr_net~=. ;
replace cnetprd = 5  if (cnetprd==3| cnetprd==4) & cnetpay>10000 &
hr_net>100 & hr_net~=. ;
replace fixtype = 17 if (cgroprd==3| cgroprd==4) & cgropay>10000 &
hr_gro>100 & hr_gro~=. ;
replace cgroprd = 5  if (cgroprd==3| cgroprd==4) & cgropay>10000 &
hr_gro>100 & hr_gro~=. ;

dogross;
donet;
oddwage;

cap log close;
log using oddwage, append;
    di "after FIX 17 ";
    tab oddwage;
log off;

/*****

```

```

*      FIX 18      *
*****/
/* there are 8 obs where cgropay and cnetpay both coded as 1 :
        clearly nothing to do with their pay, although (except
for 1 obs who "refused" to give
        period codes), period codes and hours work information
is provided.
                                Set to missing for now: can good
data for these be recovered? */

replace fixtype=18 if cnetpay==1 & cgropay==1 & hr_gro~=. ;
replace cnetpay=.  if fixtype==18;
replace znetpay=1  if fixtype==18;
replace cgropay=.  if fixtype==18;
replace zgropay=1  if fixtype==18;

donet;
dogross;
oddwage;

cap log close;
log using oddwage, append;
    di "after FIX 18 ";
    tab oddwage;
log off;

/*****
*      FIX 11: AGAIN *
*****/
/* some more individual fixes to gross and net pay */

/* WHAT TO DO??
1) Need to find a fix for serial 122904 : 10 net pay, 15 gross pay ---
multiply by 100?
    already has had net period code changed from annual to monthly to be
more in line with the gross
    period code, but this may not be the right fix in this case?

2) SERIAL 109130: monthly gross and net pay given as £55 for 56 hours per
week = 22p per hour

3) SERIAL 218767, 208683,230872, 20174: really high annual pay - see
below for comparison to wave 5 info */

li serial chrtid rw* wk_net _m if serial==218767;
li serial chrtid rw* wk_gro _m if serial==228991;
li serial chrtid rw* wk_net _m if serial==208683;
li serial chrtid rw* wk_net _m if serial==230872;
li serial chrtid rw* wk_net _m if serial==201714;

replace cgropay=ogropay/10 if serial ==228991; /* based on weekly net pay
in wave 5 of £605.305 -- more plausible that this one is in the region of
£600 not £6000 */
replace fixtype=11 if serial==228991;
drop if _m==2;

donet;
dogross;
oddwage;

```

```

cap log close;
log using oddwage, append;
    di "after more individual fixes:";
    tab oddwage;
log off;

/*****
*   FIXES 19 AND 20: NOT REALLY A FIX - *
*****/
/* if proportion of net to gross is still odd, but no fix has been made */
replace fixtype=19 if oddprop==1 & fixtype==0;

/* if gross hourly wage is still > £100 or < £2, but no fix has been made */
replace fixtype=20 if oddwage==1 & fixtype==0;

/***** SET UP SOME CATEGORY VARIABLES FOR REFERENCE *****/
/* Missing status variables for cnetpay and cgropay:
   some can be imputed from gross,
   for others, net and gross are both missing. */

/* both gross and net missing */
cap drop bothmiss;
gen bothmiss      =1          if znetpay==1 & zgropay==1;
replace bothmiss  =0          if bothmiss~=1;
replace bothmiss  =.          if employee==0;
lab var bothmiss  "both wages missing";

/* can impute net from gross */
cap drop impnet;
gen impnet        =1          if znetpay==1 & zgropay~=1;
replace impnet    =0          if impnet~=1;
replace impnet    =.          if employee==0;
lab var impnet    "can impute net from gross";

/* can impute gross from net */
cap drop impgro;
gen impgro        =1          if zgropay==1 & znetpay~=1;
replace impgro    =0          if impgro~=1;
replace impgro    =.          if employee==0;
lab var impgro    "can impute gross from net";

/* how to fix missing net pay */
cap drop fixnet;
gen fixnet        =1          if impnet==1;
replace fixnet    =2          if bothmiss==1;
replace fixnet    =0          if znetpay==0;
replace fixnet    =.          if employee==0;
lab var fixnet    "how to fix missing net pay";

```

```

lab def fixnet 0 "no fix needed" 1 "can impute from gross" 2 "set to
missing", modify;
lab val fixnet fixnet;

/*how to fix missing gross pay */
cap drop fixgro;
gen fixgro =1 if impgro==1;
replace fixgro =2 if bothmiss==1;
replace fixgro =0 if zgropay==0;
replace fixgro =. if employee==0;
lab var fixgro "how to fix missing gross wage";
lab def fixgro 0 "no fix needed" 1 "can impute from net" 2 "set to
missing", modify;
lab val fixgro fixgro;

replace fixtype=. if bothmiss==1;

/*****REASONS WHY NO HOURLY PAY
AVAILABLE *****/

cap drop zhr_net;

lab def zhr_net 0 "not missing"
1 "imputed from gross"
2 "missing pay: cannot impute"
3 "missing period code: cannot impute"
4 "missing hours, no hourly pay given" , modify;

gen zhr_net = 1 if fixnet==1;
replace zhr_net = 1 if fixnet==0 & znetprd==1 & zgropay==0 & zgroprd==0;
/* net period code is missing */
replace zhr_net = 2 if fixnet==2;
replace zhr_net = 3 if (fixnet==0| fixnet==1) & znetprd==1 & zgroprd==1;
replace zhr_net = 4 if (fixnet==0 & znetprd==0 & zhours==1)| (fixnet==1 &
zgroprd==0 & zhours==1);
replace zhr_net = 0 if employee==1 & zhr_net==.;
/* set back to not missing or can impute if pay is given hourly, or could
impute from gross pay given hourly */
replace zhr_net = 0 if cnetprd==6 & cnetprd==17 & zhours==1;
replace zhr_net = 1 if zhr_net==5 & fixnet==1 & cnetprd==6 & cnetprd==17 &
zhours==1;

replace zhr_net = . if employee==0;
lab val zhr_net zhr_net;
lab var zhr_net "missing hourly net pay";

cap drop zhr_gro;
gen zhr_gro = 1 if fixgro==1;
replace zhr_gro = 1 if fixgro==0 & zgroprd==1 & znetpay==0 & znetprd==0;
/* gross period code is missing */
replace zhr_gro = 2 if fixgro==2;
replace zhr_gro = 3 if zgroprd== 1 & znetprd ==1 & (fixgro==0 | fixgro==1);
replace zhr_gro = 4 if (fixgro==0 & zgroprd==0 & zhours==1)| (fixgro==1 &
znetprd==0 & zhours==1);
replace zhr_gro = 0 if employee==1 & zhr_gro==.;

```

```

/* set back to not missing or can impute if pay is given hourly, or could
impute from gross pay given hourly */
replace zhr_gro = 0 if cgroprd==6 & cgroprd==17 & zhours==1;
replace zhr_net = 1 if zhr_gro==5 & fixgro==1 & cnetprd==6 & cnetprd==17 &
zhours==1;

replace zhr_gro = . if employee==0;
lab def zhr_gro 0 "not missing"
1 "imputed from net"
2 "missing pay: cannot impute"
3 "missing period code: cannot impute"
4 "missing hours, no hourly pay given" , modify;

lab val zhr_gro zhr_gro;
lab var zhr_gro "missing hourly gro pay";

/***** do the imputations
*****/
do "e:\2000cohorts\do files\impute";
replace hr_gro= Ihr_gro if zhr_gro==1;
replace wk_gro= Iwk_gro if zhr_gro==1;
replace mo_gro= Imo_gro if zhr_gro==1;
replace ann_gro= Iann_gro if zhr_gro==1;

replace hr_net= Ihr_net if zhr_net==1;
replace wk_net= Iwk_net if zhr_net==1;
replace mo_net= Imo_net if zhr_net==1;
replace ann_net= Iann_net if zhr_net==1;

/***** inflate to Jan 01
*****/
/* these will have been merged in from Iintdate.dta
*/
cap drop rhr_gro;
gen rhr_gro = hr_gro/ rpi_int;
lab var rhr_gro "real gross hourly earnings, j01 prices";
cap drop lhw;
gen lhw = log(hr_gro);
lab var lhw "log real gross hourly earnings, j01 prices";

/***** code up some simple characteristics
*****/
cap drop male;
gen male =1 if cmsex==1;
replace male = 0 if cmsex==2;
lab var male "male dummy";

cap drop sampsex;
gen sampsex =1 if sample==1 & male==1;
replace sampsex =2 if sample==2 & male==1;
replace sampsex =3 if sample==1 & male==0;
replace sampsex =4 if sample==2 & male==0;

lab def sampsex 1 "bcs70 man" 2 "ncds man" 3 "bcs70 woman" 4 "ncds woman",
modify;
lab val sampsex sampsex;

cap drop bcsman;
cap drop ncdsman;

```

```

cap drop bcswo;
cap drop ncswom;

gen bcsman = 1 if sampse==1;
gen ncsmn = 1 if sampse==2;
gen bcswo = 1 if sampse==3;
gen ncswom = 1 if sampse==4;

replace bcsman = 0 if bcsman~=1;
replace ncsmn = 0 if ncsmn~=1;
replace bcswo = 0 if bcswo~=1;
replace ncswom = 0 if ncswom~=1;

/* log some results */
cap log close;

log using misspay, replace;

/* BASIC PAY VARIABLES */
tab fixnet sample, col freq;
tab fixgro sample, col freq;

/* HOURS */
tab zhours sample, col freq;

/* HOURLY PAY VARIABLES */
tab zhr_net sample, col freq;
tab zhr_gro sample, col freq;

tab zhr_net econact, col freq;
tab zhr_gro econact, col freq;

tab zhr_net cmsex, col freq;
tab zhr_gro cmsex, col freq;

/* FIXES TO PAY VARIABLES */
tab fixtype sample, col freq;

/***** hourly gross pay descriptives
*****/
su hr_gro, det;
table sampsex, c(median hr_gro);
table sampsex, c(mean hr_gro);
table sample econact, c(median hr_gro), if employee==1;
table sample econact, c(mean hr_gro), if employee==1;

table sample econact, c(median hr_gro), if cmsex==1 & employee==1;
table sample econact, c(mean hr_gro), if cmsex==2 & employee==1;

gr hr_gro if hr_gro<40 & oddwage==0 & oddprop==0, bin(50)
xlab(0,3,10,20,40) by(sampsex) ylab(0,0.04,0.08,0.12,0.16);
reg lhw ncsmn bcsman ncswom ;
log
close;

/*****FINALLY, SET NEW IMPUTED PAY AND PERIOD
VARIABLES,
AND RETURN CNETPAY, ETC. TO THEIR ORIGINALS
*****/

```

```
cap drop Inetpay;
gen Inetpay=cnetpay;
lab var Inetpay "imputed/corrected net pay variable";

cap drop Igropay;
gen Igropay=cgropay;
lab var Igropay "imputed/corrected gross pay variable";

cap drop Inetprd;
gen Inetprd=cnetprd;
lab var Inetprd "imputed/corrected net period variable";
label values Inetprd cnetprd;

cap drop Igroprd;
gen Igroprd=cgroprd;
lab var Igroprd "imputed/corrected gross period variable";
label values Igroprd cgroprd;

/* return cnet and gro variables to their originals */
replace cnetpay=onetpay;
replace cnetprd=onetprd;
replace cgropay=ogropay;
replace cgroprd=ogroprd;
```


IMPUTE.DO

```
/******IMPUTATION OF NET AND GROSS
PAY *****/
/* assume all are contracted IN to SERPs */
/* assume none get WFTC through pay packet */

# delimit;
set more 1;

/****** interview dates and tax year
******/
/* use corrected intdate variable once have got this together properly..*/
cap drop intmth;
cap drop intyr;
cap drop temp;
gen str2 temp= substr(intdate,3,2);
gen intmth= real(temp);

cap drop temp;
gen str4 temp = substr(intdate,5,4);
gen intyr= real(temp);

cap drop taxyear;
gen taxyear=1999 if ((intyr==1999)| (intyr==2000 & intmth<4));
replace taxyear=2000 if taxyear~=1999;
lab var taxyear "tax year";

/*** Because there was still the MCA in 1999, need to check if there are
any children *****/

/* sum across household members to see if any are children u19 */
cap pr drop makekid;
pr def makekid;
local loop = 2;
while `loop' <10 {;
/* include foster kids for now */
replace kidinh = kidinh+ 1 if
    reltoke`loop' ==3 & age`loop'<17 |
    reltoke`loop' ==4 & age`loop'<17 |
    reltoke`loop' ==5 & age`loop'<17 |
    reltoke`loop' ==6 & age`loop'<17 |
    reltoke`loop' ==7 & age`loop'<17;

replace maybek = maybek+1 if
    (reltoke`loop' ==3 & (age`loop'==18|age`loop'==19|
age`loop'==999)) |
    (reltoke`loop' ==4 & (age`loop'==18|age`loop'==19|
age`loop'==999)) |
    (reltoke`loop' ==5 & (age`loop'==18|age`loop'==19|
age`loop'==999)) |
    (reltoke`loop' ==6 & (age`loop'==18|age`loop'==19|
age`loop'==999)) |
    (reltoke`loop' ==7 & (age`loop'==18|age`loop'==19|
age`loop'==999));

local loop = `loop' +1;
```

```

};
end;

/***** need to know marital status and if there are children in the
household for MCA in 1999 */
cap drop kidinh;
gen kidinh=0;
lab var kidinh "kid u17 in household grid";
cap drop maybek;
gen maybek=0;
lab var maybek "kid 18,19, missing age in household grid";

makekid;

cap drop mca;
gen mca=1 if taxyear==1999 & ms==1;
replace mca = 1 if taxyear==1999 & (kidinh>0|maybek>0);
replace mca =0 if mca ~1;
lab var mca "getting married couples allowance?";

/***** some helpful indicators
*****/

/* can impute net from gross */
cap drop impnet;
gen impnet =1 if znetpay==1 & zgropay~1;
replace impnet =0 if impnet~1;
replace impnet =. if employee==0;
lab var impnet "can impute net from gross";

/* can impute gross from net */
cap drop impgro;
gen impgro =1 if zgropay==1 & znetpay~1;
replace impgro =0 if impgro~1;
replace impgro =. if employee==0;
lab var impgro "can impute gross from net";

/***** some parameters of the tax system
*****/

cap drop lel;
cap drop uel;
cap drop pa;
cap drop thresh1;
cap drop thresh2;
cap drop lowrate;
cap drop baserate;
cap drop highrate;

gen lel=67 if taxyear==2000;
replace lel=66 if taxyear==1999;
lab var lel "national insurance lower earnings limit";

gen uel =535 if taxyear==2000;

```

```

replace uel=500 if taxyear==1999;
lab var uel "national insurance upper earnings limit";

gen pa      = 4385 if taxyear==2000;
replace pa = 4335 if taxyear==1999;
lab var pa  "personal allowance";

gen thresh1 = 1520 if taxyear==2000;
replace thresh1 = 1500 if taxyear==1999;
lab var thresh1 "lower rate limit";

gen thresh2 = 28400 if taxyear==2000;
replace thresh2 = 28000 if taxyear==1999;
lab var thresh2 "basic rate limit";

gen lowrate= 0.1;
lab var lowrate "lower rate";

gen baserate = 0.22 if taxyear==2000;
replace baserate = 0.23 if taxyear==1999;
lab var baserate "basic rate";

gen highrate = 0.4;
lab var highrate "higher rate";

cap drop notax;
cap drop lowband;
cap drop baseband;
cap drop highband;

/* put people in income tax bands according to their net incomes */
gen notax =1 if (ann_net<=pa);
gen lowband =1 if (ann_net>pa) & (ann_net <= (pa + (0.9 * thresh1)));
gen baseband=1 if (ann_net > (pa + (0.9 * thresh1))) & (ann_net <= (pa +
(0.9 * thresh1) + ((1-baserate) * (thresh2-thresh1))));
gen highband=1 if (ann_net >=(pa + (0.9 * thresh1) + ((1-baserate) *
(thresh2-thresh1))));

/***** NATIONAL INSURANCE *****/
*****/

/* FROM GROSS */
cap drop wk_nig;

gen      wk_nig      = 0                                if wk_gro<l1el
;
replace wk_nig      = 0.1* (wk_gro-l1el)                if wk_gro>=l1el &
wk_gro<=uel      ;
replace wk_nig      = 0.1* (uel-l1el)                    if wk_gro>uel
;

lab var wk_nig "from gross: imputed weekly NI payments";

/* FROM NET */
cap drop wk_nin;

```

```

gen wk_nin = 0 if wk_net<=le1
;
replace wk_nin = ((wk_net-le1)/9) if wk_net>le1 &
wk_net<= uel*0.9 ;
replace wk_nin = (0.1* (uel- le1)) if wk_net>uel*0.9
;
lab var wk_nin "from net: imputed weekly NI payments";

/***** INCOME TAX *****/
/* tax is worked out separately for those whose income falls:
1. below personal allowance
2. in 10p band
3. in basic rate band
4. in higher rate band */

/* FROM GROSS */
cap drop taxann;
gen taxann = ann_gro-pa;

cap drop ann_itg;
gen ann_itg =0 if taxann <=0
;
replace ann_itg = lowrate * taxann if taxann > 0 &
(taxann < thresh1) ;
replace ann_itg = (lowrate * (thresh1)) + (baserate * (taxann -
thresh1)) if (taxann>= thresh1) & (
taxann< thresh2) ;
replace ann_itg = (lowrate * (thresh1)) + (baserate * (thresh2 -
thresh1)) + (highrate * (taxann - thresh2)) if (taxann>=thresh2)
;

/* sort out the MCA here: GIVE them the WHOLE mca for now
- ie don't try to assume split between partner or anything! */
replace ann_itg = ann_itg - 197 if mca==1;
replace ann_itg = 0 if ann_itg <0;

lab var ann_itg "from GROSS: imputed annual income tax payments";

/* FROM NET */

cap drop ann_itn;
gen ann_itn =0 if
notax==1;
replace ann_itn = (lowrate/(1-lowrate)) * (ann_net-pa) if
lowband==1;
replace ann_itn = ((baserate/(1-baserate))* (ann_net-pa)) - (((baserate-
lowrate) / (1-baserate)) * thresh1) if baseband==1;
replace ann_itn = ((highrate/(1-highrate))* (ann_net-pa)) - (((highrate-
baserate) / (1-highrate)) * thresh2) - (((baserate-lowrate) / (1-highrate)) *
thresh1) if highband==1;

replace ann_itn = ann_itn -197 if mca==1;
replace ann_itn = 0 if ann_itn<0;

lab var ann_itn "from NET: imputed annual income tax payments";

```

```

cap drop Iann_gro;
cap drop Iann_net;

gen Iann_gro = ann_net + (wk_nin*52) + ann_itn ;
lab var Iann_gro " gross annual income imputed from net ";

gen Iann_net = ann_gro - (wk_nig*52) - ann_itg ;
lab var Iann_net " net annual income imputed from gross ";

/* need to translate these back into other periods : hourly, weekly */
/* weekly */
cap drop Iwk_gro;
cap drop Iwk_net;

gen Iwk_gro= Iann_gro/52;
lab var Iwk_gro "weekly gross imputed from net";

gen Iwk_net= Iann_net/52;
lab var Iwk_net "weekly net imputed from gross";

/*monthly */
cap drop Imo_gro;
cap drop Imo_net;

gen Imo_gro= Iann_gro/12;
lab var Imo_gro "monthly gross imputed from net";

gen Imo_net = Iann_net/12;
lab var Imo_net "monthly net imputed from gross";

/* hourly */
cap drop Ihr_gro;
cap drop Ihr_net;

gen Ihr_gro = Iwk_gro/hours;
lab var Ihr_gro "hourly gross imputed from net";

gen Ihr_net = Iwk_net/hours;
lab var Ihr_net "hourly net imputed from gross";

/*****log some results
*****/
cap drop prop;
gen prop = hr_net/hr_gro;
replace prop = hr_net/Ihr_gro if zhr_gro==1;
replace prop = Ihr_net/hr_gro if zhr_net==1;

cap drop imputax*;
gen imputax1 = Ihr_net/hr_gro;
gen imputax2 = hr_net/Ihr_gro;

lab var imputax1 "ratio: imputed net given gross";
lab var imputax2 "ratio: imputed gross given net";
cap drop reported;
gen reported = hr_net/hr_gro;

cap log close;
log using imputation, replace;
su prop;

```

```
su prop if zhr_net~=1 & zhr_gro~=1, det;
su prop if zhr_net==1,det;
su prop if zhr_gro==1,det;

su hr_gro Ihr_gro hr_net Ihr_net;
sum imputax1 imputax2 reported;

cap log close;
```

PARTNER.DO

```
/* this sorts out partner's income
   1. set up missing variable
   2. put into weekly/ annual
   3. any fixes?
   4. impute gross from net
*/
#delimit;
set more 1;

/***** set original in stone
*****/

cap drop opnet;
gen opnet = pnetpay;
lab var opnet "original pnetpay";

cap drop opnetprd;
gen opnetprd = pnetprd;
lab var opnetprd "original pnetprd";

cap drop opnpred;
gen opnpred = pnetpred;
lab var opnpred "original pnetpred";

/**** EARNINGS STATUS :   this is different from the cm coding, as there
                           are earnings figs for people across all peconact,
                           not just employees *****/

cap drop pearns;
gen pearns      =1          if pnetpay~=. ;
replace pearns  =0          if pnetpay==. ;
lab var pearns  "partner's earnings reported";

/***** missing partner's pay
*****/

cap drop zpnetpay;
gen zpnetpay    =1          if opnet== 99999998|opnet==99999999|opnet==1;
replace zpnetpay =0          if zpnetpay~=1;
replace zpnetpay =.          if pnetpay==. ;
lab var zpnetpay "missing pnetpay";

/***** period codes
*****/

/*the coding frame for pnetprd:
   1 one week
   2 a fortnight
   3 four weeks
   4 a calendar month
   5 a year or
   6 other period

the coding frame for pnetpred:
   1 one week
   2 a fortnight/2 weeks
   3 four weeks
   4 a calendar month
   5 a year (12 months/52 weeks)
```

```

        6 three weeks
        7 five weeks
        8 six weeks
        9 seven weeks
       10 eight weeks
       11 two calendar months
       12 eight times a year
       13 nine times a year
       14 ten times a year
       15 three months/13 weeks
       16 six months/26 weeks
       17 hourly
       18 daily
       19 one off lump sum
       20 some other period
       21 varies
       22 refused
       23 some other comment
       24 irrelevant/unspecific response */

/***** NET PAY VARIABLES
*****/

/* First make sure not to start manipulating missing values 9999997 and
9999998 ! */

replace pnetpay =. if zpNetpay==1;

/* need to guess
   days per week (5 give their pay daily)
   hours per week (3 give their pay hourly)

   assume the number of hours (and days based on this)
   is the median employee hours of the cohort members--
   this is 20 for part time males and females,
   40 for full-time females,
   45 for full-time males

   (assume partner's sex is opp of cm sex -- too much effort to get proper
code from household
   grid for now -- out of these 8 its pretty much certain to be true!)
*/

cap drop phours;
gen phours = 20 if (peconact==2|peconact==4);
replace phours = 40 if (peconact==1|peconact==3) & cmsex==1;
replace phours = 45 if (peconact==1|peconact==3) & cmsex==2;
lab var phours "partner's hours: estimated";

cap drop pdays;
gen pdays = 2.5 if (peconact==2|peconact==4);
replace pdays = 5 if (peconact==1|peconact==3);
lab var pdays "partner's days: estimated";

/* this program converts net pay into different periods depending on the
period code */
cap pr drop dopnet;
pr def dopnet;

```



```
cap drop ann_pnet;
cap drop wk_pnet;
```

```
cap drop zpNetprd;
```

```
gen wk_pnet          =          (pnetpay)                if  pnetprd==1
& zpNetpay==0;
replace wk_pnet      =          (pnetpay / 2)            if  pnetprd==2
& zpNetpay==0;
replace wk_pnet      =          (pnetpay / 4)            if  pnetprd==3
& zpNetpay==0;
replace wk_pnet      =          (pnetpay / (13/3))        if  pnetprd==4
& zpNetpay==0;
replace wk_pnet      =          (pnetpay / 52)           if  pnetprd==5
& zpNetpay==0;
replace wk_pnet      =          9999997                 if  pnetprd==. &
pearns==1          & zpNetpay==0;

replace wk_pnet      =          (pnetpay)                if pnetprd==6 &
pnetpred==1        & zpNetpay==0;
replace wk_pnet      =          (pnetpay / 2)            if pnetprd==6 &
pnetpred==2        & zpNetpay==0;
replace wk_pnet      =          (pnetpay / 3)            if pnetprd==6 &
pnetpred==3        & zpNetpay==0;
replace wk_pnet      =          (pnetpay / (13/3))        if pnetprd==6 &
pnetpred==4        & zpNetpay==0;
replace wk_pnet      =          (pnetpay / 52)           if pnetprd==6 &
pnetpred==5        & zpNetpay==0;
replace wk_pnet      =          (pnetpay / 3)            if pnetprd==6 &
pnetpred==6        & zpNetpay==0;
replace wk_pnet      =          (pnetpay / 5)            if pnetprd==6 &
pnetpred==7        & zpNetpay==0;
replace wk_pnet      =          (pnetpay / 6)            if pnetprd==6 &
pnetpred==8        & zpNetpay==0;
replace wk_pnet      =          (pnetpay / 7)            if pnetprd==6 &
pnetpred==9        & zpNetpay==0;
replace wk_pnet      =          (pnetpay / 8)            if pnetprd==6 &
pnetpred==10       & zpNetpay==0;
replace wk_pnet      =          (pnetpay / (26/3))        if pnetprd==6 &
pnetpred==11       & zpNetpay==0;
replace wk_pnet      =          (pnetpay / (52/8))        if pnetprd==6 &
pnetpred==12       & zpNetpay==0;
replace wk_pnet      =          (pnetpay / (52/9))        if pnetprd==6 &
pnetpred==13       & zpNetpay==0;
replace wk_pnet      =          (pnetpay / (52/10))       if pnetprd==6 &
pnetpred==14       & zpNetpay==0;
replace wk_pnet      =          (pnetpay / 13)           if pnetprd==6 &
pnetpred==15       & zpNetpay==0;
replace wk_pnet      =          (pnetpay / 26)           if pnetprd==6 &
pnetpred==16       & zpNetpay==0;
replace wk_pnet      =          (pnetpay * phours)        if pnetprd==6 &
pnetpred==17       & zpNetpay==0;
replace wk_pnet      =          (pnetpay * pdays)        if pnetprd==6 &
pnetpred==18       & zpNetpay==0;
replace wk_pnet      =          9999997                 if pnetprd==6 &
pnetpred>=19 & pnetpred~=. & zpNetpay==0;
replace wk_pnet      =          9999997                 if pnetprd==. &
pearns==1;
```

```

lab var wk_pnet    "partner weekly net pay";

/* missing net period */
gen zpnetprd      =          1 if wk_pnet==9999997;
replace zpnetprd  =          0 if pearn==1 & zpnetprd~1;
replace zpnetprd  =          . if pearn==0;
lab var zpnetprd  "missing net period code";

replace wk_pnet   =          . if zpnetprd==1;

/* annual net pay */
gen ann_pnet      =          (wk_pnet*52)          if zpnetprd==0;
replace ann_pnet.=          if zpnetprd==1;
lab var ann_pnet  "partner annual net pay";

/* ROUGH guess at hourly net pay -- by dividing by estimated hours */
cap drop hr_pnet;
gen hr_pnet = wk_pnet/phours;
replace hr_pnet =. if zpnetprd==1;

end;
dopnet;

cap pr drop podd;
pr def podd;
cap drop podd;
gen podd=1 if (hr_pnet<2 | hr_pnet>100) & ann_pnet~=. ;
lab var podd "odd partner's wage";
tab podd;
end;
podd;

/***** FIXES TO PARTNER'S PAY *****/
/* FIX ONE
   A LOAD CODED AS WEEKLY WHICH I AM CHANGING TO ANNUAL ---- SOME COULD
   HAVE THE DECIMAL POINT IN THE WRONG PLACE,
   ITS IMPOSSIBLE TO KNOW WHICH, BUT IT MATTERS SINCE WHAT I AM DOING
   DIVIDES BY 52, WHEREAS THE OTHER DIVIDES BY 100!!!! */

   replace pnetprd=5 if (serial ==111207 | serial == 205291 | serial ==
100875 | serial == 127091 |
                           serial ==107752 | serial == 108363 | serial ==
211375);

/* FIX TWO
   A LOAD CODED AS ANNUAL WHICH MUST BE MONTHLY */

   replace pnetprd = 4 if hr_pnet<1 & pnetpay>500 & pnetprd==5;

/* FIX 3: these need to be multiplied by 1000 I think ! */
replace pnetpay=opnet*1000 if serial==231141|serial==220423|serial==218974;

/* A LOAD CODED AS ANNUAL WHICH I THINK ARE WEEKLY - ALTHOUGH THEY COULD BE
HOURLY AND WITHOUT A DECIMAL POINT...! */

```

```

replace pnetprd=1 if
serial==223385|serial==208807|serial==128051|serial==230870|serial==208819|
serial==223287;

/* FIX 4: these are coded as monthly which should be annual */
replace pnetprd = 5 if (serial ==
129075|serial==226193|serial==200516|serial==208649);
su pnetprd;

/* FIX 5 : decimal point has gone missing: divide by 100 */
replace pnetpay = opnet /100 if serial==109177|serial==228795;

/* FIX 6: some individual fixes */
replace pnetpay =1100 if serial==129849;

/* Fix 7: these ought to be monthly ! */
replace pnetprd = 4 if  serial==230227|serial==111346|

    serial==228451|serial==206491|serial==108097|serial==220322|

serial==110103|serial==227417|serial==222487|serial==130222|
serial==218190|serial==112998|serial==209574|serial==105600|

serial==222518|serial==123032|serial==229425|serial==120296;

/* these need to be *100 */
replace pnetpay = opnet *100 if serial==
216308|serial==201426|serial==111346|serial==222511;

/* these need to be * 10 */
replace pnetpay= opnet * 10 if
serial==113089|serial==214352|serial==211187|serial==104141|

serial==209738|serial==219897|serial==228918|serial==111609|

serial==222537|serial==104729|serial==125905|serial==111127|

serial==217108|serial==111811|serial==212263;
dopnet;
replace pnetpay = pnetpay*10 if hr_pnet>=0.5 & hr_pnet<=1;

/* these look like hourly */
replace pnetprd=6 if
serial==225242|serial==105684|serial==200634|serial==112685;
replace pnetpred=17 if
serial==225242|serial==105684|serial==200634|serial==112685;

/* change to missing - can't really tell what they should be...*/
replace pnetpay=. if
serial==113463|serial==222918|serial==219680|serial==117299|serial==130637|

serial==202656|serial==219418|serial==124541|serial==127480|serial==214513;

/* these look like weekly */
replace pnetprd=1 if pnetprd==4 & hr_pnet>=1.02 & hr_pnet<2;

dopnet;
podd;

```

```

/* lots more needed, but give up for now.... /
JUST IMPUTE A GROSS FIGURE FROM THIS NET -- although the net is given net of
pension contributions, so gross will not be accurate! */
/***** NATIONAL INSURANCE
*****/

/* FROM NET */
cap drop wk_nip;

gen wk_nip = 0 if wk_pnet<=le1
;
replace wk_nip = ((wk_pnet-le1)/9) if wk_pnet>le1 &
wk_pnet<= uel*0.9 ;
replace wk_nip = (0.1* (uel- le1)) if
wk_pnet>uel*0.9 ;
lab var wk_nip "from net: partner's imputed weekly NI payments";

/***** INCOME TAX
*****/
/* tax is worked out separately for those whose income falls:
1. below personal allowance
2. in 10p band
3. in basic rate band
4. in higher rate band */

cap drop pnotax;
cap drop plowb;
cap drop pbaseb;
cap drop phighb;

/* put partners in income tax bands according to their net incomes */
gen pnotax =1 if (ann_pnet<=pa);
gen plowb =1 if (ann_pnet>pa) & (ann_pnet <= (pa + (0.9 * thresh1)));
gen pbaseb=1 if (ann_pnet > (pa + (0.9 * thresh1))) & (ann_pnet <= (pa +
(0.9 * thresh1) + ((1-baserate) * (thresh2-thresh1)));
gen phighb=1 if (ann_pnet >=(pa + (0.9 * thresh1) + ((1-baserate) *
(thresh2-thresh1)));

/* FROM NET */

cap drop ann_itp;
gen ann_itp =0 if
notax==1;
replace ann_itp = (lowrate/(1-lowrate)) *(ann_pnet-pa) if
lowband==1;
replace ann_itp = ((baserate/(1-baserate))*(ann_pnet-pa)) - ((baserate-
lowrate) / (1-baserate)) * thresh1 if baseband==1;
replace ann_itp = ((highrate/(1-highrate))*(ann_pnet-pa)) - ((highrate-
baserate)/(1-highrate)) * thresh2 - ((baserate-lowrate)/(1-highrate)) *
thresh1 if highband==1;

```

```

/* don't know if its right to give the whole MCA to the partner -- when
imputing before I gave it all to the cohort memeber
   think about doing this properly later. */
replace ann_itp = ann_itp -197 if mca==1;
replace ann_itp = 0           if ann_itp<0;

lab var ann_itp "from NET: partner's imputed annual income tax payments";

cap drop ann_pgro;

gen ann_pgro = ann_pnet + (wk_nip*52) + ann_itp ;
lab var ann_pgro "partner's gross annual income imputed from net ";

/* need to translate these back into other periods : hourly, weekly */
/* weekly */
cap drop wk_pgro;
gen wk_pgro= ann_pgro/52;
lab var wk_pgro "partner's weekly gross imputed from net";

/* hourly */
cap drop hr_pgro;

gen hr_pgro = wk_pgro/24;
lab var hr_pgro "partner's hourly gross imputed from net";

/***** put originals back *****/
#delimit;
cap drop Ipnetpay;
cap drop Ipnetprd;
cap drop Ipnpred;

rename pnetpay Ipnetpay;
rename pnetprd Ipnetprd;
rename pnetpred Ipnpred;

rename opnet pnetpay;
rename opnetprd pnetprd;
rename opnpred pnetpred;

label values pnetprd pnetprd;
label values pnetpred pnetpred;

```

APPENDIX 5: Basic Skills and other variables

Documentation Samantha Parsons of the Centre for Longitudinal Studies relating to basic skills and other variables is reproduced below.

Queries about the contents of this appendix should be directed to the User Support Group
(cohort@cls.ioe.ac.uk)

PRELIMINARY ANALYSIS OF NCDS/BCS70 NEW 1999/2000 SURVEY DATA

- i. **Age CM and CM partner left full-time education**
- ii. **Maths difficulties**
- iii. **Reading difficulties**
- iv. **Writing difficulties**
- v. **Other difficulties of having 3R problem**
- vi. **Skills questions**
- vii. **Course attendance**
- viii. **Computer use at home and work**
- ix. **Illegal drug use**
- x. **Diet, food and exercise**
- xi. **Eating problems**
- xii. **Smoking**
- xiii. **Drinking**
- xiv. **Accidents**

AGE LEFT FULL-TIME EDUCATION QUESTIONS

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
actagel2	607	2.7			^1	0	22072
<i>actagel2 details age CM left full-time education</i>							
agelfte2	66	0.3	6	0	36	0.2	22572
<i>agelfte2 captures CMs who were in full-time education at time of interview</i>							
furthed2	607	2.7					22073
<i>furthed2 details CMs who returned to full-time education less than 3 years after initially leaving full-time education</i>							
lftmore2	20522	90.5			^2		2156
<i>lftmore2 details age CM left second period of full-time education</i>							
plefted	5981	26.4	^12	0.1	^347	2.1	16340
<i>plefted details age CM partner left full-time education</i>							

Derived AGE LEFT FULL-TIME EDUCATION Variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
Ageftedu	109	0.5					22571
<i>6 category variables combining information from actagel2 and agelfte2. Values: 1'before 16' 2'at 16' 3'post 16' 4'post 18' 5'post 21' 6'still in ft edu'.</i>							
Agelfted	109	0.5					22571
<i>continuous variable combining information from actagel2, agelfte2, furthed2 and lftmore2. Value range: 14 to 42, 99=still in education</i>							
Ageledu	109	0.5					22571
<i>6 category variable collapsing information in agelfted. Values: 1'before 16' 2'at 16' 3'post 16' 4'post 18' 5'post 21' 6'still in ft edu'</i>							
plefted1	6340	28.0					16340
<i>5 category variable collapsing information in plefted. Values: 1'before 16' 2'at 16' 3'post 16' 4'post 18' 5'post 21'</i>							

SPSS SYNTAX

age CM left full-time education

frequencies

variables=actagel2 agelfte2 furthed2 lftmore2.

missing values actagel2 lftmore2 (99).

missing values agelfte2 (8,9).

compute ageftedu = -1.

if (actagel2 < 16) ageftedu = 1.

if (actagel2 = 16) ageftedu = 2.

if (actagel2 = 17 | actagel2 = 18) ageftedu = 3.

if (actagel2 = 19 | actagel2 = 20 | actagel2 = 21) ageftedu = 4.

if (actagel2 > 21 and agelfte2 = 1) ageftedu = 5.

if (missing(actagel2) and agelfte2 = 2) ageftedu = 6.

missing values ageftedu (-1).

variable labels ageftedu 'age CM left ft continuous education'.

value labels ageftedu 1'before 16' 2'at 16' 3'post 16' 4'post 18' 5'post 21' 6'still in ft edu'.

frequencies

variables=ageftedu.

age CM left full time education accounting for a 'gap' of no more than 3 years

**two cases said had an additional period of education but did not say when finished, although they were not still in ft education - for these two cases I used the age of leaving ft education in ACTAGEL2 variable.

compute agefted = actagel2.

if (furthed2 = 1) agefted = lftmore2.

if (agelfte2 = 2) agefted = 99.

if (serial = 109993 | serial = 217003) agefted = actagel2.

variable labels agefted 'age CM finally left ft education'.

value labels agefted 99'still in ed'.

frequencies

variables=agefted.

compute ageledu = -1.

if (agefted < 16) ageledu = 1.

if (agefted = 16) ageledu = 2.

if (agefted = 17 | agefted = 18) ageledu = 3.

if (agefted = 19 | agefted = 20 | agefted = 21) ageledu = 4.

if (agefted > 21) ageledu = 5.

if (agefted = 99) ageledu = 6.

missing values ageledu (-1).

variable labels ageledu 'age CM finally left ft education'.

value labels ageledu 1'before 16' 2'at 16' 3'post 16' 4'post 18' 5'post 21' 6'still in ft edu'.

frequencies

variables=ageftedu ageledu.

age CM partner left full-time education

frequencies


```
variables= plefted.  
missing values plefted (98,99).  
recode plefted (14,15 = 1) (16 = 2) (17,18 = 3) (19 thru 21 = 4) (22 thru 46 = 5) (sysmis, 98, 99 = -1)  
into plefted1.  
missing values plefted1 (-1).  
variable labels plefted1 'age CM partner finally left education'.  
value labels plefted1 1'before 16' 2'at 16' 3'post 16' 4'post 18' 5'post 21'.  
frequencies  
variables= plefted1.
```

MATHS QUESTIONS

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
mathsprb	66	0.3	4	0	21	0.1	22589
<i>549 (2.4%) had difficulty/were not able to work out change from £10. However, only 426 (78% of 549) asked additional 5 questions on type of maths difficulties CM had:</i>							
mprbtype							426
addup							426
subtract							426
multiply					1	0	425
subtract					1	0	425
mathcors	66	0.3	4	0	20	0.1	22590
<i>429 (1.9%) had done a course to improve their maths. 429 asked if doing course now or previously</i>							
mathsnow	22251	98.1					429
mathimp	139	0.6			3	0	22538
<i>6234 (27.7%) ever wanted to improve their maths. 6234 asked multi-coded questions i) why they wanted to improve their maths and ii) where/how they would improve them</i>							
mthlike1 –	16446	72.5			2		6232
mthlike5							
mthplac1-	16446	72.5			^4		6230
mthplac9							
datesprb	1377	6.1	1	0	15	0.1	21271
<i>696 (3.3%) had difficulty/were not able to work out dates/use a calendar</i>							
mathconf	11421	50.4	3	0	20	0.1	11236
<i>1536 (13.7%) not confident helping their child(ren) with maths</i>							
mathskid	11421	50.4	3	0	16	0.1	11240
<i>1093 (9.7%) did not help their child(ren) with maths</i>							
partmath	12733	56.1	4	0	22	0.1	9921
<i>1439 (14.5%) of CM partner did not help their child(ren) with maths</i>							

Derived MATHS variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
bsamd	1394	6.1					21284
<p>4 category variable combining information given in mathsprb and datesprb. Only for CMs who answered both questions. CM reporting 'difficulties' or 'no, can't do' treated as one category. Values: 1'no difficulties' 2'maths difficulties' 3'date/calendar difficulties' 4'maths and 3'date/calendar difficulties'.</p>							
bsamp1	22254	98.1					426
<p>6 category variable counting the specific maths problems of CMs who had reported maths difficulties in mathsprb. Included mprbtype (recognising numbers), addup, subtract, multiply, and divide. Values: 0'not these problems' 1'one problem' 2'two problems' 3'three problems' 4'four problems' 5'all five problems'.</p>							
bsamp2	22254	98.1					426
<p>4 category variable counting the specific maths problems of CMs who had reported maths difficulties in mathsprb. Included addup, subtract, multiply, and divide. Values: 0'not these problems' 1'one problem' 2'two problems' 3'three problems' 4'all four problems'.</p>							
<p>Derived variables from multi-coded mthlike1-mthlike5: reasons why CM wants to improve their maths: for their children, to get a job, to get a promotion, get a better job, for self satisfaction.</p>							
impmkid	16448	72.5					6232
impmgjob							
impmprom							
impmbjob							
impmself							
<p>Derived variables from multi-coded mthplac1-mthplac9: means by which CM would improve their maths: on a day college course, on an evening college course, on a weekend college course, on ANY college course, at a community centre, by using local library resources, by using PC packages, by using programmes on TV, by using programmes on the radio, by using books at home.</p>							
domcolld	16450	72.5					6230
domcolle							
domcollw							
domcoll							
domcomc							
domllib							
dompc							
domtv							
domradio							
dombook							

SPSS SYNTAX

missing values mathsprb mprbtype mathcors mthimp addup subtract multiply divide datesprb
mathconf mathskid mathsnow partmath (8,9).

Maths and Date problems

```
compute bsamd = -1.  
if (mathsprb = 1 and datesprb = 1) bsamd = 0.  
if (mathsprb >= 2 and datesprb = 1) bsamd = 1.  
if (mathsprb = 1 and datesprb >= 2) bsamd = 2.  
if (mathsprb >= 2 and datesprb >= 2) bsamd = 3.  
missing values bsamd (-1).
```

```
variable labels bsamd '2000: CM has maths and calendar problems'.  
value labels bsamd 0'no probs' 1'maths diffs' 2'date diffs' 3'maths and date diffs'.
```

```
frequencies  
  variables=bsamd.
```

Number of math problems.

```
do if (mprbtype >= 1).  
count bsamp1 = mprbtype addup subtract multiply divide (1).  
end if.  
variable labels bsamp1 '2000: CM has probs with math calculations'.  
value labels bsamp1 0'no problems' 1'one problem' 2'two problems' 3'three problems'  
4'four problems' 5'all five problems'.
```

```
do if (mprbtype >= 1).  
count bsamp2 = addup subtract multiply divide (1).  
end if.  
variable labels bsamp2 '2000: CM has probs with 4 types math calculations'.  
value labels bsamp2 0'no problems' 1'one problem' 2'two problems' 3'three problems'  
4'all four problems'.
```

```
frequencies  
  variables= bsamp1 bsamp2.
```

Reasons why want to improve math skills

missing values mthlike1 mthlike2 mthlike3 mthlike4 mthlike5 (8,9).

```
do if (mthlike1 > 0).  
compute impmkid = 0.  
compute impmgjob = 0.  
compute impmprom = 0.  
compute impmbjob = 0.  
compute impmself = 0.  
do repeat x = mthlike1 mthlike2 mthlike3 mthlike4 mthlike5 .  
if (x = 1) impmkid = 1.  
if (x = 2) impmgjob = 1.  
if (x = 3) impmprom = 1.  
if (x = 4) impmbjob = 1.  
if (x = 5) impmself = 1.  
end repeat.  
end if.  
variable labels impmkid 'CM wants to improve maths to help their children'.  
variable labels impmgjob 'CM wants to improve maths to help get a job'.
```

variable labels impmprom 'CM wants to improve maths to help get a promotion at work'.
variable labels impmbjob 'CM wants to improve maths to help get a better job'.
variable labels impmself 'CM wants to improve maths to study for own satisfaction'.

value labels impmkid impmgjob impmprom impmbjob impmself 0'not this reason' 1'yes, for this reason'.

frequencies
variables=impmkid impmgjob impmprom impmbjob impmself .

Place where CM would go to improve math skills

missing values mthplac1 (99).

```
do if (mthplac1 > 0).  
  compute domcolld = 0.  
  compute domcolle = 0.  
  compute domcollw = 0.  
  compute domcoll = 0.  
  compute domcomc = 0.  
  compute domllib = 0.  
  compute dompc = 0.  
  compute domtv = 0.  
  compute domradio = 0.  
  compute dombook = 0.  
do repeat x = mthplac1 mthplac2 mthplac3 mthplac4 mthplac5 mthplac6 mthplac7 mthplac8  
mthplac9.  
if (x = 1) domcolld = 1.  
if (x = 2) domcolle = 1.  
if (x = 3) domcollw = 1.  
if (x <= 3) domcoll = 1.  
if (x = 4) domcomc = 1.  
if (x = 5) domllib = 1.  
if (x = 6) dompc = 1.  
if (x = 7) domtv = 1.  
if (x = 8) domradio = 1.  
if (x = 9) dombook = 1.  
end repeat.  
end if.
```

variable labels domcolld 'CM would improve their maths on a DAY college course'.
variable labels domcolle 'CM would improve their maths on an EVENING college course'.
variable labels domcollw 'CM would improve their maths on a WEEKEND college course'.
variable labels domcoll 'CM would improve their maths on ANY college course'.
variable labels domcomc 'CM would improve their maths at a community centre'.
variable labels domllib 'CM would improve their maths by using local library resources'.
variable labels dompc 'CM would improve their maths by using PC packages'.
variable labels domtv 'CM would improve their maths by using programmes on TV'.
variable labels domradio 'CM would improve their maths by using programmes on the RADIO'.
variable labels dombook 'CM would improve their maths by using books at home'.

value labels domcolld domcolle domcollw domcoll domcomc domllib dompc domtv domradio
dombook 0'not using this method' 1'this method'.

frequencies
variables=domcolld domcolle domcollw domcoll domcomc domllib dompc domtv domradio
dombook.

(see COURSE section for derived reading course variable)

READING QUESTIONS

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
readprb1 <i>679 (3.2%) had difficulty/were not able read/understand magazine/newspaper text. 680 asked additional 2 questions on other reading difficulties (includes CM with value of 9 in readprb1)</i>	1377	6.1			1	0	21302
readprb2	22000	97.0					680
readprb3	22003	97.0			3	0	677
readprb4 <i>Asks if CM feels ability to read paperwork has improved over the last decade. 789 (3.7%) feel it has got worse.</i>	1377	6.1			3	0	21300
readcors <i>197 (0.9%) had done a course to improve their reading. 197 asked if doing course now or previously</i>	66	0.3	4	0	20	0.1	22590
readnow	22483	99.1					197
readimp <i>1856 (8.2%) ever wanted to improve their reading. 1856 asked multi-coded questions i) why they wanted to improve their reading and ii) where/how they would improve them</i>	128	0.6			2	0	22550
redlike1 – redlike5	20824	91.8					1856
redplac1- redplac9 <i>(redplac8 and redplac9 hold no information)</i>	20824	91.8			^2		1854
kidrdcnf <i>561 (5.0%) not confident helping their child(ren) with reading</i>	11421	50.4	3	0	21	0.1	11235
readkid <i>822 (7.3%) did not read to their child(ren)</i>	11421	50.4	3	0	15	0.1	11241
partread <i>1271 (12.8%) of CM partner did not read to their child(ren)</i>	12733	56.1	4	0	21	0.1	9922

Derived READING variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
bsaread1	1378	6.1					21302
<i>4 category variable combining information given in readprb1, readprb2 and readprb3. CM reporting 'difficulties' or 'no, can't do' treated as one category. Values: 1'no reading difficulties' 2'1 type of reading difficulty' 3'2 types of reading difficulties' 4'3 types of reading difficulties'.</i>							
bsaread2	22003	97.0					677
<i>3 category variable combining information given in readprb1, readprb2 and readprb3 for CMs who reported 'difficulties' or 'no, can't do' in readprb1 and answered readprb1, readprb2 and readprb3. CM reporting 'difficulties' or 'no, can't do' treated as one category. Values: 1'no reading difficulties' 2'1 type of reading difficulty' 3'2 types of reading difficulties' 4'3 types of reading difficulties'.</i>							
Derived variables from multi-coded redlike1-redlike5: reasons why CM wants to improve their reading: for their children, to get a job, to get a promotion, get a better job, for self satisfaction.							
imprkid	20824	91.8					1856
imprgjob							
imprprom							
imprbjob							
imprself							
Derived variables from multi-coded redplac1-redplac9: means by which CM would improve their reading: on a day college course, on an evening college course, on a weekend college course, on ANY college course, at a community centre, by using local library resources, by using PC packages, by using programmes on TV, by using programmes on the radio, by using books at home.							
dorcollid	20826	91.8					1854
dorcolle							
dorcollw							
dorcoll							
dorcomc dorllib							
dorpc							
dortv							
dorradio							
dorbook							

SPSS SYNTAX

missing values readprb1 readprb2 readprb3 readprb4 readcors readimp partread readkid readnow kidrdcnf (8,9).

Number of reading problems CM has

```
do if (readprb1 > 0).
count bsaread1 = readprb1 readprb2 readprb3 (2,3) /
end if.
```

```
do if (readprb1 > 1).
count bsaread2 = readprb1 readprb2 readprb3 (2,3) /
end if.
```

```
variable labels bsaread1 'CM not able to do or has difficulties with reading tasks'.
variable labels bsaread2 'CM not able to do or has difficulties with reading tasks'.
```

```
frequencies
  variables=bsaread1 bsaread2 .
```

Reasons CM wants to improve reading skills

missing values redlike1 redlike2 redlike3 redlike4 redlike5 (8,9).

```
do if (redlike1 > 0).
compute imprkid = 0.
compute imprgjob = 0.
compute imprprom = 0.
compute imprbjob = 0.
compute imprself = 0.
do repeat x = redlike1 redlike2 redlike3 redlike4 redlike5 .
if (x = 1) imprkid = 1.
if (x = 2) imprgjob = 1.
if (x = 3) imprprom = 1.
if (x = 4) imprbjob = 1.
if (x = 5) imprself = 1.
end repeat.
end if.
```

```
variable labels imprkid 'CM wants to improve reading to help their children'.
variable labels imprgjob 'CM wants to improve reading to help get a job'.
variable labels imprprom 'CM wants to improve reading to help get a promotion at work'.
variable labels imprbjob 'CM wants to improve reading to help get a better job'.
variable labels imprself 'CM wants to improve reading to study for own satisfaction'.
```

```
value labels imprkid imprgjob imprprom imprbjob imprself 0'not this reason' 1'yes, for this reason'.
```

```
frequencies
  variables=imprkid imprgjob imprprom imprbjob imprself .
```


Place where CM would go to improve reading skills

missing values redplac1 (99).

```
do if (redplac1 > 0).
compute dorcolld = 0.
compute dorcolle = 0.
compute dorcollw = 0.
compute dorcoll = 0.
compute dorcomc = 0.
compute dorllib = 0.
compute dorpc = 0.
compute dortv = 0.
compute dorradio = 0.
compute dorbook = 0.
do repeat x = redplac1 redplac2 redplac3 redplac4 redplac5 redplac6 redplac7 redplac8 redplac9 .
if (x = 1) dorcolld = 1.
if (x = 2) dorcolle = 1.
if (x = 3) dorcollw = 1.
if (x <= 3) dorcoll = 1.
if (x = 4) dorcomc = 1.
if (x = 5) dorllib = 1.
if (x = 6) dorpc = 1.
if (x = 7) dortv = 1.
if (x = 8) dorradio = 1.
if (x = 9) dorbook = 1.
end repeat.
end if.
```

variable labels dorcolld 'CM would improve their reading on a DAY college course'.

variable labels dorcolle 'CM would improve their reading on an EVENING college course'.

variable labels dorcollw 'CM would improve their reading on a WEEKEND college course'.

variable labels dorcoll 'CM would improve their reading on ANY college course'.

variable labels dorcomc 'CM would improve their reading at a community centre'.

variable labels dorllib 'CM would improve their reading by using local library resources'.

variable labels dorpc 'CM would improve their reading by using PC packages'.

variable labels dortv 'CM would improve their reading by using programmes on TV'.

variable labels dorradio 'CM would improve their reading by using programmes on the RADIO'.

variable labels dorbook 'CM would improve their reading by using books at home'.

value labels dorcolld dorcolle dorcollw dorcoll dorcomc dorllib dorpc dortv dorradio dorbook 0'not using this method' 1'this method'.

frequencies

variables=dorcolld dorcolle dorcollw dorcoll dorcomc dorllib dorpc dortv dorradio dorbook.

(see COURSE section for derived reading course variable)

WRITING QUESTIONS

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
writeprb <i>1087 (3.2%) had difficulty/were not able to write a thank-you letter to a friend. 1087 asked additional question on spelling difficulties (includes CM with value of 9 in writeprb)</i>	1377	6.1			1	0	21302
wprbtype <i>70 (6.4%) never tried to write a thank-you letter to a friend. 1017 asked additional questions on handwriting legibility and difficulties putting down in words what want to say</i>	21593	95.2			1	0.1	1086
hwritprb	21663	95.5			1	0.1	1016
wordsprb	21663	95.5					1017
writcors <i>315 (1.4%) had done a course to improve their writing. 315 asked if doing course now or previously</i>	66	0.3	4	0	21	0.1	22589
writenow	22365	98.6					315
writimp <i>2636 (11.7%) ever wanted to improve their writing. 2636 asked multi-coded questions i) why they wanted to improve their reading and ii) where/how they would improve them</i>	138	0.6			3	0	22539
wrilike1 – wrilike5	20044	88.4			1	0	2635
wriplac1- wriplac9	20044	88.4			^5		2631
writconf <i>806 (7.2%) not confident helping their child(ren) with writing</i>	11421	50.4	4	0	26	0.1	11229
writekid <i>1440 (12.8%) did not help their child(ren) with writing</i>	11421	50.4	3	0	14	0.1	11242
partwrit <i>1687 (17.0%) of CM partner did help their child(ren) with writing</i>	12733	56.1	4	0	21	0.1	9922

Derived WRITING variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
bsawrit1	1379	6.1					21301
<p><i>8 category variable combining information given in writeprb, hwritprb, wprbtype and wordsprb. CM reporting 'difficulties' or 'no, can't do' treated as one category. Values: 0'no probs' 1'writing spelling handwriting articulating' 2'writing spelling handwriting' 3'writing spelling articulating' 4'writing handwriting articulating' 5'writing spelling' 6'writing articulating' 7'writing handwriting' 8'writing only'.</i></p>							
bsawrit2	21594	95.2					1086
<p><i>4 category variable counting the specific writing problems of CMs who had reported 'difficulties' or 'no, can't do' in writeprb. Included writeprb, wprbtype (spelling), hwritprb (handwriting), and wordsprb (articulating). CM reporting 'difficulties' or 'no, can't do' in writeprb treated as one category. Values: 1'1 difficulty' 2'2 types of writing difficulties' 3'3 types of writing difficulties' 4'4 types of writing difficulties'.</i></p>							
<p>Derived variables from multi-coded redlike1-redlike5: reasons why CM wants to improve their reading: for their children, to get a job, to get a promotion, get a better job, for self satisfaction. Values: 0'not this reason' 1'yes, for this reason'.</p>							
impwkid	20045	88.4					2635
impwgjob							
impwprom							
impwbjob							
impwself							
<p>Derived variables from multi-coded redplac1-redplac9: means by which CM would improve their reading: on a day college course, on an evening college course, on a weekend college course, on ANY college course, at a community centre, by using local library resources, by using PC packages, by using programmes on TV, by using programmes on the radio, by using books at home. Values: 0'not using this method' 1'this method'.</p>							
dowcollid	20049	88.4					2631
dowcolle							
dowcollw							
dowcoll							
dowcomc							
dowllib							
dowpc							
dowtv							
dowradio							
dowbook							

SPSS SYNTAX

Type of writing difficulties CM has

missing values writeprb hwritprb wprbtype wordsprb writimp writcors writconf writekid writenow partwrit (8,9).

```
compute bsawrit1 = -1.  
if (writeprb = 1) bsawrit1 = 0.  
if (writeprb >= 2 and (wprbtype = 1 | wprbtype = 3) and (missing(hwritprb) | hwritprb = 1) and  
(missing(wordsprb) | wordsprb = 1)) bsawrit1 = 1.  
if (writeprb >= 2 and (wprbtype = 1 | wprbtype = 3) and (missing(hwritprb) | hwritprb = 1) and wordsprb  
= 2) bsawrit1 = 2.  
if (writeprb >= 2 and (wprbtype = 1 | wprbtype = 3) and hwritprb = 2 and (missing(wordsprb) |  
wordsprb = 1)) bsawrit1 = 3.  
if (writeprb >= 2 and wprbtype = 2 and (missing(hwritprb) | hwritprb = 1) and (missing(wordsprb) |  
wordsprb = 1)) bsawrit1 = 4.  
if (writeprb >= 2 and (wprbtype = 1 | wprbtype = 3) and hwritprb = 2 and wordsprb = 2) bsawrit1 = 5.  
if (writeprb >= 2 and wprbtype = 2 and hwritprb = 2 and (missing(wordsprb) | wordsprb = 1)) bsawrit1  
= 6.  
if (writeprb >= 2 and wprbtype = 2 and (missing(hwritprb) | hwritprb = 1) and wordsprb = 2) bsawrit1 =  
7.  
if (writeprb >= 2 and wprbtype = 2 and hwritprb = 2 and wordsprb = 2) bsawrit1 = 8.  
missing values bsawrit1 (-1).
```

variable labels bsawrit1 'type of writing problems CM has'.
value labels bsawrit1 0'no probs' 1'writing spelling handwriting articulating' 2'writing spelling
handwriting' 3'writing spelling articulating'
4'writing handwriting articulating' 5'writing spelling' 6'writing articulating' 7'writing handwriting' 8'writing
only'.

Number of writing difficulties CM has

```
recode writeprb (1=0) (2,3=1) into write.  
recode wprbtype (2=0) (1,3=1) into wprb.
```

```
do if (write = 1).  
count bsawrit2 = write wprb hwritprb wordsprb (1).  
end if.
```

variable labels bsawrit2 'Number of writing difficulties CM has'.
value labels bsawrit2 0'no probs' 1'one difficulty' 2'two difficulties' 3'three difficulties' 4'all four
difficulties'.

```
frequencies  
variables=bsawrit1 bsawrit2.
```

Reasons why CM wants to improve writing skills

frequencies

```
variables=writimp wrilike1 wrilike2 wrilike3 wrilike4 wrilike5 wriplac1 wriplac2 wriplac3  
wriplac4 wriplac5 wriplac6 wriplac7 wriplac8 wriplac9 .
```

missing values wrilike1 wrilike2 wrilike3 wrilike4 wrilike5 (8,9).

```
do if (wrilike1 > 0).  
compute impwkid = 0.  
compute impwgjob = 0.  
compute impwprom = 0.  
compute impwbjob = 0.  
compute impwself = 0.  
do repeat x = wrilike1 wrilike2 wrilike3 wrilike4 wrilike5 .  
if (x = 1) impwkid = 1.  
if (x = 2) impwgjob = 1.  
if (x = 3) impwprom = 1.  
if (x = 4) impwbjob = 1.  
if (x = 5) impwself = 1.  
end repeat.  
end if.
```

variable labels impwkid 'CM wants to improve writing to help their children'.

variable labels impwgjob 'CM wants to improve writing to help get a job'.

variable labels impwprom 'CM wants to improve writing to help get a promotion at work'.

variable labels impwbjob 'CM wants to improve writing to help get a better job'.

variable labels impwself 'CM wants to improve writing to study for own satisfaction'.

value labels impwkid impwgjob impwprom impwbjob impwself 0'not this reason' 1'yes, for this reason'.

frequencies

```
variables=impwkid impwgjob impwprom impwbjob impwself .
```

Place where CM would go to improve writing skills

missing values wriplac1 (99).

```
do if (wriplac1 > 0).  
compute dowcolld = 0.  
compute dowcolle = 0.  
compute dowcollw = 0.  
compute dowcoll = 0.  
compute dowcomc = 0.  
compute dowllib = 0.  
compute dowpc = 0.  
compute dowtv = 0.  
compute dowradio = 0.  
compute dowbook = 0.  
do repeat x = wriplac1 wriplac2 wriplac3 wriplac4 wriplac5 wriplac6 wriplac7 wriplac8 wriplac9.  
if (x = 1) dowcolld = 1.  
if (x = 2) dowcolle = 1.  
if (x = 3) dowcollw = 1.  
if (x <= 3) dowcoll = 1.  
if (x = 4) dowcomc = 1.  
if (x = 5) dowllib = 1.  
if (x = 6) dowpc = 1.  
if (x = 7) dowtv = 1.  
if (x = 8) dowradio = 1.
```

```
if (x = 9) dowbook = 1.  
end repeat.  
end if.
```

```
variable labels dowcolld 'CM would improve their writing on a DAY college course'.  
variable labels dowcolle 'CM would improve their writing on an EVENING college course'.  
variable labels dowcollw 'CM would improve their writing on a WEEKEND college course'.  
variable labels dowcoll 'CM would improve their writing on ANY college course'.  
variable labels dowcomc 'CM would improve their writing at a community centre'.  
variable labels dowllib 'CM would improve their writing by using local library resources'.  
variable labels dowpc 'CM would improve their writing by using PC packages'.  
variable labels dowtv 'CM would improve their writing by using programmes on TV'.  
variable labels dowradio 'CM would improve their writing by using programmes on the RADIO'.  
variable labels dowbook 'CM would improve their writing by using books at home'.
```

```
value labels dowcolld dowcolle dowcollw dowcoll dowcomc dowllib dowpc dowtv dowradio dowbook  
0'not using this method' 1'this method'.
```

```
frequencies
```

```
variables=dowcolld dowcolle dowcollw dowcoll dowcomc dowllib dowpc dowtv dowradio dowbook  
writenow.
```

(see COURSE section for derived writing course variable)

DIFFICULTIES OF HAVING BASIC SKILLS PROBLEMS

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
getjob <i>620 (38.5%) CMs with 3R problem felt affected them getting new job</i>	21047	92.8			21	0.1	1612
copejob <i>391 (24.2%) CMs with 3R problem felt affected them coping in current job</i>	21047	92.8			15	0.1	1618
gtpromot <i>570 (35.7%) CMs with 3R problem felt affected them getting a promotion</i>	21047	92.8			38	0.2	1595
copehmbs <i>253 (15.6%) CMs with 3R problem felt affected them in managing household business</i>	21047	92.8			6	0	1627
helpkids <i>419 (25.7%) CMs with 3R problem felt affected them being able to help their children lear. In addition, 409 (25.1%) never helped their children to read or learn</i>	21047	92.8			3	0	1630
copeleis <i>341 (20.9%) CMs with 3R problem felt affected them pursuing other interests</i>	21047	92.8			4	0	1629

SKILLS (self-completed)

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
SKILL1A <i>22344 CM reported 'good' 'okay' or 'poor' communicating skills. 21544 (96%) reported if they used this skill at work.</i>	283	1.2	3		5		
SKILL1B	1136	5.0			1		21543
SKILL2A <i>22268 CM reported 'good' 'okay' or 'poor' number/calculation skills. 21477 (96%) reported if they used this skill at work.</i>	283	1.2	3		4		
SKILL2B	2103	5.3			1		21476
SKILL3A <i>19726 CM reported 'good' 'okay' or 'poor' computer/IT skills. 19173 (86%) reported if they used this skill at work.</i>	283	1.2	3		7		
SKILL3B	3507	15.5			1		19172
SKILL4A <i>22219 CM reported 'good' 'okay' or 'poor' team working skills. 21469 (96%) reported if they used this skill at work.</i>	283	1.2	3		7		
SKILL4B	1211	5.3			3		21466
SKILL5A <i>22297 CM reported if they were 'good' 'okay' or 'poor' at learning new skills. 21517 (96%) reported if they used this skill at work.</i>	283	1.2	3		6		
SKILL5B	1163	5.1			1		21516
SKILL6A <i>22308 CM reported 'good' 'okay' or 'poor' problem solving skills. 21515 (96%) reported if they used this skill at work.</i>	283	1.2	3		7		
SKILL6B	1165	5.1			1		21514
SKILL7A <i>21909 CM reported 'good' 'okay' or 'poor' tool use skills. 21149 (94%) reported if they used this skill at work.</i>	283	1.2	3		6		
SKILL7B	1531	6.8			1		21148
SKILL8A <i>20975 CM reported 'good' 'okay' or 'poor' caring skills. 20217 (90%) reported if they used this skill at work.</i>	283	1.2	3		19	0.1%	
SKILL8B	2463	10.9			1		21216
SKILL9A <i>21228 CM reported 'good' 'okay' or 'poor' finance/accounting skills. 20523 (92%) reported if they used this skill at work.</i>	283	1.2	3		5		
SKILL9B	2157	9.5			1		20522

COURSES

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
failqual	66	0.3	3		34	0.2	22577
<i>3404 (8.8%) failed/cot completed course. 3404 asked how many courses they had failed/not completed</i>							
numfqual	19276	84.9			^3		3401
yts (bcs70 only)	11454	50.5	3		12	0.1	11211
<i>3303 (29.5%) CMs had done a yts course. 3303 asked how many they had done, and if they were on a course now</i>							
numyts	19377	85.4	^1		^1		3301
ytsnow	19377	85.4	1				3302
othgov	66	0.3	4		22	0.1	22588
<i>550 (2.4%) CMs had done an 'other' type of Government course. 550 asked how many they had done, and if they were on a course now</i>							
numgov	22130	97.6			^3	0.5	547
govnow	22130	97.6					550
aptrain	66	0.3	4		24	0.1	22586
<i>329 (1.5 %) CMs had done a modern apprenticeship course. 329 asked how many they had done, and if they were on a course now</i>							
numap	22351	98.5					329
apnow	22351	98.5					329
actrain	66	0.3	4		23	0.1	22587
<i>352 (1.6%) CMs had done an access course. 3303 asked how many they had done, and if they were on a course now</i>							
numac	22328	98.4					352
acnow	22328	98.4					352
wrktrain	66	0.3	4		24	0.1	22586
<i>7871 (34.7%) CMs had done a work-related training course. 7871 asked how many they had done, and if they were on a course now</i>							
numwrktr	14809	65.3			^66	0.8	7805
wrktrnow	14809	65.3			1	0	7870
leiscors	66	0.3	5		23	0.1	22586
<i>5528 (24.4%) CMs had done a course for interest or leisure. 5528 asked how many they had done, and if they were on a course now</i>							
numleis	17152	75.6			^4	0.1	5524
leisnow	17152	75.6					5528

Derived COURSE variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
yts1 <i>3 category variable combining information from yts and ytsnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	11470	50.6					11210
gov1 <i>3 category variable combining information from othgov and govnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	94	0.4					22588
apt1 <i>3 category variable combining information from apttrain and apnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	92	0.4					22586
access1 <i>3 category variable combining information from actrain and acnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	95	0.4					22587
wrt1 <i>3 category variable combining information from wrktrain and wrktrnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	93	0.4					22585
leis1 <i>3 category variable combining information from leiscors and leisnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	94	0.4					22586
read1 <i>3 category variable combining information from readcors and readnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	90	0.4					22590
writ1 <i>3 category variable combining information from writcors and writenow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	91	0.4					22589
math1 <i>3 category variable combining information from mathcors and mathsnow. Values: 0'no' 1'yes, previously' 2'yes, now'.</i>	90	0.4					22590

SPSS SYNTAX

missing values failqual yts ytsnow othgov govnow aptrain apnow actrain acnow wrktrain wrktrnow
leiscors leisnow readcors readnow writcors writenow mathcors mathsnow (8,9).

missing values numfqual numyts numgov numap numac numwrktr numleis numread numwrite
nummaths (98,99).

```
compute yts1 = -1.  
if (yts = 2) yts1 = 0.  
if (ytsnow = 2) yts1 = 1.  
if (ytsnow = 1) yts1 = 2.  
missing values yts1 (-1).  
variable labels yts1 'CM been on YTS?'.  
value labels yts1 0'no' 1'yes, previously' 2'yes, now'.
```

```
compute gov1 = -1.  
if (othgov = 2) gov1 = 0.  
if (govnow = 2) gov1 = 1.  
if (govnow = 1) gov1 = 2.  
missing values gov1 (-1).  
variable labels gov1 'CM been on other government course?'.  
value labels gov1 0'no' 1'yes, previously' 2'yes, now'.
```

```
compute apt1 = -1.  
if (aptrain = 2) apt1 = 0.  
if (apnow = 2) apt1 = 1.  
if (apnow = 1) apt1 = 2.  
missing values apt1 (-1).  
variable labels apt1 'CM been on modern apprenticeship?'.  
value labels apt1 0'no' 1'yes, previously' 2'yes, now'.
```

```
compute access1 = -1.  
if (actrain = 2) access1 = 0.  
if (acnow = 2) access1 = 1.  
if (acnow = 1) access1 = 2.  
missing values access1 (-1).  
variable labels access1 'CM been on access course'.  
value labels access1 0'no' 1'yes, previously' 2'yes, now'.
```

```
compute wrt1 = -1.  
if (wrktrain = 2) wrt1 = 0.  
if (wrktrnow = 2) wrt1 = 1.  
if (wrktrnow = 1) wrt1 = 2.  
missing values wrt1 (-1).  
variable labels wrt1 'CM been on work-related training course'.  
value labels wrt1 0'no' 1'yes, previously' 2'yes, now'.
```

```
compute leis1 = -1.  
if (leiscors = 2) leis1 = 0.  
if (leisnow = 2) leis1 = 1.  
if (leisnow = 1) leis1 = 2.  
missing values leis1 (-1).  
variable labels leis1 'CM been on leisure course'.  
value labels leis1 0'no' 1'yes, previously' 2'yes, now'.
```

```
compute readc1 = -1.  
if (readcors = 2) readc1 = 0.  
if (readnow = 2) readc1 = 1.  
if (readnow = 1) readc1 = 2.  
missing values readc1 (-1).  
variable label readc1 'CM been on reading improvement course'.  
value labels readc1 0'no' 1'yes, previously' 2'yes, now'.
```

```
compute writ1 = -1.  
if (writcors = 2) writ1 = 0.  
if (writenow = 2) writ1 = 1.  
if (writenow = 1) writ1 = 2.  
missing values writ1 (-1).  
variable labels writ1 'CM been on writing improvement course'.  
value labels writ1 0'no' 1'yes, previously' 2'yes, now'.
```

```
compute math1 = -1.  
if (mathcors = 2) math1 = 0.  
if (mathsnow = 2) math1 = 1.  
if (mathsnow = 1) math1 = 2.  
missing values math1 (-1).  
variable labels math1 'CM been on basic maths improvement course'.  
value labels math1 0'no' 1'yes, previously' 2'yes, now'.
```

```
frequencies  
  variables=yts1 apt1 gov1 wrt1 access1 leis1 read1 writ1 math1.
```

COMPUTER USE AT HOME and WORK

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
pchome 13379 (59.0%) CM had a PC at home. 13379 asked how often they used the	66	0.3	4		20	0.1	22590
hpcuse 11289 (84.4%) CMs used the PC at home. 11289 asked about the ways they use a PC	9301	41.0			4	0	13375
howuse01-howuse33 (howuse18-howuse33 hold no information)	11391	50.2			^8	0.1	11281
pcwork 12026 (64.1%) CM use a PC at work. 12026 asked how often use a PC and the ways they use a PC	3915	17.3	2		9	0	18754
wpcuse howuse34-howuse46	10654	47.0	1		^2	0	12025 12024

Derived HOME COMPUTER variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
pchome1 6 category variable combining information from pchome and hpcuse . Values: 0'no' 1'yes, but do not use it' 2'yes, use it less than once a week' 3'yes, use it once a week' 4'yes, use it 2-4 times a week' 5'yes, use it daily'	94	0.4					22586
Derived variables from multi-coded howuse01-howuse33. For 11289 (50.0%) CMs who have and use a home computer. Ways CM uses home computer: for word processing, WWW, e-mail, data analysis, data base work, design packages, playing games, sending/receiving faxes, using CD ROM or Encyclopedia, composing music, listening to music, photography, programming, managing home finances, spreadsheets, webb design, scanning, other things, unspecified things. Values: 0'not use home PC this way' 1'yes, uses home PC this way'.							
hwp hwww hemail hdatan hdatan hdesign hgames hfax hencyrom hmusicc hmusicl hphoto hprog hhomefin hspread hwebdes hscan hothor hunspec	11391	50.2					11289

Derived WORK COMPUTER variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
pcwork1 5 category variable combining information from pcwork and wpcuse . Values 0'no' 1'yes, use it less than once a week' 2'yes, use it once a week' 3'yes, use it 2-4 times a week' 4'yes, use it daily'.	3927	17.3					18753

Derived variables from multi-coded howuse34-howuse46. For 12025 (64.1%) CMs who use a PC at work. Ways CM uses work computer: *for word processing, WWW, e-mail, data analysis, data base work, design packages, playing games, sending/receiving faxes, using CD ROM or Encyclopedia, composing music, listening to music, photography, programming, spreadsheets, webb design, scanning, other things, unspecified things. Values: 0'not use work PC this way' 1'yes, uses work PC this way'.*

wwp	10655	47.0					12025
www							
wemail							
wdata							
wdatab							
wdesign							
wgames							
wfax							
wencyrom							
wmusicc							
wmusicl							
wphoto							
wprog							
wother							

SPSS SYNTAX

Computer use at home

```
compute pchome1 = -1.  
if (pchome = 2) pchome1 = 0.  
if (pchome = 1 and hpcuse = 1) pchome1 = 1.  
if (pchome = 1 and hpcuse = 5) pchome1 = 2.  
if (pchome = 1 and hpcuse = 4) pchome1 = 3.  
if (pchome = 1 and hpcuse = 3) pchome1 = 4.  
if (pchome = 1 and hpcuse = 2) pchome1 = 5.  
missing values pchome1 (-1).  
variable labels pchome1 'Does CM have and use a PC at home?'.  
value labels pchome1 0'no' 1'yes, but do not use it' 2'yes, use it less than once a week' 3'yes, use it  
once a week' 4'yes, use it 2-4 times a week' 5'yes, use it daily'.
```

```
frequencies  
  variables=pchome1.
```

```
do if (pchome1 >= 0).  
  compute hwp = 0.  
  compute hwww = 0.  
  compute hemail = 0.  
  compute hdatan = 0.  
  compute hdatab = 0.  
  compute hdesign = 0.  
  compute hgames = 0.  
  compute hfax = 0.  
  compute hencyrom = 0.  
  compute hmusicc = 0.  
  compute hmusicl = 0.  
  compute hphoto = 0.  
  compute hprog = 0.  
  compute hhomefin = 0.  
  compute hspread = 0.  
  compute hwebdes = 0.  
  compute hscan = 0.  
  compute hother = 0.  
  compute hunspec = 0.  
do repeat x = howuse01 howuse02 howuse03 howuse04 howuse05 howuse06 howuse07 howuse08  
howuse09 howuse10 howuse11 howuse12 howuse13 howuse14 howuse15 howuse16 howuse17 .  
if (x = 1) hwp = 1.  
if (x = 2) hwww = 1.  
if (x = 3) hemail = 1.  
if (x = 4) hdatan = 1.  
if (x = 5) hdatab = 1.  
if (x = 6) hdesign = 1.  
if (x = 7) hgames = 1.  
if (x = 8) hfax = 1.  
if (x = 9) hencyrom = 1.  
if (x = 10) hmusicc = 1.  
if (x = 11) hmusicl = 1.  
if (x = 12) hphoto = 1.  
if (x = 13) hprog = 1.  
if (x = 14) hhomefin = 1.  
if (x = 15) hspread = 1.  
if (x = 16) hwebdes = 1.  
if (x = 17) hscan = 1.  
if (x = 18) hother = 1.  
if (x = 19) hunspec = 1.
```

```
end repeat.  
end if.
```

```
variable labels hwp 'CM uses home PC for word processing'.  
variable labels hwww 'CM uses home PC for WWW'.  
variable labels hemail 'CM uses home PC for e-mail'.  
variable labels hdatan 'CM uses home PC for data analysis'.  
variable labels hdatab 'CM uses home PC for data base work'.  
variable labels hdesign 'CM uses home PC for design packages'.  
variable labels hgames 'CM uses home PC for playing games'.  
variable labels hfax 'CM uses home PC for sending/receiving faxes'.  
variable labels hencyrom 'CM uses home PC for using CD ROM or Encyclopedia'.  
variable labels hmusicc 'CM uses home PC for composing music'.  
variable labels hmusicl 'CM uses home PC for listening to music'.  
variable labels hphoto 'CM uses home PC for photography'.  
variable labels hprog 'CM uses home PC for programming'.  
variable labels hhomefin 'CM uses home PC for managing home finances'.  
variable labels hspread 'CM uses home PC for using spreadsheets'.  
variable labels hwebdes 'CM uses home PC for webb design'.  
variable labels hscan 'CM uses home PC for scanning'.  
variable labels hother 'CM uses home PC for other things'.  
variable labels hunspec 'CM uses home PC for unspecified things'.  
value labels hwp hwww hemail hdatan hdatab hdesign hgames hfax hencyrom hmusicc hmusicl  
hphoto hprog hhomefin hspread hwebdes hscan hother hunspec 0'not use home PC this way' 1'yes,  
uses home PC this way'.
```

frequencies

```
variables=pchome1 hwp hwww hemail hdatan hdatab hdesign hgames hfax hencyrom hmusicc  
hmusicl hphoto hprog hhomefin  
hspread hwebdes hscan hother hunspec .
```

Computer use at work

```
compute pcwork1 = -1.  
if (pcwork = 2) pcwork1 = 0.  
if (pcwork = 1 and wpcuse = 4) pcwork1 = 1.  
if (pcwork = 1 and wpcuse = 3) pcwork1 = 2.  
if (pcwork = 1 and wpcuse = 2) pcwork1 = 3.  
if (pcwork = 1 and wpcuse = 1) pcwork1 = 4.  
missing values pcwork1 (-1).  
variable labels pcwork1 'Does CM use a PC at work?'.  
value labels pcwork1 0'no' 1'yes, use it less than once a week' 2'yes, use it once a week'  
3'yes, use it 2-4 times a week' 4'yes, use it daily'.
```

frequencies

```
variables=pcwork1.
```

```
do if (pcwork1 >= 0).  
compute wwp = 0.  
compute wwww = 0.  
compute wemail = 0.  
compute wdatan = 0.  
compute wdatab = 0.  
compute wdesign = 0.  
compute wgames = 0.  
compute wfax = 0.  
compute wencyrom = 0.  
compute wmusicc = 0.
```



```

compute wmusicl = 0.
compute wphoto = 0.
compute wprog = 0.
compute wother = 0.
do repeat x = howuse34 howuse35 howuse36 howuse37 howuse38 howuse39 howuse40 howuse41
  howuse42 howuse43 howuse44 howuse45 howuse46 .
if (x = 1) wwp = 1.
if (x = 2) wwwww = 1.
if (x = 3) wemail = 1.
if (x = 4) wdatan = 1.
if (x = 5) wdatab = 1.
if (x = 6) wdesign = 1.
if (x = 7) wgames = 1.
if (x = 8) wfax = 1.
if (x = 9) wencyrom = 1.
if (x = 10) wmusicc = 1.
if (x = 11) wmusicl = 1.
if (x = 12) wphoto = 1.
if (x = 13) wprog = 1.
if (x = 14) wother = 1.
end repeat.
end if.

```

```

variable labels wwp 'CM uses work PC for word processing'.
variable labels wwwww 'CM uses work PC for WWW'.
variable labels wemail 'CM uses work PC for e-mail'.
variable labels wdatan 'CM uses work PC for data analysis'.
variable labels wdatab 'CM uses work PC for data base work'.
variable labels wdesign 'CM uses work PC for design packages'.
variable labels wgames 'CM uses work PC for playing games'.
variable labels wfax 'CM uses work PC for sending/receiving faxes'.
variable labels wencyrom 'CM uses work PC for using CD ROM or Encyclopedia'.
variable labels wmusicc 'CM uses work PC for composing music'.
variable labels wmusicl 'CM uses work PC for listening to music'.
variable labels wphoto 'CM uses work PC for photography'.
variable labels wprog 'CM uses work PC for programming'.
variable labels wother 'CM uses work PC for other things'.

```

```

value labels wwp wwwww wemail wdatan wdatab wdesign wgames wfax wencyrom wmusicc wmusicl
wphoto wprog wother
0'not use' 1'uses'.

```

frequencies

```

variables=wwp wwwww wemail wdatan wdatab wdesign wgames wfax wencyrom wmusicc wmusicl
wphoto wprog wother .

```

DRUG QUESTIONS

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
CANNABIS	283	1.2	13	0.1	6	0.0	22378
ECSTASY	283	1.2	12	0.1	7	0.0	22378
AMPHET	283	1.2	11	0	6	0.0	22380
LSD	283	1.2	11	0	6	0.0	22380
POPPER	283	1.2	11	0	8	0.0	22378
MAGMUSH	283	1.2	11	0	7	0.0	22379
COCAINE	283	1.2	11	0	7	0.0	22379
TEMAZ	283	1.2	11	0	9	0.0	22377
SEMERON	283	1.2	11	0	10	0.0	22376
KETAMINE	283	1.2	11	0	8	0.0	22378
CRACK	283	1.2	11	0	6	0.0	22380
HEROIN	283	1.2	11	0	6	0.0	22380
METHAD	283	1.2	11	0	6	0.0	22380
OTHRUG	283	1.2	11	0	7	0.0	22379

DRUG-DRUG6: multi-coded string variables for the 614 CMs who answered 'yes' to **othdrug**. 22378 are 'numeric missing' leaving just 302 individual answers in drug, 37 in drug2, 13 in drug3, 8 in drug4, 4 in drug5, 2 in drug6. Each answer needs to be converted from string to numeric.

DERIVED DRUG VARIABLES

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
<i>If CMs said they had taken SEMERON they were omitted from derived variables on illegal drug use. 43 (0.2%) said they had taken the drug SEMERON.</i>							
anydrug1	347	1.5					22333
<i>Takes information from the 12 individual illegal drug questions (information in othdrug not coded yet). Counts the number of illegal drugs CM has taken in the last 12 months. 13.5% of CMs have taken at least 1 illegal drug in last 12 months.</i>							
drugs1	347	1.5					22333
<i>Collapses anydrug1 variable into 0'not taken any drug' 1'yes, taken at least 1 illegal drug in last 12 months'.</i>							
anydrug2	347	1.5					22333
<i>Takes information from the 12 individual illegal drug questions (information in othdrug not coded yet). Counts the number of illegal drugs CM has ever taken. 43.5% of CMs have taken at least 1 illegal drug in their lifetime.</i>							
drugs2	347	1.5					22333
<i>Collapses anydrug2 variable into 0'never taken an illegal drug' 1'yes, taken at least 1 illegal drug in lifetime'.</i>							

SPSS SYNTAX

missing values cannabis ecsacy amphet lsd popper magmush cocaine temaz semeron ketamine
crack heroin methad othdrug (8,9).

```
do if (semeron = 1).  
count anydrug1 = cannabis ecsacy amphet lsd popper magmush cocaine temaz ketamine crack  
heroin methad (3) /  
    anydrug2 = cannabis ecsacy amphet lsd popper magmush cocaine temaz ketamine crack  
heroin methad (2, 3).  
end if.
```

```
variable labels anydrug1 'CM taken any illegal drug in last 12 months'.  
variable labels anydrug2 'CM ever taken any illegal drug'.
```

```
frequencies  
    variables=anydrug1 anydrug2.
```

```
recode anydrug1 (0=0) (1 thru 12=1) into drugs1 .  
recode anydrug2 (0=0) (1 thru 12=1) into drugs2.  
variable labels drugs1 'CM taken any illegal drug in last 12 months'.  
variable labels drugs2 'CM ever taken any illegal drug'.  
value labels drugs1 drugs2 0'no' 1'yes'.
```

```
frequencies  
    variables=drugs1 drugs2.
```

DIET FOOD EXERCISE

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
fruit	66	0.3	9		28	0.1	22577
eggs	66	0.3	10		29	0.1	22575
salads	66	0.3	10		28	0.1	22576
cookdveg	66	0.3	10		26	0.1	22578
oilfried	66	0.3	11		37	0.2	22566
chops	66	0.3	11		45	0.2	22558
sweets	66	0.3	11		27	0.1	22576
cakes	66	0.3	11		26	0.1	22577
whlbread	66	0.3	11		27	0.1	22576
othbread	66	0.3	11		26	0.1	22577
redmeat	66	0.3	11		26	0.1	22577
poultry	66	0.3	11		26	0.1	22577
fish	66	0.3	11		26	0.1	22577
pulses	66	0.3	11		27	0.1	22576
veggy	66	0.3	9		26	0.1	22579
<i>808 (3.6%) CMs vegetarian. 808 asked the type of vegetarian they were</i>							
vegtype							808
spshdiet	66	0.3	9		26	0.1	22579
<i>1061 (4.7%) CMs followed a special diet. 1061 asked the type of special diet they followed. Only 1009 (95% of 1061) asked if the diet had been recommended by a doctor</i>							
diettype							1061
dietdoc							1009
exercise	66	0.3	10		27	0.1	22577
<i>17263 (76.3%) CMs do regular exercise. 17263 asked how often they take regular exercise and if they become breathless/sweaty during exercis</i>							
breathle					2	0	17261
sweat					3	0	17260

EATING PROBLEMS

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
eatprob	66	0.3	6		21	0.1	22587
<i>729 (3.2%) CMs reported an eating disorder. 729 asked name of eating problem(s) they have had in multi-coded variables</i>							
eating1-eating4 (eating4 holds no information)	21951	96.8	1				728
el1age	21951	96.8			⁵	0.7	724
el112m	21951	96.8	1	0.1	1	0.1	727
<i>349 (47.9%) CMs had an eating disorder in last 12 months. 349 asked if they had seen a doctor re: eating disorder</i>							
el1doc	22331						349

Derived EATING PROBLEMS variables

variable	system-missing		user-missing: 8 (^ indicates 98)		user-missing: 9 (^ indicates 99)		valid n
	n	%	n	%	n	%	
eatprnow	95	0.4					22585
<i>3 category variable combining information given in eatprob and el112m. Values: 0'no eating problems' 1'previous eating problem' 2'current eating problem'</i>							
bulemia	93	0.4					22587
anorexia	93						22587
swallow	93						22587
otheatpr	93						22587
<i>2 category variables. If CM gave valid answer to eatprob, used information in multi-coded eating1-eating4 variables. Values: 0'not this eating problem' 1'yes, this eating problem'</i>							

SPSS SYNTAX

```
missing values eatprob eating1 eating2 eating3 eating4 el112m el1doc (8,9).  
missing values el1age (98,99).
```

```
compute eatprnow = -1.  
if (eatprob = 2) eatprnow = 0.  
if (eatprob = 1 and el112m = 2) eatprnow = 1.  
if (eatprob = 1 and el112m = 1) eatprnow = 2.  
missing values eatprnow (-1).  
variable labels eatprnow 'CM suffers from eating disorders?'.  
value labels eatprnow 0'no eating problems' 1'previous eating problem' 2'current eating problem'.
```

```
frequencies  
  variables=eatprnow.
```

```
do if (eatprob <= 2).  
  compute bulemia = 0.  
  compute anorexia = 0.  
  compute swallow = 0.  
  compute otheatpr = 0.  
  do repeat x = eating1 eating2 eating3 eating4 .  
    if (x = 1) bulemia = 1.  
    if (x = 2) anorexia = 1.  
    if (x = 3) swallow = 1.  
    if (x = 4) otheatpr = 1.  
  end repeat.  
end if.  
variable labels bulemia 'cm has suffered from bulemia'.  
variable labels anorexia 'cm has suffered from anorexia'.  
variable labels swallow 'cm has suffered from swallowing difficulties'.  
variable labels otheatpr 'cm has suffered from otheatpr'.  
value labels bulemia anorexia swallow otheatpr 0'not this eating problem' 1'yes, this eating problem'.
```

```
frequencies  
  variables=bulemia anorexia swallow otheatpr .
```

SMOKING QUESTIONS

variable	system-missing		user-missing: 8 (^ indicates 98; ^^indicates 998)		user-missing: 9 (^ indicates 99; ^^indicates 999)		valid n
	n	%	n	%	n	%	
smoking 6222 (27.4%) CMs smoke every day. 6222 sked how many they smoke a day	66	0.3	8		26	0.1	22580
nofcigs 6347 (28%) CMs ex-smokers or occasional smokers. 6347 asked if ever smoked cigarettes regularly	16458	72.6			^^16	0.3	6206
exsmoker 4532 (71.4%) CMs had smoked regularly. 4532 asked age last smoked regularly	16333	72.0					6347
agequit	18148	80.0			^6	0.1	4526
othsmoke 5864 (25.9%) live with someone who smokes in CM home. 5864 asked whether this was spouse/partner or someone else	2382	10.5	8		29	0.1	20261
whosmoke 4521 (77.1%) CMs said partner smoked. 4521 asked how many spouse/partner smoked a day	16816	74.1					5864
partcigs	18159	80.1	1		27	0.1	4493

Derived SMOKING variable

variable	system-missing		user-missing: 8 (^ indicates 98; ^^indicates 998)		user-missing: 9 (^ indicates 99; ^^indicates 999)		valid n
	n	%	n	%	n	%	
smoke 6 category variables combining information from smoking and nofcigs . Values:0'never smoked' 1'ex-smoker' 2'occasional smoker' 3'up to 10 a day' 4'11 to 20 a day' 5'more than 20 a day'.	116	0.5					22564

SPSS SYNTAX

missing values smoking othsmoke partcigs (8,9).
missing values nofcigs (999).
missing values agequit (99).

```
compute smoke = -1.  
if (smoking = 1) smoke = 0.  
if (smoking = 2) smoke = 1.  
if (smoking = 3) smoke = 2.  
if (smoking = 4 and nofcigs <= 10) smoke = 3.  
if (smoking = 4 and (nofcigs > 10 and nofcigs <=20)) smoke = 4.  
if (smoking = 4 and nofcigs > 20) smoke = 5.  
missing values smoke (-1).  
variable labels smoke 'CM smoking habits'.  
value labels smoke 0'never smoked' 1'ex smoker' 2'occassional smoker' 3'up to 10 a day' 4'11 to 20 a  
day' 5'more than 20 a day'.
```

```
frequencies  
  variables=smoke.
```


DRINKING QUESTIONS

variable	system-missing		user-missing: 8 (^ indicates 98; ^^indicates 998)		user-missing: 9 (^ indicates 99; ^^indicates 999)		valid n
	n	%	n	%	n	%	
drinks	66	0.3	7		27	0.1	22580
<i>18260 (80.7%) CMs drank alcohol regularly (daily, weekly or monthly). 18260 asked how many of each type of alcoholic drink they had in last 7 days prior to interview</i>							
beer	4420	19.5			^^^16	0.1	18244
spirits	4420	19.5	^^^1		^^^16	0.1	18243
wine	4420	19.5	^^^1		^^^3		18256
sherry	4420	19.5	^^^1		^^^1		18258
pops	4420	19.5	^^^1		^^^1		18258
othdrink	4420	19.5	^^^1				18259
cage1	493	2.2			5		22182
<i>5521 (24.9%) CMs ever felt should cut down on their drinking. 5521 asked if felt this in last year</i>							
cage2	17159	75.7			1		5520
cage3	493	2.2					22187
<i>2193 (9.7%) CMs ever felt criticised about their drinking. 2193 asked if felt this in last year</i>							
cage4	20487	90.3					2193
cage5	493	2.2					22187
<i>2693 (11.9%) CMs ever felt bad or guilty about their drinking. 2693 asked if felt this in last year</i>							
cage6	19987	88.1					2693
cage7	493	2.2	1		1		22185
<i>874 (3.9%) CMs ever had a drink first thing in A.M. to steady hands. 874 asked if they had in last year</i>							
cage8	21806	96.1					874
cage9	4208	18.6			4	0	18468
<i>2539 (11.2%) CMs ever drink during breaks at work. 2539 asked how often they did this</i>							
cage10	20141	88.8			2		2537

Derived DRINKING variables

variable	system-missing		user-missing: 8 (^ indicates 98; ^^indicates 998)		user-missing: 9 (^ indicates 99; ^^indicates 999)		valid n
	n	%	n	%	n	%	
drkunits	44250						18230
<i>Continuous variable giving total number of alcohol units CM had in 7 days prior to interview. Combines information given in beer, spirits, wine, sherry and pops. Range from 0-427. Mean: 19.26, Median: 11.00, sd: 25.77</i>							
drkcut	499	2.2					22181
<i>3 category variable combining cag1 and cage2. Values: 0'no' 1'yes, previously' 2'yes, last 12 months'.</i>							
drkcrit	493	2.2					22187
<i>3 category variable combining cag3 and cage4. Values: 0'no' 1'yes, previously' 2'yes, last 12 months'.</i>							
drkbad	493	2.2					22187
<i>3 category variable combining cag5 and cage6. Values: 0'no' 1'yes, previously' 2'yes, last 12 months'.</i>							
drkhand	495	2.2					22185
<i>3 category variable combining cag7 and cage8. Values: 0'no' 1'yes, previously' 2'yes, last 12 months'.</i>							
drkwork	4214	18.6					18466
<i>3 category variable combining cag9 and cage10. Values: 1'yes, special occasions' 2'yes, 2 or 3 times a month' 3'yes, 2 or 3 times a week' 4'yes, most days'.</i>							

SPSS SYNTAX

missing values drinks (8,9).
missing values beer spirits wine sherry pops (998,999).
missing values othdrink (8).

compute drkunits = ((beer * 2) + spirits + wine + sherry + pops).
variable labels drkunits 'CM alcohol units in a week'.

frequencies
variables=drkunits.

missing values cage1 cage2 cage3 cage4 cage5 cage6 cage7 cage8 cage9 cage10 (9).

compute drkcut = -1.
if (cage1 = 2) drkcut = 0.
if (cage1 = 1 and cage2 = 2) drkcut = 1.
if (cage1 = 1 and cage2 = 1) drkcut = 2.
missing values drkcut (-1).
variable labels drkcut 'CM felt should cut down on drinking?'.

compute drkcrit = -1.
if (cage3 = 2) drkcrit = 0.
if (cage3 = 1 and cage4 = 2) drkcrit = 1.
if (cage3 = 1 and cage4 = 1) drkcrit = 2.
missing values drkcrit (-1).
variable labels drkcrit 'CM felt criticised about drinking?'.

compute drkbad = -1.
if (cage5 = 2) drkbad = 0.
if (cage5 = 1 and cage6 = 2) drkbad = 1.
if (cage5 = 1 and cage6 = 1) drkbad = 2.
missing values drkbad (-1).
variable labels drkbad 'CM felt bad/guilty about drinking?'.

compute drkhand = -1.
if (cage7 = 2) drkhand = 0.
if (cage7 = 1 and cage8 = 2) drkhand = 1.
if (cage7 = 1 and cage8 = 1) drkhand = 2.
missing values drkhand (-1).
variable labels drkhand 'CM had drink to steady hands?'.

compute drkwork = -1.
if (cage9 = 2) drkwork = 0.
if (cage9 = 1 and cage10 = 4) drkwork = 1.
if (cage9 = 1 and cage10 = 3) drkwork = 2.
if (cage9 = 1 and cage10 = 2) drkwork = 3.
if (cage9 = 1 and cage10 = 1) drkwork = 4.
missing values drkwork (-1).
variable labels drkwork 'CM had drink at work?'.

value labels drkcut drkcrit drkbad drkhand 0'no' 1'yes, previously' 2'yes, last 12 months'.
value labels drkwork 0'no' 1'yes, special occasions' 2'yes, 2 or 3 times a month' 3'yes, 2 or 3 times a week' 4'yes, most days'.

frequencies
variables=drkcut drkcrit drkbad drkhand drkwork .

ACCIDENT QUESTIONS

variable	system-missing		user-missing: 98 (^ indicates 998)		user-missing: 99 (^ indicates 999)		valid n
	n	%	n	%	n	%	
<i>Multi-coded questions ask if CM had an accident, and if yes, what type of accident.</i>							
accidan1	66	0.3	8	0	24	0.1	22582
accidan2							
accidan3							
accidan4							
accidan5							
accidan6							
accidan7							
accidan8							
accidan9							
<i>12458 (55%) reported no accidents with 10124 (45%) reporting they had some kind of accident. These CMs asked how many accidents they had had in total, the age and type of their most recent accident</i>							
accidno	12524	55.2	^10	0.1	^42	0.4	10104
accage	12576	55.4	^1	0	^18	0.2	10085
accwhy	12576	55.4	1	0	2	0	10101

Derived ACCIDENT variables

variable	system-missing		user-missing: 98 (^ indicates 998)		user-missing: 99 (^ indicates 999)		valid n
	n	%	n	%	n	%	
<i>2-category derived variables from multi-coded accidan1-accidan9. Each variable gives % experiencing type of accident: road accident (pedestrian), road accident (driver/passenger), at work, at home, at school/college, playing sport, other type, violent assault, sexual assault. Values: 0'not experienced' 1'yes, experienced'.</i>							
accroadp	98	0.4					22582
accroadd							
accwork							
acchome							
accschcl							
accsport							
accother							
vassault							
assex							

SPSS SYNTAX

missing values accidan1 (98,99).

frequencies

variables=accidan1 accidan2 accidan3 accidan4 accidan5 accidan6 accidan7 accidan8 accidan9
accidno accage accwhy.

```
do if (accidan1 >= 0).  
compute accroadp = 0.  
compute accroadd = 0.  
compute accwork = 0.  
compute acchome = 0.  
compute accschcl = 0.  
compute accsport = 0.  
compute accother = 0.  
compute vassault = 0.  
compute asssex = 0.  
do repeat x = accidan1 accidan2 accidan3 accidan4 accidan5 accidan6 accidan7 accidan8 accidan9.  
if (x = 1) accroadp = 1.  
if (x = 2) accroadd = 1.  
if (x = 3) accwork = 1.  
if (x = 4) acchome = 1.  
if (x = 5) accschcl = 1.  
if (x = 6) accsport = 1.  
if (x = 7) accother = 1.  
if (x = 8) vassault = 1.  
if (x = 9) asssex = 1.  
end repeat.  
end if.
```

variable labels accroadp 'CM experienced road accident as pedestrian'.
variable labels accroadd 'CM experienced road accident as driver/passenger'.
variable labels accwork 'CM experienced accident at work'.
variable labels acchome 'CM experienced accident at home'.
variable labels accschcl 'CM experienced accident at school/college'.
variable labels accsport 'CM experienced accident when doing sport'.
variable labels accother 'CM experienced other sort of accident'.
variable labels vassault 'CM experienced violent assault or mugging'.
variable labels asssex 'CM experienced sexual assault'.

value labels accroadp accroadd accwork acchome accschcl accsport accother vassault asssex 0'not
experienced' 1'yes, experienced'.

frequencies

variables=accroadp accroadd accwork acchome accschcl accsport accother vassault asssex .

APPENDIX 6: Problems with reporting of GCSEs

Documentation produced by Andrew Jenkins of the Institute of Education relating to the reporting of GCSEs is reproduced below.

Queries about the contents of this appendix should be directed to the User Support Group
(cohort@cls.ioe.ac.uk)

Report on GCSE date errors in NCDS/BCS 2000 (Andrew Jenkins)

This report summarises the findings of a check on the dates at which cohort members in NCDS/BCS 2000 claim to have obtained their GCSEs.

1. The Extent of the Problem

The earliest date at which it appears to be possible to have obtained a GCSE is 1988, as indicated by the following quotations:

'GCSE was introduced in September 1986, with the first exams taking place in Summer 1988. It replaced GCE 'O' level and CSE' (British Qualifications, 25 ed, 1994, Kogan Page).

'Pupils leaving school in Summer 1988 sat either GCSEs or a mixture of 'O' levels and CSEs (and some sat all three types of exam)'. Dept of Education and Science, *Statistics of School Examinations*, 1990.

'GCSE examinations were introduced in 1988. The final GCE 'O' levels examinations for home students were held in the winter session 1987, with summer 1987 being the final session for CSE exams'. Dept for Education, *Statistics of School Examinations*, 1991.

Some 94 NCDS cohort members and over 1,000 BCS cohort members (CMs) are reporting having obtained one or more GCSEs before 1988 in the 2000 sweep of the data. All of the 94 NCDS CMs and the first 200 BCS CMs have been checked against earlier sweeps of the NCDS/BCS surveys.

2. GCSE Date Errors for the NCDS

In the 2000 sweep of the data, those NCDS CMs who had participated in the previous sweep in 1991 were asked to supply information on all qualifications which they had obtained since 1991; those who were absent from the previous sweep were asked about all qualifications which they had obtained since 1974, or school leaving age. Asking for information over such a long spell of time is very likely to introduce inaccuracies due to imperfect recall. However, for NCDS, the 2000 sweep represents the sixth follow-up on the lives of cohort members which means that there is potentially a lot of information in earlier sweeps which we can use for checking. Back in 1978, schools and colleges attended by cohort members were contacted and asked to supply details of the qualifications which NCDS CMs had obtained. This is a valuable resource, since it represents an objective source of information on qualifications obtained. The main drawback is that, obviously, it only covers the period up to 1978. In what follows it is referred to as the EXAMS file. It contains data on the number of qualifications obtained, and also some data on the subjects of qualifications. NCDS IV (1981) appears to contain little or nothing on qualifications obtained, but NCDS V (1991) is more informative. In NCDS V CMs were asked (a) about qualifications ever obtained and (b) about qualifications obtained since 1981. Note that the NCDS questionnaire did ask for GCSEs, as well as O levels and CSEs. The main drawback of NCDS V is that it does not tell us about the number of qualifications.

For example, CMs were asked if they had CSEs at grades 2-5, but were not asked to report how many they had obtained. It also, of course, has problems of recall errors. The EXAMS file, supplemented by NCDS V, then, are the main sources which we can use to check the accuracy of the data in NCDS VI (2000).

A detailed case-by-case listing for each of the 94 NCDS records in question is contained in Appendix 1, and this should be referred to when considering the summary given here.

Fortunately, most of the 94 cases were present in the EXAMS file, and since the majority of erroneous cases were reporting that they got their GCSEs in the 1973 to 1976 period, the data contained in the EXAMS file is relevant. In more detail, all but ten of the 94 cases were present in the EXAMS file. The cases which were not in EXAMS were cases 2,16,26,45,62,74,80,83,91,93. Out of these, five (16,26,74,80,83) were present in NCDS 5. There is no data for the remaining five.

For some of the remaining cases, EXAMS file data was found to be not relevant because they were reporting GCSE qualifications after 1978 but fortunately, this applied to only a small number of cases. These cases were (3,11,18,30,77,81,82). Some of these (3,30,82) were covered by NCDS V, but the remaining 4 were not.

In sum, then, there were only 9 cases where data could not be obtained from earlier sweeps to check the GCSE errors in NCDS VI.

The majority of those for whom data is available would appear to have CSEs rather than the GCSEs reported in NCDS VI. However, some have O levels, and some a mixture of O levels and CSEs. To give some specific examples, cases 13 and 15 are just two examples of a straightforward mix-up between GCSEs and CSEs (for case 13 there is also a small difference in number of GCSE/CSEs reported – 7 in the EXAMS file and 6 in 2000 sweep). In cases 9 and 37, on the other hand, GCSEs and O levels would appear to have been muddled up by the CMs concerned. Although a good many cases appear to show confusion between GCSEs and either O levels or CSEs, some cases are less simple. For instance, in case 94, there is a huge discrepancy between the number of O levels in the EXAMS file and the number of GCSEs listed in NCDS VI. Cases 19 and 52 are other cases where there is little evidence in the EXAMS file that the qualifications reported in NCDS VI were actually obtained. On the other hand, some cases can be easily sorted out even in the absence of data. Case 2 shows duplication and the GCSE data can be deleted.

It is good that we have some data to check NCDS VI against for most of the 94 problem cases. Most of the GCSEs would appear to be CSEs, in fact, but some cases they are O levels or a mixture of O levels and CSEs, and in some cases it is not simple to decide what the best answer is. Nonetheless, it is probably best to deal with them on a case by case basis rather than adopting a simple but inaccurate solution such as simply recoding all GCSEs to O levels or all GCSEs to CSEs.

3. GCSE date errors among a sample of BCS 2000 CMs.

Because there are such a large number of GCSE errors among the BCS cohort, it has only been possible to look at the first 200 of these (rather less than a fifth of the total).

The BCS CMs were one of the last cohorts to sit O levels and CSEs as their school-leaving exams before they were replaced by GCSEs. BCS CMs should only have GCSEs if they stayed in education to at least age 18, or subsequently returned to study for a qualification.

The only source of data for checking the responses in BCS 2000 is the earlier BCS survey conducted in 1996. Questions about qualifications in the BCS 1996 survey were asked in the order: CSE, O level, GCSE, rather than the order GCSE, O level, CSE in BCS 2000. Thus checking BCS 2000 against BCS 1996 should help to root out errors which may have arisen from the order in which the questions were asked in BCS 2000. However, BCS 1996 had a relatively small sample size (n = 9,003), so many of those in BCS 2000 may not be present in BCS 1996.

Of the 200 BCS CMs checked, some 134 were also present in BCS 96. Out of these 134, 42 were also reporting GCSEs in BCS 96 (although not necessarily the same number as in BCS 2000); 17 provided no information about qualifications in BCS 96, and 75 did not report any GCSEs in BCS 96. Detailed information on each case is presented in Appendix 2.

Among the 134 for whom 1996 data are available it is usually possible to make some reasonable guess as to what should be done about the erroneous GCSE data in BCS 2000. For example, it is clear that in cases 8 and 11 the GCSE data should be recoded to O levels, while in case 41 the GCSEs should be recoded as CSEs. Other cases are not quite so obvious, and some, where there are large discrepancies between 1996 and 2000 (e.g. cases 23, 30) or conversely where errors have been repeated (e.g. 31) are fairly puzzling. The real difficulty is, however, that we have a substantial number of cases not present in 1996, and so there is no data to check the BCS 2000 results against.

4. O levels and CSEs after 1988 reported by CMs in NCDS/BCS 2000

I have also had a quick look at whether any NCDS cohort members are reporting having obtained O levels after they were (presumably) phased out in 1988. Among all the NCDS CMs, some 59 had reported one or more O levels with a date post 1988.

The dates for the reported O levels were all in the 1990s. The source of the error is likely to be either that (a) they have actually got GCSEs rather than O levels or (b) a mistake was made inputting the data, so that the O level was actually obtained in, say, 1974 but this was coded as 1994.

Now 47 of the 59 were also present in NCDS V, which means that they were (or should have been) only asked questions about qualifications obtained since 1991. This would make it unlikely that error (b) occurred. The data do indeed show that most of these 47 were only reporting O levels obtained in the 1990s: 28 reported a single O level obtained in the 1990s and no earlier O levels; a further 13 reported two O levels obtained in the 1990s and no O levels. However, the remaining six of the 47, despite being present in NCDS V in 1991, also reported O levels obtained in both the 1970s and the 1990s.

For the 12 not present in 1991, it is difficult to discriminate between the (a) and (b) errors listed above.

So it is not possible to resolve all the cases, but for many recoding to GCSE would be a sensible solution.

Only two NCDS CMs reported CSEs after 1988.

For BCS, there appear to be 133 cases in which CMs are reporting O levels after 1988, and 32 cases of CSEs after 1988 but these have not been studied in any detail.

Problems in Using the NCDS/BCS 2000 Data (Andrew Jenkins 27/3/2001)

1) The fact that there are three different sub-groups in the data, namely, the BCS whose qualifications are recorded from 1986; some NCDS who also participated in 1991 and so data on them is from 1991 onwards, and other NCDS who missed out last time and so have data in current sweep back to 1974 caused us some confusion. I think this will not be clear to other potential users and needs to be flagged very clearly in the documentation.

2) As you know, merging data with earlier sweeps is far from straightforward because of confusion about chrtid, and serial variables in the different sweeps. This has been a major headache for the QCA project where we needed to merge different sweeps.

3) Many of the qualification variables have 98 or 99 (or in some cases 8,9 or 9998, 9999) as value labels, but these are not defined. Presumably these are missing values/could not say/would not say, but they need to be defined. This problem is fairly minor for some variables e.g the year people got their CSEs (if they got them in a single year) = cseyrq and this has 6 people coded as having got their CSEs in either 9998 or 9999; but much more serious for other variables e.g 'how many CSEs did you get at grade 1' (edcse1) has 120 people coded as 99 and one coded as 98. Some cases (2,164 for this variable) are also coded as zero but this obviously doesn't cover all the people who have not got a CSE. Why are some coded zero and not others?

4) Problem 3 also contributes to making it very difficult to get *consistent* answers to relatively simple questions such as how many people in the sample have got a CSE.

One way to answer this is to use the derived variable numcse. This reports that 4,621 have a value for this answer and we can deduct the number reporting zero CSEs (251) to work out the number with one or more CSEs (4621-251= 4370).

Another approach to the same question is to look at the variables edqtyp14 to edqtyp26 which report whether individual has got particular qualifications. If we use these variables to create a dummy variable, ncse, taking value 1 if individual has got CSE and zero if not, as follows

```
compute ncse=0.  
if any(3,edqtyp14,edqtyp15,edqtyp16,edqtyp17,edqtyp18,  
edqtyp19,edqtyp20,edqtyp21,edqtyp22,edqtyp23,edqtyp24,edqtyp25,  
edqtyp26) ncse=1.
```

```
execute .
```

(note that 3 is the code for CSE for these variables).

Doing this and running frequencies yields the answer that 4,621 have got (at least one) CSE. Thus we have two answers to the question how many have got one or more CSEs – is it 4,621 or 4,370?

On closer inspection, it looks as if some of those reporting a qualification in the edqtyp group of variables have not in fact got it, so we have tended to use variables such as numcse, knocking out the zeros and 98s etc. But it would be good if we could get consistent answers to the question whichever way it was looked at.

Note that this issue arises with the academic educational qualifications (edqtyp14 to edqtyp26) and with a similar set of vocational qualifications voctyp12 to voctyp22.

5) If the variable Dmpart=1 this implies that NCDS CM was included in 1991 sweep and therefore should not have pre 91 qualifications included on current sweep. However, there are individuals with Dmpart=1 and qualifications recorded with pre 91 dates. For example, with A levels, if we take edyear56, the year in which people got their first A level, then 105 of the 179 individuals with an A

level in NCDS and with Dmpart=1 are recorded as having got their A level before 1991 (mostly in 1976). This problem also arises with other qualifications.

Is this because some individuals who did participate in 1991 have nonetheless had their qualifications recorded again in 2000 or is the Dmpart variable faulty in that it is misallocating significant numbers of people (my money's on the former, I think).

6) As already mentioned, large nos of the BCS people are claiming to have obtained GCSEs in 1986 when they were 16. My impression is that GCSE was not introduced until about 1988 and that many respondents have just got mixed up with O levels and CSEs. Conversely, some people are reported as having got CSEs and O levels after the late 1980s when presumably they had been abolished.

7) Proxy individuals have also caused a little confusion since the variables in which they report qualifications have very similar names to the variable names for the rest of the sample. Again this is probably just be a case of flagging very clearly that they are there in the documentation.

8) There are a couple of typos, as well:

Labelling of the variable eddiped incorrect. Should be 'Does CM know the year s/he got their Dip H Ed' not 'Year obtained Dip H Ed'.

The variable edgcse2 is labelled as 'number of GCSEs obtained at grades D-E'; should be 'number of GCSEs obtained at grades D-G', (?) as with other qualifications such as AS levels.

Summary of Results of Checking Dates at which CM's obtained GCSEs in NCDS/BCS 2000.

Earliest Dates at which a GCSE could have been obtained:

The earliest date at which it appears to be possible to have obtained a GCSE is 1988 (see the quotations below). Some 94 NCDS cohort members and over 1,000 BCS cohort members are reporting having obtained one or more GCSEs before 1988 in the 2000 sweep of the data.

'GCSE was introduced in September 1986, with the first exams taking place in Summer 1988. It replaced GCE 'O' level and CSE' (British Qualifications, 25 ed, 1994, Kogan Page).

'Pupils leaving school in Summer 1988 sat either GCSEs or a mixture of 'O' levels and CSEs (and some sat all three types of exam)'. Dept of Education and Science, *Statistics of School Examinations*, 1990.

'GCSE examinations were introduced in 1988. The final GCE 'O' levels examinations for home students were held in the winter session 1987, with summer 1987 being the final session for CSE exams'. Dept for Education, *Statistics of School Examinations*, 1991.

NCDS

For the 94 NCDS cohort members examing NCDS 5 and the EXAMS file from 1978 provided information on the qualifications obtained by all but nine of the 94.

The following were not in the EXAMS file: cases 2,16,26,45,62,74,80,83,91,93.

Some of these (16,26,74,80,83) were present in NCDS 5.

There is no data for the remaining five.

For some of the remaining cases, EXAMS file data was found to be not relevant because they were reporting GCSE qualifications after 1978. These cases were (3,11,18,30,77,81,82). Some of these (3,30,82) were covered by NCDS 5, but the remaining 4 were not.

In total, then, there were only 9 cases where data could not be obtained from earlier sweeps to check the GCSE errors in NCDS 6.

The majority of those for whom data is available would appear to have CSEs rather than the GCSEs reported in NCDS 6. However, some have O levels, and some a mixture of O levels and CSEs.

BCS

I have looked at the first 200 cases for BCS. Of these, some 134 were present in BCS 96. Out of these 134, 42 were also reporting GCSEs in BCS 96 (although not necessarily the same number as in BCS 2000); 17 provided no information about qualifications in BCS 96, and 75 did not report any GCSEs in BCS 96.

NCDS VI – Results of Checking those with Erroneous Dates for GCSEs against Earlier NCDS Sweeps

There are 94 cases in NCDS VI in which NCDS cohort members have recorded a date for their GCSEs prior to the introduction of the first GCSE exams (1988). These 94 cases are as follows.

Case: 1

Chrtid: 581058E

NCDS VI GCSE/CSE/GCE Qualifications: one GCSE in 1974, also 3 O levels in 1974.

Present in NCDS V:

Present in EXAMS file: Yes

EXAMS file GCE/CSE Qualifications: 3 CSEs at grades 2-5.

Case: 2

Chrtid: 330031E

NCDS VI GCSE/CSE/GCE Qualifications: GCSEs from 1975 in homecrafts, English literature and English language. Also CSEs from 1976 in English Language, English Literature, Geography, Drama and Homecraft. Also O levels from 1977 in English Language, English Literature, Geography, Drama and Religious Knowledge.

Present in NCDS V: No

Present in EXAMS file: No

EXAMS file GCE/CSE Qualifications: N/A.

Case: 3

Chrtid: 750128Q

NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs in 1982, 2 at A-C grades, 1 at lower grade. No O levels or CSEs.

Present in NCDS V: Yes

NCDS V Qualifications: Recorded as having 'other' qualification(s), but not GCSE, CSE or O level.

Present in EXAMS file: Yes

EXAMS file GCE/CSE Qualifications: No qualifications.

Case: 4

Chrtid: 845002U

NCDS VI GCSE/CSE/GCE Qualifications: 8 GCSEs in 1973, 4 at A-C grades and 4 at lower grades.

Present in NCDS V: No

Present in EXAMS file: Yes

EXAMS file GCE/CSE Qualifications: 1 O level/CSE grade 1, 2 CSEs at grade 4, 2 CSEs at grade 5.

Case: 5

Chrtid: 516081F

NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1974, 4 at A-C grades, 1 at lower grade. No O levels or CSEs.

Present in NCDS V:

Present in EXAMS file: Yes

EXAMS file GCE/CSE Qualifications: 1 CSE grade 4, 2 CSE grade 5.

Case: 6

Chrtid: Y30263X

NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1974, 3 at A-C grades, 2 at lower grades. 2 O levels in 1974, no CSEs.

Present in NCDS V:

Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 O level/CSE grade 1; 2 CSE grade 4.

Case: 7
Chrtid: Y30072Q
NCDS VI GCSE/CSE/GCE Qualifications: 6 GCSEs all at A-C grades in 1974. No O levels or CSEs.
Present in NCDS V:
Present in EXAMS file:
EXAMS file GCE/CSE Qualifications: 2 CSEs at grade 3; 2 CSEs at grade 4, 2 CSEs at grade 5.

Case:8
Chrtid:X66014T
NCDS VI GCSE/CSE/GCE Qualifications:3 GCSEs in 1974, no O levels or CSEs.
Present in NCDS V:Yes
Present in EXAMS file:yes
EXAMS file GCE/CSE Qualifications:1 O level at grade A-C.

Case:9
Chrtid:X83029D
NCDS VI GCSE/CSE/GCE Qualifications:7 GCSEs in 1974, all at A-C grades, no O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file:Yes
EXAMS file GCE/CSE Qualifications:7 O levels at A-C, no CSEs.

Case:10
Chrtid:048054V
NCDS VI GCSE/CSE/GCE Qualifications:6 GCSEs in 1974, no O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file:Yes
EXAMS file GCE/CSE Qualifications:1 CSE grade 5.

Case:11
Chrtid:044015N
NCDS VI GCSE/CSE/GCE Qualifications:4 GCSEs – 1 obtained in 1980, 1 in 1994, 2 in 1995
Present in NCDS V: No
Present in EXAMS file:Yes
EXAMS file GCE/CSE Qualifications: No O levels or CSEs.

Case: 12
Chrtid: 620094X
NCDS VI GCSE/CSE/GCE Qualifications: 2 GCSEs in 1974, both at A-C grades, no O levels and no CSEs.
Present in NCDS V: No
Present in EXAMS file:
EXAMS file GCE/CSE Qualifications: 2 CSEs, 1 at grade 2; 1 at grade 4.

Case: 13
Chrtid:100052D
NCDS VI GCSE/CSE/GCE Qualifications: 6 GCSEs in 1974, all at A-C grades
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 3 CSE grade 2; 3 CSE grade 3; 1 CSE grade 5.

Case:14
Chrtid: 517088B
NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs in 1974, 2 at A-C grades and 1 at D-G. Also 1 CSE grade 1 in 1973.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: Took 3 exams in total; passed two: 1 CSE grade 3, 1 CSE grade 4.

Case:15
Chrtid: 081020Q
NCDS VI GCSE/CSE/GCE Qualifications: 1 GCSE in 1974 at A-C grade, no O levels and no CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE at grade 2.

Case: 16
Chrtid: 835005U
NCDS VI GCSE/CSE/GCE Qualifications: 10 GCSEs in 1974, all at grades A-C.
Present in NCDS V: Yes
Qualifications in NCDS V: CSEs at grade 1, and CSEs at grades 2-5.
Present in EXAMS file: No
EXAMS file GCE/CSE Qualifications: N/A.

Case: 17
Chrtid: 093052W
NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs in 1973, 2 of them at A-C grade, and 1 at a lower grade; 1 O level in 1973 at A-C grade; 3 CSEs in 1973, 2 at grade 1, and 1 at a lower grade.
Present in NCDS V: Yes.
Qualifications in NCDS V: CSEs at grade 1, and O levels at A-C grades.
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 1, 1 CSE grade 4, 2 CSE grade 5.

Case: 18
Chrtid: 093070Y
NCDS VI GCSE/CSE/GCE Qualifications: 1 GCSE in 1986; 5 O levels in 1974
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 6 CSEs at grades 2-5: 1 at grade 2, 2 at grade 3, 1 at grade 4, 2 at grade 5.

Case: 19
Chrtid: 0930175P
NCDS VI GCSE/CSE/GCE Qualifications: 6 GCSEs in 1974, all at below grade C.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: No exams taken or passed.

Case: 20

Chrtid: 222005A
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1974 at A-C; 4 CSEs in 1974 at grade 1.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE at grade 4.

Case: 21
Chrtid: 054006T
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1974, 3 at A-C grades, 1 at a lower grade. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE at grade 2, 3 CSEs at grade 5.

Case: 22
Chrtid: 055031X
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1974, all at A-C grades; no O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 4 exams taken, 3 passed: 2 CSE grade 3, 1 CSE grade 4

Case: 23
Chrtid: 093071A
NCDS VI GCSE/CSE/GCE Qualifications: 2 GCSEs in 1974, 1 at grade A-C; no O levels, no CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 1; 1 CSE grade 4.

Case: 24
Chrtid: 083047X
NCDS VI GCSE/CSE/GCE Qualifications: 1 GCSE in 1976 at A-C grade; 6 CSEs in 1974, 3 at grade 1, 3 at grades 2-5.
Present in NCDS V: No
Present in EXAMS file:
EXAMS file GCE/CSE Qualifications: 6 CSEs: 3 at grade 1, 1 at grade 2, 1 at grade 3, 1 at grade 5.

Case: 25
Chrtid: 099010M
NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1974, 4 at A-C grades, 1 at lower grades. No O levels and no CSEs.
Present in NCDS V: Yes
Qualifications in NCDS V: City & Guilds qualifications only.
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: No O levels or CSEs obtained.

Case: 26
Chrtid: 094022S
NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1974 (3 at A-C grades, 2 at lower grades). No O levels, 3 CSEs in 1974, all at grades 2-5.
Present in NCDS V: Yes
Qualifications in NCDS V: CSEs at grades 2-5 only.

Present in EXAMS file: No
EXAMS file GCE/CSE Qualifications: N/A.

Case: 27
Chrtid: 055110T
NCDS VI GCSE/CSE/GCE Qualifications: 8 GCSEs in 1976, 5 at A-C grades, 3 at lower grades. No O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 4 O level/CSE grade 1 passes. 1 CSE grade 2, 1 CSE grade 3.

Case: 28
Chrtid: 092153X
NCDS VI GCSE/CSE/GCE Qualifications: 1 GCSE in 1974 at D-G grade. Also 5 CSEs all at grade 1.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 4 CSEs at grade 1; 1 CSE grade 3.

Case: 29
Chrtid: 984038V
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1974, 2 at A-C grades, 2 at D-G grades. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 2, 1 CSE grade 4, 1 CSE grade 5.

Case: 30
Chrtid: Y30042E
NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs at D-G in 1985. No O levels or CSEs.
Present in NCDS V: Yes
Qualifications in NCDS V: CSEs at grades 2-5; O levels at A-C grades; also RSA and City & Guilds qualifications.
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 5.

Case: 31
Chrtid: 282068R
NCDS VI GCSE/CSE/GCE Qualifications: 2 GCSEs in 1974, 3 more in 1985: 3 at A-C grades, 2 at D-G grades. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 1, 1 CSE grade 3, 1 CSE grade 4.

Case: 32
Chrtid: 110315S
NCDS VI GCSE/CSE/GCE Qualifications: 2 GCSEs at A-C grades, 1 in 1974, the other in 1977; no O levels or CSEs.
Present in NCDS V: Yes

Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 2, 4 CSEs at grade 4.

Case: 33
Chrtid: 330002X
NCDS VI GCSE/CSE/GCE Qualifications: 1 GCSE in 1977 at A-C grade; also 5 CSEs in 1974, 1 at grade 1, and the other 4 at grades 2-5.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 1, 3 CSEs at grade 3, 1 CSE grade 5.

Case: 34
Chrtid: 330092B
NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1974 all at A-C grades; no O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 2 CSEs at grade 3, 2 CSEs at grade 4, 1 CSE at grade 5.

Case: 35
Chrtid: 330009N
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1974, 2 of these at A-C grades, 2 at D-G. 1 O level in 1974 at grades D-E, no CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: No O levels recorded; 1 CSE grade 3, 1 CSE grade 4, 2 CSEs at grade 5.

Case: 36
Chrtid: 330074Z
NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1974, all at grades D-G; also 2 O levels in 1974 at A-C grades; no CSEs.
Present in NCDS V: Yes
Qualifications in NCDS V: O levels at A-C grades.
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 3, 1 CSE grade 5.

Case: 37
Chrtid: 382013W
NCDS VI GCSE/CSE/GCE Qualifications: 2 GCSEs in 1974, both at A-C grades; also 5 CSEs in 1974, all at grades 2-5.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 2 O levels at grades A-C; 1 CSE grade 3, 1 CSE grade 4, 2 CSE grade 5.

Case: 38
Chrtid: 238018Z
NCDS VI GCSE/CSE/GCE Qualifications: 6 GCSEs in 1974; 4 at A-C grades, 2 at D-G grades. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 2 CSE grade 2, 3 CSE grade 3, 2 CSE grade 4.

Case: 39
Chrtid: Y20012P
NCDS VI GCSE/CSE/GCE Qualifications: 7 GCSEs in 1976, 5 at A-C grades, 2 at D-G. No O levels and no CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 2, 3 CSE grade 3, 3 CSE grade 4.

Case: 40
Chrtid: 110106H
NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1974, all at A-C grades. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 3, 1 CSE grade 4, 1 CSE grade 5.

Case: 41
Chrtid: 120011B
NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs in 1974, all at A-C. No O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 2, 1 CSE grade 3, 2 CSE grade 4, 1 CSE grade 5.

Case: 42
Chrtid: 052018Q
NCDS VI GCSE/CSE/GCE Qualifications: 1 GCSE in 1975, 1 in 1996, both at A-C grades. No O levels and no CSEs.
Present in NCDS V: Yes
Qualifications in NCDS V: O levels at A-C grades.
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: No exams taken or passed.

Case: 43
Chrtid: 184008Z
NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs in 1974, 2 at A-C grades, 1 at D-G.
Present in NCDS V: Yes
Qualifications in NCDS V: CSEs at grades 2-5.
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: No exams taken or passed.

Case: 44
Chrtid: 937018X
NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1974, all at A-C grades. No O levels and no CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 2 O level/CSE grade 1, 1 CSE grade 3, 2 CSE grade 5.

Case: 45
Chrtid: 935018M
NCDS VI GCSE/CSE/GCE Qualifications: 1 GCSE in 1974 at A-C grade; no O levels and no CSEs.
Present in NCDS V: No
Present in EXAMS file: No
EXAMS file GCE/CSE Qualifications: N/A.

Case: 46
Chrtid: Y30070L
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1974, all at A-C grades. No O levels or CSEs.
Present in NCDS V: Yes
Qualifications in NCDS V: CSEs at grades 2-5.
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: No exams taken or passed.

Case: 47
Chrtid: 937014P
NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs in 1974, all at A-C grades. No O levels and no CSEs.
Present in NCDS V: Yes
Qualifications in NCDS V: Recorded as having no qualifications.
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: No exams taken or passed.

Case: 48
Chrtid: 381033X
NCDS VI GCSE/CSE/GCE Qualifications: 6 GCSEs in 1974, 1 at A-C grade, 5 at D-G grades. No O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 3, 2 CSE grade 4, 3 CSE grade 5.

Case: 49
Chrtid: 381064K
NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs in 1976, all at A-C grades. No O levels, 3 CSE grade 1 also in 1976.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 4 CSEs at grades 2-5; 1 at grade 2, 1 at grade 3, 1 at grade 4, 1 at grade 5.

Case: 50
Chrtid: 282113T
NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1973, 1 at an unspecified date, all at A-C grades. No O levels, no CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 O level/CSE grade 1, 2 CSE grade 2, 1 CSE grade 3.

Case: 51
Chrtid: 385059P
NCDS VI GCSE/CSE/GCE Qualifications: 13 GCSEs, 10 in 1974, 2 in 1975, 1 in 1995; 12 at A-C grades, 1 at D-G. No O levels, and no CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 8 O levels at A-C grades; no CSEs.

Case: 52
Chrtid: 092346K
NCDS VI GCSE/CSE/GCE Qualifications: 2 GCSEs in 1974, both at A-C grades. No O levels and no CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: No Exams taken or passed.

Case: 53
Chrtid: 527049X
NCDS VI GCSE/CSE/GCE Qualifications: 7 GCSEs in 1974, all at A-C grades. No O levels and no CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 7 exams taken, 6 passes. 5 CSEs at grade 3; 1 CSE at grade 4.

Case: 54
Chrtid: 825099A
NCDS VI GCSE/CSE/GCE Qualifications: 2 GCSEs in 1974, 2 in 1985. 2 of these were at A-C grades, 2 at lower grades. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 3 O level/CSE 1. 2 at CSE grade 2.

Case: 55
Chrtid: 982023V
NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs in 1974, all at A-C grades. No O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 2 CSE grade 4, 1 CSE grade 5.

Case: 56
Chrtid: 982081L
NCDS VI GCSE/CSE/GCE Qualifications: 1 GCSE in 1974, A-C grade. 3 CSEs at grade 1, 2 CSEs at grades 2-5.
Present in NCDS V: No

Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 O level/CSE 1; 1 CSE grade 2, 1 CSE grade 3, 1 CSE grade 4, 1 CSE grade 5.

Case: 57
Chrtid: 960080K
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1974, all at A-C grades; 1 O level in 1974, no CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 1, 2 CSE grade 3, 3 CSE grade 4.

Case: 58
Chrtid: Y30017F
NCDS VI GCSE/CSE/GCE Qualifications: 8 GCSEs in 1974, all at A-C grades; no O levels and no CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 8 O levels, no CSEs listed.

Case: 59
Chrtid: 509226R
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1976, all at A-C grades; no O levels; 7 CSEs in 1974, all at grades 2-5.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 3 O level/CSE grade 1, 2 CSEs at grade 2, 4 CSEs at grade 3, 1 CSE at grade 4.

Case: 60
Chrtid: 511141V
NCDS VI GCSE/CSE/GCE Qualifications: 2 GCSEs in 1973, both at A-C grades. No O levels reported, 6 CSEs in 1974 at grades 2-5.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 O level/CSE grade 1, 1 CSE grade 2, 2 CSE grade 3, 4 CSE grade 5.

Case: 61
Chrtid: Y30098K
NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1974, all at D-G grades; no O levels, 3 CSEs in 1974, all at grades 2-5.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE at grade 3, 2 CSEs at grade 5.

Case: 62
Chrtid: 520034V
NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs in 1974, 1 at A-C grades, 2 at lower grades. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: No
EXAMS file GCE/CSE Qualifications: N/A

Case: 63
Chrtid: 815061S
NCDS VI GCSE/CSE/GCE Qualifications: 12 GCSEs, 11 obtained in 1974 and 1 in 1975, 6 at A-C grades and 6 at lower grades. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 12 exams taken; 9 passes at O level/CSE grade 1.

Case: 64
Chrtid: 500421Q
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1973, all at A-C grades; 2 O levels, 1 at A-C grades, 1 at a lower grade; no CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE at grade 3, 1 CSE at grade 4, 1 CSE at grade 5.

Case: 65
Chrtid: 684015K
NCDS VI GCSE/CSE/GCE Qualifications: 6 GCSEs in 1974, 1 at A-C grade, the rest at lower grades; no O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: No exams taken, none passed

Case: 66
Chrtid: 180025C
NCDS VI GCSE/CSE/GCE Qualifications: 8 GCSEs in 1974, 4 at A-C grades, 4 lower. No O levels and no CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 3 O level/CSE grade 1.

Case: 67
Chrtid: 528005F
NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1974, 4 at A-C, 1 at a lower grade. No O levels and no CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 2 O level/CSE 1.

Case: 68
Chrtid: Y30061K
NCDS VI GCSE/CSE/GCE Qualifications: 2 GCSEs in 1974, both at A-C grades. No O levels and no CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 4, 1 CSE grade 5.

Case: 69
Chrtid: 610020K
NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs in 1976, 2 in 1974, all at A-C grades. 1 CSE grade 1, 2 CSEs at grades 2-5.
Present in NCDS V: Yes

Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 O level/CSE grade 1, 2 CSEs at grade 2, 2 CSEs at grade 3, 1 CSE grade 4.

Case: 70
Chrtid: 583014U
NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1977, all at A-C grades. No O levels, 4 CSEs at below grade 1.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE at grade 5.

Case: 71
Chrtid: 682037K
NCDS VI GCSE/CSE/GCE Qualifications: 2 GCSEs in 1974, both at A-C grades. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 2 O level/CSE grade 1, 4 CSEs at grade 3, 1 CSE grade 5.

Case: 72
Chrtid: 782118S
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1975, all at A-C grades. No O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 3, 1 CSE grade 4.

Case: 73
Chrtid: 782170U
NCDS VI GCSE/CSE/GCE Qualifications: 8 GCSEs in 1974, 5 at A-C grades, 3 at lower grades. Also 2 O levels at A-C grades, no CSEs.
Present in NCDS V: Yes
Qualifications in NCDS V: CSE at grades 2-5, O levels at A-C.
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: No exams taken, no qualifications.

Case: 74
Chrtid: 823501Q
NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs in 1974, all at A-C grades. No O levels or CSEs.
Present in NCDS V: Yes
Qualifications in NCDS V: CSEs at grade 1 and CSEs at grades 2-5.
Present in EXAMS file: No
EXAMS file GCE/CSE Qualifications: N/A

Case: 75
Chrtid: 840002S
NCDS VI GCSE/CSE/GCE Qualifications: 13 GCSEs in 1975, all at A-C grades. No O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 13 exams taken, 8 passes at O level/CSE grade 1.

Case: 76
Chrtid: 840017F

NCDS VI GCSE/CSE/GCE Qualifications: 13 GCSEs in 1974, 10 at A-C grades, 3 at lower grades. No O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 9 O levels/CSE grade 1.

Case: 77
Chrtid: 560023R
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs , 2 in 1986, 2 in 1987, all at A-C grades. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 3 CSEs at grade 3.

Case: 78
Chrtid: 882083H
NCDS VI GCSE/CSE/GCE Qualifications: 9 GCSEs in 1974, all at A-C grades. No O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 8 O levels/CSE grade 1.

Case: 79
Chrtid: 824514E
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1975, all at lower grades. No O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE at grade 2, 1 CSE at grade 3, 1 CSE at grade 4, 2 CSEs at grade 5.

Case: 80
Chrtid: 882031M
NCDS VI GCSE/CSE/GCE Qualifications: 5 GCSEs in 1974, all at A-C grades. No O levels. 1 CSE grade 1 in 1974.
Present in NCDS V: Yes
Qualifications in NCDS V: CSEs at grades 2-5; O levels at grades A-C; GCSEs at grades A-C.
Present in EXAMS file: No
EXAMS file GCE/CSE Qualifications: N/A.

Case: 81
Chrtid: 365003W
NCDS VI GCSE/CSE/GCE Qualifications: 2 GCSEs, 1 in 1986, 1 in 1996, both at A-C grades. 8 O levels, all at A-C grades, no CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 7 O levels/CSE grade 1.

Case: 82
Chrtid: 610012L
NCDS VI GCSE/CSE/GCE Qualifications: 1 GCSE in 1985, at A-C grade. No O levels or CSEs.
Present in NCDS V: Yes
Qualifications in NCDS V: CSEs at grade 1; O levels at grades A-C.
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 3, 1 CSE grade 4, 1 CSE grade 5.

Case: 83
Chrtid: 550237E
NCDS VI GCSE/CSE/GCE Qualifications: 3 GCSEs in 1974, all at A-C grades. No O levels or CSEs.

Present in NCDS V: Yes
Qualifications in NCDS V: CSEs at grades 2-5.
Present in EXAMS file: No
EXAMS file GCE/CSE Qualifications: N/A

Case: 84
Chrtid: 933007V
NCDS VI GCSE/CSE/GCE Qualifications: 7 GCSEs in 1974, all at A-C grades. No O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 3 O level/CSE grade 1, 2 CSE grade 3, 1 CSE grade 4.

Case: 85
Chrtid: Y01195T
NCDS VI GCSE/CSE/GCE Qualifications: 6 GCSEs in 1974, 1 at A-C grades, 5 at lower grades. No O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE at grade 2, 2 CSEs at grade 3, 1 CSE at grade 4, 2 CSEs at grade 5.

Case: 86
Chrtid: 517102S
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1974, all at A-C grades. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE grade 4, 1 CSE grade 5.

Case: 87
Chrtid: 500073T
NCDS VI GCSE/CSE/GCE Qualifications: 8 GCSEs in 1974, all at A-C grades. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 3 CSE at grade 3, 3 CSE at grade 4.

Case: 88
Chrtid: 825035X
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs, 3 in 1986 and 1 in 1975, 1 is at A-C grade, others at lower grade. No O levels, 6 CSEs in 1975, 1 at grade 1, 5 at grades 2-5.
Present in NCDS V: Yes
Qualifications in NCDS V: CSEs at grades 2-5; CSEs at grade 1, O levels at A-C grade.
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 O level/CSE grade 1, 1 CSE grade 2, 3 CSE grade 3, 1 CSE grade 4, 1 CSE grade 5.

Case: 89
Chrtid: 180007A
NCDS VI GCSE/CSE/GCE Qualifications: 1 GCSE in 1973, at below grade C. No O levels, 5 CSEs in 1974 at grades 2-5.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: No exams taken, no qualifications.

Case: 90
Chrtid: 465013H
NCDS VI GCSE/CSE/GCE Qualifications: 1 GCSE in 1974, at A-C grade. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: No exams taken, no qualifications.

Case: 91
Chrtid: 592014V
NCDS VI GCSE/CSE/GCE Qualifications: 7 GCSEs in 1974, 5 at A-C grades, 2 at lower grades. No O levels or CSEs.
Present in NCDS V: No
Present in EXAMS file: No
EXAMS file GCE/CSE Qualifications: N/A.

Case: 92
Chrtid: 514058A
NCDS VI GCSE/CSE/GCE Qualifications: 4 GCSEs in 1974, all at A-C grades. No O levels or CSEs.
Present in NCDS V: Yes
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 CSE at grade 2, 2 CSEs at grade 3, 3 CSEs at grade 4.

Case: 93
Chrtid: 514071S
NCDS VI GCSE/CSE/GCE Qualifications: 6 GCSEs in 1974, all at A-C grades. 1 O level in 1974, no CSEs.
Present in NCDS V: No
Present in EXAMS file: No
EXAMS file GCE/CSE Qualifications: N/A

Case: 94
Chrtid: 503018A
NCDS VI GCSE/CSE/GCE Qualifications: 12 GCSEs in 1973, 6 at A-C grades, 6 at lower grades. No O levels and no CSEs.
Present in NCDS V: Yes
Qualifications in NCDS V: O levels at A-C grades.
Present in EXAMS file: Yes
EXAMS file GCE/CSE Qualifications: 1 O level/CSE grade 1.

BCS70

1. CHRTID/SERIAL: 17347090

BCS 2000 Qualifications: 6 GCSEs, all in 1986. No O levels or CSEs.
Present in BCS 1996: No

2. CHRTID/SERIAL: 17345113

BCS 2000 Qualifications: 5 GCSEs, all in 1986, 4 at grades A-C, 1 at a lower grade. No O levels or CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 1 O level at grade A-C; 7 GCSEs all at A-C grades.

3. CHRTID/SERIAL: 10482048

BCS 2000 Qualifications: 1 GCSE in 1987, grade A-C; 6 (or 8 as suggested by summary data) O levels in 1986; no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 8 O levels at grade A-C; also 1 O level at a lower grade. 1 GCSE.

4. CHRTID/SERIAL: 7246058

BCS 2000 Qualifications: 3 GCSEs in 1986, all at A-C grade. No O levels or CSEs.

Present in BCS 1996: No

5. CHRTID/SERIAL: 5078026

BCS 2000 Qualifications: 2 GCSEs in 1986, both at A-C grades; no O levels; 7 CSEs in 1986.

Present in BCS 1996: Yes.

BCS 1996 Qualifications: 1 O level at A-C grade; 2 other O levels at lower grades; 1 CSE grade 1; 4 CSEs at below grade 1. No GCSEs.

6. CHRTID/SERIAL: 3535025

BCS 2000 Qualifications: 6 GCSEs in 1986, 4 at A-C grade, 2 at lower grades. No O levels or CSEs.

Present in BCS 1996: No

7. CHRTID/SERIAL: 5110044

BCS 2000 Qualifications: 5 GCSEs in 1986, 3 at A-C grades, 2 at lower grades.

Present in BCS 1996: No

8. CHRTID/SERIAL: 5840013

BCS 2000 Qualifications: 10 GCSEs in 1985. 10 O levels – 1 in 1983, 1 in 1984, 2 in 1985, 6 in 1986. No CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 10 O levels, all at A-C grades, and 5 GCSEs at A-C grades.

9. CHRTID/SERIAL: 5045072

BCS 2000 Qualifications: 5 GCSEs in 1986. No O levels or CSEs.

Present in BCS 1996: No

10. CHRTID/SERIAL: 12072019

BCS 2000 Qualifications: 1 GCSE in 1987. 7 O levels - 3 in 1987 and 4 in 1986. 2 CSEs in 1986.

Present in BCS 1996: No

11. CHRTID/SERIAL: 6266032

BCS 2000 Qualifications: 6 GCSEs in 1986, 5 at A-C grades, 1 at lower grade. No O levels or CSEs.

Present in BCS 1996: Yes.

BCS 1996 Qualifications: 5 O levels at A-C grades; no GCSEs and no CSEs.

12. CHRTID/SERIAL: 6287060
BCS 2000 Qualifications: 4 GCSEs in 1986, all at A-C grades; no O levels or CSEs.
Present in BCS 1996: No
13. CHRTID/SERIAL: 21453000
BCS 2000 Qualifications: 5 GCSEs in 1987, 3 at A-C grades, 2 at lower grades. No O levels or CSEs.
Present in BCS 1996: No
14. CHRTID/SERIAL: 3159070
BCS 2000 Qualifications: 1 GCSE in 1986, at A-C grade. 9 O levels, 8 in 1986 and 1 in 1985, all at A-C grades. No CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 10 O levels at A-C grades; no GCSEs.
15. CHRTID/SERIAL: 4539004
BCS 2000 Qualifications: 1 GCSE in 1986, at A-C grade. 3 O levels in 1986, at A-C grades. 3 CSEs in 1986, 2 at grade 1, and 1 at lower grade.
Present in BCS 1996: No
16. CHRTID/SERIAL: 17549095
BCS 2000 Qualifications: 8 GCSEs in 1986, 7 at A-C grades and 1 at lower grade. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 7 O levels at grades A-C. No GCSEs.
17. CHRTID/SERIAL: 10418066
BCS 2000 Qualifications: 9 GCSEs, all at A-C grades, 8 obtained in 1986 and 1 in 1987. No O levels or CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 9 O levels, all at A-C grades, no CSEs or GCSEs.
18. CHRTID/SERIAL: 15428033
BCS 2000 Qualifications: 2 GCSEs, both at below grade C, in 1986. No O levels. 3 CSEs in 1986, 2 at grade 1, 1 at lower grade.
Present in BCS 1996: No
19. CHRTID/SERIAL: 4169024
BCS 2000 Qualifications: 3 GCSEs in 1987, all at A-C grades; 2 O levels in 1987, at A-C grades, no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 4 O levels, all at below grade C.
20. CHRTID/SERIAL: 4251023
BCS 2000 Qualifications: 1 GCSE in 1987, at A-C grade; 3 O levels in 1986, 2 at A-C grades, 1 at lower grade; 6 CSEs in 1986, 2 at grade 1, 4 at lower grades.
Present in BCS 1996: No
21. CHRTID/SERIAL: 16306031
BCS 2000 Qualifications: 2 GCSEs in 1986, 1 at A-C grades, 1 at a lower grade. 9 O levels in 1986, 7 at A-C grades, 2 at lower grades. No CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 8 O levels at A-C; coded as -3 for both O levels below grade C and GCSEs.

22. CHRTID/SERIAL: 921022
BCS 2000 Qualifications: 3 GCSEs in 1986, 2 at A-C grades, 1 at lower grade; no O levels; 2 CSEs in 1986, both at below grade 1.
Present in BCS 1996: No

23. CHRTID/SERIAL: 16442012
BCS 2000 Qualifications: 7 GCSEs in 1987, 3 at A-C grades, 4 lower grades; 8 O levels in 1987, 3 at A-C grades, 5 at lower grades; 7 CSEs in 1986, 3 at grade 1, 4 at lower grades.
Present in BCS 1996: Yes
BCS 1996 Qualifications: No CSEs or GCSEs, 2 O levels at A-C grades.

24. CHRTID/SERIAL: 11496003
BCS 2000 Qualifications: 2 GCSEs, 1 in 1986, 1 in 1987, 1 of these is at A-C grade, other at lower grade. No O levels; 4 CSEs in 1986, all at grade 1.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 3 CSEs at grade 1; 2 CSEs at lower grades; also 1 GCSE at grades A-C.

25. CHRTID/SERIAL: 12954012
BCS 2000 Qualifications: 2 GCSEs in 1986, both at A-C grades. 7 O levels, 2 in 1985, 5 in 1986. 5 of them at A-C grades, 2 at lower grades. No CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 7 O levels at A-C grades, 2 O levels at lower grades; no CSEs.

26. CHRTID/SERIAL: 6649014
BCS 2000 Qualifications: 11 GCSEs, all at A-C grade, 9 in 1986, 2 in 1987. No O levels, no CSEs.
Present in BCS 1996: No

27. CHRTID/SERIAL: 10911025
BCS 2000 Qualifications: 10 GCSEs in 1986, all at A-C grades. No O level, no CSE.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 9 O levels at A-C grade. 1 GCSE at A-C.

28. CHRTID/SERIAL: 13100091
BCS 2000 Qualifications: 5 GCSEs in 1986. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 5 O levels at A-C grades; no CSEs; no GCSEs.

29. CHRTID/SERIAL: 14836090
BCS 2000 Qualifications: 4 GCSEs in 1986, all at A-C grades; 1 O level in 1986 at A-C grade; 1 CSE in 1986 at grade 1.
Present in BCS 1996: No

30. CHRTID/SERIAL: 3845033
BCS 2000 Qualifications: 8 GCSEs in 1986, all at below grade C. No O levels. 3 CSEs in 1986, at below grade 1.
Present in BCS 1996: Yes
BCS 1996 Qualifications: Has ticked the no qualifications box.

31. CHRTID/SERIAL: 16997056
BCS 2000 Qualifications: 11 GCSEs in 1986, 10 at A-C grades, 1 at a lower grade. No O levels, no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: Also 10 GCSEs at A-C, and 1 at a lower grade. No O levels or CSEs.

32. CHRTID/SERIAL: 12429074
BCS 2000 Qualifications: 7 GCSEs in 1986, 6 at A-C grades, 1 at a lower grade. No O levels or CSEs.
Present in BCS 1996: Yes

BCS 1996 Qualifications: 2 CSEs at grade 1; 6 CSEs at other grades, 1 O level at A-C. No GCSEs.

33. CHRTID/SERIAL: 15685019

BCS 2000 Qualifications: 3 GCSEs in 1986, all at A-C grades. No O levels or CSEs.
Present in BCS 1996: No

34. CHRTID/SERIAL: 14449083

BCS 2000 Qualifications: 1 GCSE at A-C grade in 1986. No O levels or CSEs.
Present in BCS 1996: No

35. CHRTID/SERIAL: 16441037

BCS 2000 Qualifications: 8 GCSEs in 1986, 6 at A-C grades, 2 at lower grades. No O levels or CSEs.
Present in BCS 1996: No

36. CHRTID/SERIAL: 6990050

BCS 2000 Qualifications: 8 GCSEs in 1986, 6 at A-C grades, 2 at lower grades. 1 O level at A-C grade in 1987. No CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 8 CSEs all below grade 1; 5 O levels at A-C grades; no GCSEs.

37. CHRTID/SERIAL: 10291095

BCS 2000 Qualifications: 8 GCSEs, 7 of them obtained in 1986 and 1 in 1987, all of them were at A-C grades. 8 O levels, 7 in 1986 and 1 in 1987, all at A-C grades. No CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 8 O levels at A-C grades.

38. CHRTID/SERIAL: 12162020

BCS 2000 Qualifications: 7 GCSEs, 6 in 1986, and 1 in 1988, 4 of them at A-C grades and 3 at lower grades. No O levels or CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 3 O levels at A-C grades, 1 further O level at a lower grade; 5 CSEs all below grade 1. No GCSEs.

39. CHRTID/SERIAL: 20344000

BCS 2000 Qualifications: 6 GCSEs in 1986, 4 at A-C grades, 2 at lower grades. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 1 CSE grade 1, and 4 CSEs at lower grades; 1 O level at below grade C. No GCSEs.

40. CHRTID/SERIAL: 2796034

BCS 2000 Qualifications: 5 GCSEs, 3 in 1986 and 2 in 1988, all at A-C grades. 1 O level in 1987, at A-C grade. 6 CSEs in 1986, 1 at grade 1 and 5 at lower grades.
Present in BCS 1996: No

41. CHRTID/SERIAL: 11994039

BCS 2000 Qualifications: 8 GCSEs in 1986, 4 at A-C grades, and 4 at lower grades. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 8 CSEs, all at below grade 1; No O levels and no GCSEs.

42. CHRTID/SERIAL: 12873012

BCS 2000 Qualifications: 4 GCSEs in 1987, 3 at A-C and 1 at a lower grade. 2 O levels in 1988 both at A-C grades, no CSEs.
Present in BCS 1996: No

43. CHRTID/SERIAL: 14521254

BCS 2000 Qualifications: 5 GCSEs in 1986, all at below grade C; 1 O level in 1986 at A-C grade, no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 3 CSEs at grade 1; no O levels or GCSEs.

44. CHRTID/SERIAL: 5627033

BCS 2000 Qualifications: 3 GCSEs, in 1986, 2 at A-C grades and 1 at a lower grade. 5 O levels in 1986, all at A-C grades; 2 CSEs in 1986, 1 of them at grade 1 and 1 at a lower grade.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 5 O levels at A-C; 3 GCSEs at A-C 1 of them at A-C and 1 at a lower grade; 2 CSEs – 1 at grade 1, 1 at a lower grade.

45. CHRTID/SERIAL: 17770000

BCS 2000 Qualifications: 6 GCSEs in 1986, 5 at A-C grades and 1 at a lower grade. 1 O level in 1986 at below C grade; no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 5 CSEs at below grade 1, 1 O level below grade C; no GCSEs.

46. CHRTID/SERIAL: 6897000

BCS 2000 Qualifications: 8 GCSEs in 1986, 6 at A-C grades, 2 at lower grades; 8 O levels in 1986, 6 at A-C, 2 at lower grades. No CSEs.

Present in BCS 1996: No

47. CHRTID/SERIAL: 120029

BCS 2000 Qualifications: 2 GCSEs in 1986, both at A-C grades; 8 O levels, 6 in 1986, 2 in 1987, all at A-C grades. No CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 11 O levels at A-C grades; 1 GCSE at A-C grade.

48. CHRTID/SERIAL: 2036036

BCS 2000 Qualifications: 8 GCSEs in 1986, 7 at A-C grades, 1 at a lower grade. No O levels, 1 CSE at grade 1.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 1 CSE grade 1, 4 O levels at A-C, 2 GCSEs at A-C grade.

49. CHRTID/SERIAL: 5043021

BCS 2000 Qualifications: . 9 GCSEs, with 8 in 1986 and 1 in 1991, all at A-C grades. 1 O level in 1990 at A-C grade; no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 9 O levels at A-C grades, no GCSEs or CSEs.

50. CHRTID/SERIAL: 12989017

BCS 2000 Qualifications: 1 GCSE in 1987, at A-C grade; 8 O levels in 1986, 7 at A-C grades, 1 at a lower grade; 1 CSE in 1986 at grade 2 or lower.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 1 CSE at grade 2 or below; 7 O levels at A-C grades, 1 O level at a lower grade; 1 GCSE at A-C grade.

51. CHRTID/SERIAL: 11458049

BCS 2000 Qualifications: 4 GCSEs, all at A-C grades, 2 obtained in 1986 and 2 in 1987. No O levels. 4 CSEs in 1986, all at below grade 1.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 4 GCSEs at A-C grades; 5 GCSEs at other grades. Note left school and full time education at 17, so unlikely to have GCSEs.

52. CHRTID/SERIAL: 14811061
BCS 2000 Qualifications: 1 GCSE in 1987, at A-C grade. 10 O levels in 1985, 9 at A-C grades, 1 at a lower grade. No CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 9 O levels all at A-C grades, no GCSEs.

53. CHRTID/SERIAL: 15304002
BCS 2000 Qualifications: 3 GCSEs in 1987, all at A-C grades; 7 O levels in 1986, all at A-C grades; 4 CSEs in 1986, all at grade 1.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 2 CSEs at grade 1; 8 O levels at A-C grades; 2 GCSEs at A-C grades.

54. CHRTID/SERIAL: 13760011
BCS 2000 Qualifications: 5 GCSEs, all at A-C grades, 2 obtained in 1987 and 3 in 1986. No O levels or CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 7 CSEs – all at below grade 1; 5 O levels at A-C grades; no GCSEs.

55. CHRTID/SERIAL: 13970042
BCS 2000 Qualifications: 4 GCSEs in 1986, all at A-C grades; no O levels or CSEs.
Present in BCS 1996: No

56. CHRTID/SERIAL: 13624030
BCS 2000 Qualifications: 7 GCSEs in 1987, 2 at A-C grades and 5 at lower grades.
Present in BCS 1996: No

57. CHRTID/SERIAL: 11025086
BCS 2000 Qualifications: 3 GCSEs, 2 in 1987 and 1 in 1986, 2 at A-C grades, 1 at a lower grade.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 1 CSE at below grade 1; 4 O levels at A-C grades; 2 GCSEs at A-C grades.

58. CHRTID/SERIAL: 10086040
BCS 2000 Qualifications: 8 GCSEs, all at A-C grades, 1 in 1985 and 7 in 1986. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 8 O levels at A-C grades; no CSEs or GCSEs.

59. CHRTID/SERIAL: 5048098
BCS 2000 Qualifications: 9 GCSEs in 1986, all at A-C grades. No CSEs or O levels.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 9 O levels at A-C grades, no GCSEs.

60. CHRTID/SERIAL: 12324020
BCS 2000 Qualifications: 8 GCSEs (5 at A-C grade, 3 at lower grades), 5 of them in 1986, 2 in 1988 and 1 in 1993. No O levels or CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 4 CSEs: 1 at grade 1, 3 at lower grades; 1 O level at A-C grade; 3 GCSEs at A-C grades.

61. CHRTID/SERIAL: 16488069
BCS 2000 Qualifications: 7 GCSEs in 1986, 4 at A-C grades, 3 at lower grades. No O levels and no CSEs.
Present in BCS 1996: No

62. CHRTID/SERIAL: 1270066
BCS 2000 Qualifications: 5 GCSEs in 1985, all at A-C grades; 2 CSEs in 1986, both below grade 1; no O levels.

Present in BCS 1996: Yes
BCS 1996 Qualifications: 5 O levels at A-C grades; no GCSEs or CSEs.

63. CHRTID/SERIAL: 10165090
BCS 2000 Qualifications: 10 GCSEs in 1986, all at A-C grades; no O levels or CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 10 O levels at A-C grades, no CSEs or GCSEs.

64. CHRTID/SERIAL: 10263265
BCS 2000 Qualifications: 9 GCSEs (6 at A-C grades, 3 at lower grades), 8 obtained in 1986 and 1 in 1990.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 2 CSEs: 1 at grade 1, 1 at a lower grade; 4 O levels at A-C grades, 3 O levels at lower grades; 1 GCSE at A-C grade.

65. CHRTID/SERIAL: 2619073
BCS 2000 Qualifications: 7 GCSEs, all at A-C grades, 5 in 1986 and 2 in 1987. 1 O level, A-C grade, in 1985; 1 CSE in 1986 at below grade 1.
Present in BCS 1996: No

66. CHRTID/SERIAL: 11731001
BCS 2000 Qualifications: 9 GCSEs in 1986, 3 at A-C grades, 6 at lower grades.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 3 GCSEs at A-C grades and 6 at lower grades. Note age of leaving school and full-time education is recorded as 16.

67. CHRTID/SERIAL: 1552096
BCS 2000 Qualifications: 9 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications.

68. CHRTID/SERIAL: 3678033
BCS 2000 Qualifications: 3 GCSEs: 1 in 1986, 1 in 1987, 1 in 1989, all at A-C grades. 8 O levels, all A-C grades, 7 obtained in 1986 and 1 in 1987. No CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 10 O levels, all at below grade C; 1 GCSE at A-C grade.

69. CHRTID/SERIAL: 11878010
BCS 2000 Qualifications: 8 GCSEs (5 at A-C grades, 3 at lower grades): 4 in 1986, 2 in 1988, 1 in 1989, 1 in 1994. No O levels or CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications.

70. CHRTID/SERIAL: 6807013
BCS 2000 Qualifications: 4 GCSEs in 1986, 3 at A-C grades, 1 at a lower grade. No O levels or CSEs.
Present in BCS 1996: No

71. CHRTID/SERIAL: 8027004
BCS 2000 Qualifications: 4 GCSEs (3 at A-C grades, 1 lower grade), 1 in 1986, 1 in 1987, 2 in 1989; 6 O levels in 1986, 1 at A-C grade and 5 at lower grades; no CSEs.
Present in BCS 1996: No

72. CHRTID/SERIAL: 2137089
BCS 2000 Qualifications: 2 GCSEs, both at A-C grades, 1 in 1987 and 1 in 1988. 3 O levels (2 at A-C grades and 1 at lower grade), 2 in 1986 and 1 in 1987; 6 CSEs in 1986, all below grade 1.
Present in BCS 1996: Yes

BCS 1996 Qualifications: 6 CSEs, all below grade 1; 1 O level at A-C grade; 3 GCSEs at A-C grade.

73. CHRTID/SERIAL: 4197002

BCS 2000 Qualifications: 1 GCSE in 1986 at A-C grade; 6 O levels in 1986, 5 at A-C grades, and 1 at a lower grade; no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 7 O levels: 6 at A-C grades and 1 at a lower grade; 1 CSE at below grade 1.

74. CHRTID/SERIAL: 2772081

BCS 2000 Qualifications: 1 GCSE in 1986, at A-C grade. No O levels or CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at 16; -3 for CSEs at below grade 1 ??.

No other qualifications recorded.

75. CHRTID/SERIAL: 12690085

BCS 2000 Qualifications: 1 GCSE in 1986, at A-C grades; 4 O levels in 1996, 3 at A-C grades and 1 at a lower grade; 3 CSEs in 1986, 1 at grade 1, 2 at lower grades.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 2 CSEs: 1 at grade 1, 1 at a lower grade. 5 O levels: 4 of them at A-C grades, 1 at a lower grade. No GCSEs.

76. CHRTID/SERIAL:15338032

BCS 2000 Qualifications: 7 GCSEs in 1986, all at A-C grades; 1 O level in 1986, at A-C grade; no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 7 O levels at A-C grades; 1 CSE grade 1, no GCSEs.

77. CHRTID/SERIAL: 1714096

BCS 2000 Qualifications: 9 GCSEs, 8 in 1986, 1 in 1987; no O levels and no CSEs. Summary data says 10 GCSEs, all at A-C grades.

Present in BCS 1996: No

78. CHRTID/SERIAL: 9419066

BCS 2000 Qualifications: 6 GCSEs in 1986, 5 at A-C grades, 1 at lower grade. No O levels and no CSEs.

Present in BCS 1996: No

79. CHRTID/SERIAL: 14358006

BCS 2000 Qualifications: 8 GCSEs (6 at A-C grades, 2 at lower grades), 7 in 1986, 1 in 1987. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 6 GCSEs at A-C grades.

80. CHRTID/SERIAL: 5585063

BCS 2000 Qualifications: 3 GCSEs in 1986, all at A-C grades; no O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: (Left school and full-time education at 16, making it very unlikely that any GCSEs). 1 CSE at below grade 1.

81. CHRTID/SERIAL: 9222061

BCS 2000 Qualifications: 2 GCSEs in 1986, both at A-C grades. 2 O levels in 1986 at A-C grades; 1 CSE in 1986 at grade 1.

Present in BCS 1996: Yes

BCS 1996 Qualifications: (left school and full-time education in 1986). 3 O levels at A-C grades; 3 GCSEs at A-C grades.

82. CHRTID/SERIAL: 14487037
BCS 2000 Qualifications: 2 GCSEs in 1986, both at A-C grades; no O levels or CSEs.
Present in BCS 1996: No
83. CHRTID/SERIAL: 1504065
BCS 2000 Qualifications: 9 GCSEs in 1986, all at A-C grades. No O levels or CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications
84. CHRTID/SERIAL: 5348029
BCS 2000 Qualifications: 9 GCSEs, 8 in 1986 and 1 in 1994. Summary says 7 GCSEs, 7 at A-C and 2 at lower grades. No O levels or CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 1 CSE at below grade 1, left school and full-time education at 16.
85. CHRTID/SERIAL: 5793043
BCS 2000 Qualifications: 3 GCSEs in 1985, all at A-C grades. No O levels and no CSEs.
Present in BCS 1996: No
86. CHRTID/SERIAL: 9219011
BCS 2000 Qualifications: 9 GCSEs, (7 at A-C grades, 2 at lower grades), 7 in 1986, 2 in 1987; no O levels or CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications.
87. CHRTID/SERIAL: 1367094
BCS 2000 Qualifications: 5 GCSEs in 1986, 4 at A-C and 1 at lower grade. No O levels and no CSEs.
Present in BCS 1996: No
88. CHRTID/SERIAL: 12163096
BCS 2000 Qualifications: 10 GCSEs, all at A-C grades, 8 in 1986, 1 in 1987, 1 in 1988. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications
89. CHRTID/SERIAL: 14486062
BCS 2000 Qualifications: 6 GCSEs in 1986, 3 at A-C grades, and 3 at lower grades. 2 O levels in 1986, both at A-C grades; no CSEs.
Present in BCS 1996: No
90. CHRTID/SERIAL: 15934095
BCS 2000 Qualifications: 2 GCSEs in 1986, 1 at A-C grades and 1 at lower grades.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 2 O levels at A-C; no GCSEs.
91. CHRTID/SERIAL: 3935034
BCS 2000 Qualifications: 5 GCSEs, all at A-C grades, 4 in 1987 and 1 in 1989. 7 O levels in 1986, all at A-C grades; no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications
92. CHRTID/SERIAL: 15648040
BCS 2000 Qualifications: 7 GCSEs, all at A-C grades; 6 in 1986 and 1 in 1991. No O levels and no CSEs.
Present in BCS 1996: No

93. CHRTID/SERIAL: 10366019
BCS 2000 Qualifications: 6 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications.
94. CHRTID/SERIAL: 9232037
BCS 2000 Qualifications: 8 GCSEs in 1986, all at A-C grades.
Present in BCS 1996: No
95. CHRTID/SERIAL: 11337044
BCS 2000 Qualifications: 5 GCSEs in 1986, all at A-C grades. No O levels and no CSEs
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications. (Left school and full-time education at 17, before introduction of GCSEs).
96. CHRTID/SERIAL: 1412013
BCS 2000 Qualifications: 7 GCSEs: 5 in 1986 and 2 in 1987. Summary data states 8 GCSEs, 6 at A-C grades and 2 lower grades. No O levels and no CSEs
Present in BCS 1996: No
97. CHRTID/SERIAL: 9519043
BCS 2000 Qualifications: 1 GCSE in 1986, at A-C grade. No O levels and no CSEs
Present in BCS 1996: Yes
BCS 1996 Qualifications: 4 GCSEs at below grade C.
98. CHRTID/SERIAL:13705030
BCS 2000 Qualifications: 4 GCSEs in 1984, 1 at A-C grade, 3 at lower grades. No O levels and no CSEs
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications. Left school and full-time education at 16.
99. CHRTID/SERIAL: 9064062
BCS 2000 Qualifications: 7 GCSEs (6 at A-C grades, 1 at lower grade), 6 in 1986 and 1 in 1987. No O levels and no CSEs
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications.
100. CHRTID/SERIAL: 1701094
BCS 2000 Qualifications: 1 GCSE in 1986 at A-C grade; 4 O levels in 1986, 2 at A-C grades and 2 at lower grades. No CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications.
101. CHRTID/SERIAL: 12625027
BCS 2000 Qualifications: 1 GCSE in 1986, at A-C grade. 3 O levels in 1986, all at A-C grades; no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications.
102. CHRTID/SERIAL: 10168015
BCS 2000 Qualifications: 1 GCSE in 1986, at A-C grades; 3 O levels in 1986, all at A-C grades; 4 CSEs in 1986, all at grade 1.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 4 CSEs at grade 1; 4 CSEs at other grades; 3 O levels at A-C grades; 3 GCSEs, 2 at A-C grades, 1 at a lower grade.
103. CHRTID/SERIAL: 1674076

BCS 2000 Qualifications: 7 GCSEs in 1986, 4 at A-C grades, 3 at lower grades. No O levels and no CSEs.

Present in BCS 1996: No

104. CHRTID/SERIAL: 14718010

BCS 2000 Qualifications: 6 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.

Present in BCS 1996: No

105. CHRTID/SERIAL: 13305021

BCS 2000 Qualifications: 1 GCSE in 1987, at A-C grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 5 CSEs, 2 at grade 1, 3 at other grades; 7 O levels, 5 at A-C grades, 2 at other grades. 1 GCSE at A-C grade.

106. CHRTID/SERIAL: 13764012

BCS 2000 Qualifications: 4 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 3 CSEs, 1 at grade 1, 2 at other grades; 5 O levels, 2 at A-C and 3 at other grades. No GCSEs.

107. CHRTID/SERIAL: 16291064

BCS 2000 Qualifications: 2 GCSEs in 1987, both at A-C grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 2 GCSEs at A-C grades, no other qualifications.

108. CHRTID/SERIAL: 17341038

BCS 2000 Qualifications: 7 GCSEs in 1986, 3 at A-C grades; 4 at lower grades. No O levels and no CSEs.

Present in BCS 1996: No

109. CHRTID/SERIAL: 17437091

BCS 2000 Qualifications: 4 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 4 CSEs: 1 at grade 1, 3 at other grades. No GCSEs or O levels.

110. CHRTID/SERIAL: 12605075

BCS 2000 Qualifications: 1 GCSE in 1986, at A-C grade; 4 O levels in 1986, 1 at A-C grade, 3 at lower grades; 4 CSEs in 1986, all below grade 1.

Present in BCS 1996: Yes

BCS 1996 Qualifications: Left school and full-time education at 16; 5 CSEs all below grade 1; 3 O levels, 2 at A-C grades, 1 at lower grade. No GCSEs.

111. CHRTID/SERIAL: 5905062

BCS 2000 Qualifications: 2 GCSEs in 1986, both at A-C grades. No O levels, 4 CSEs in 1986, all at grade 1.

Present in BCS 1996: No

112. CHRTID/SERIAL: 11919005

BCS 2000 Qualifications: 10 GCSEs in 1986, 9 at A-C grades, 1 at a lower grade. 10 O levels in 1986, 9 at A-C and 1 at lower grade. No CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 8 O levels at A-C grade, 1 CSE grade 1. No GCSEs.

113. CHRTID/SERIAL: 13335050

BCS 2000 Qualifications: 6 GCSEs in 1986, 4 at A-C and 2 at lower grades. No O levels and no CSEs.

Present in BCS 1996: Yes
BCS 1996 Qualifications: Left school and full-time education at age 16; 5 GCSEs, 2 at A-C grade; 3 at lower grades; no O levels or CSEs.

114. CHRTID/SERIAL: 11676005
BCS 2000 Qualifications: 2 GCSEs in 1986, 1 at A-C and 1 at lower grade. 2 O levels in 1986, 1 at A-C and 1 at lower grade; 1 CSE in 1986 at grade 1.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 1 CSE grade 1; 3 O levels at A-C grades. No GCSEs.

115. CHRTID/SERIAL: 8604090
BCS 2000 Qualifications: 7 GCSEs in 1986, 5 at A-C grades, and 2 at lower grades. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: left school and full-timed education at 16, 3 CSEs at grade 1, 4 O levels, all below grade C.

116. CHRTID/SERIAL: 16639266
BCS 2000 Qualifications: 8 GCSEs in 1986, 2 at A-C grades, and 6 at lower grades. No O levels and no CSEs.
Present in BCS 1996: No

117. CHRTID/SERIAL: 14676040
BCS 2000 Qualifications: 1 GCSE in 1987, at A-C grade. 7 O levels in 1986, all at A-C grade; no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 8 O levels at A-C grades; 1 CSE at below grade 1.

118. CHRTID/SERIAL: 15215077
BCS 2000 Qualifications: 6 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: (left school and full-time education at age 16), 6 O levels at A-C grades, no GCSEs.

119. CHRTID/SERIAL: 262062
BCS 2000 Qualifications: 4 GCSEs in 1986, all at A-C grades.
Present in BCS 1996: Yes
BCS 1996 Qualifications: claims to have left school and full-time education at age 15, and has ticked the box for 'no qualifications'.

120. CHRTID/SERIAL: 12693010
BCS 2000 Qualifications: 1 GCSE in 1986, at A-C grade; 7 O levels in 1986, with 6 at A-C grades and 1 at a lower grade. No CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 6 O levels at A-C grades, 1 GCSE at A-C grade.

121. CHRTID/SERIAL: 10441068
BCS 2000 Qualifications: 4 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: left school and full-time education at age 16, 5 CSEs all at below grade 1, no GCSEs.

122. CHRTID/SERIAL: 15814065
BCS 2000 Qualifications: 6 GCSEs in 1986, 3 at A-C grades and 3 at lower grades. No O levels and no CSEs.
Present in BCS 1996: No

123. CHRTID/SERIAL: 16007075

BCS 2000 Qualifications: 2 GCSEs in 1986, both at A-C grades. No O levels, 1 CSE in 1986, at below grade 1.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 3 CSEs at below grade 1, no GCSEs.

124. CHRTID/SERIAL: 5773091

BCS 2000 Qualifications: 3 GCSEs in 1986, all at A-C grades. No O levels; 4 CSEs in 1986, 1 at grade 1, 3 at lower grades.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 7 GCSEs at below grade C; no O levels or CSEs.

125. CHRTID/SERIAL: 10481073

BCS 2000 Qualifications: 6 GCSEs in 1986; summary data suggests 9 GCSEs, 8 at A-C grades, 1 at a lower grade. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 8 O levels at A-C grades; no GCSEs.

126. CHRTID/SERIAL: 2832053

BCS 2000 Qualifications: 6 GCSEs in 1986; summary data suggests 9 GCSEs all at A-C grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 9 O levels at A-C grades; no GCSEs.

127. CHRTID/SERIAL: 14225075

BCS 2000 Qualifications: 1 GCSE in 1987, at below grade C. No O levels. 4 CSEs in 1986, all at grade 1.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 3 CSEs, 2 at grade 1, and 1 at a lower grade; 1 O level at A-C grade; 1 GCSE at below grade C.

128. CHRTID/SERIAL: 2816001

BCS 2000 Qualifications: 3 GCSEs in 1987, all at A-C grades. No O levels and no CSEs.

Present in BCS 1996: No

129. CHRTID/SERIAL: 7103050

BCS 2000 Qualifications: 8 GCSEs in 1986, 4 at A-C grades and 4 at lower grades. No O levels and no CSEs.

Present in BCS 1996: No

130. CHRTID/SERIAL: 3522023

BCS 2000 Qualifications: 1 GCSE in 1986 at below grade C; 6 O levels in 1986, 2 at A-C grades and 4 at lower grades. 7 CSEs in 1986, 2 at grade 1 and 5 at lower grades.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 7 CSEs, 2 at grade 1 and 5 at lower grades; 6 O levels, 2 at A-C grades and 4 at lower grades. No GCSEs.

131. CHRTID/SERIAL: 3524074

BCS 2000 Qualifications: 4 GCSEs in 1986, 3 at A-C grades and 1 at a lower grade. No O levels and no CSEs.

Present in BCS 1996: No

132. CHRTID/SERIAL: 3323044

BCS 2000 Qualifications: 8 GCSEs in 1986, 1 at A-C grades and 7 at lower grades.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16; 7 CSEs all below grade 1; no O levels or GCSEs.

133. CHRTID/SERIAL: 160034

BCS 2000 Qualifications: 5 GCSEs in 1986, 5 at A-C grades and 1 at a lower grade. 2 O levels in 1986, both at A-C grades; no CSEs.

Present in BCS 1996: No

134. CHRTID/SERIAL: 2740001

BCS 2000 Qualifications: 6 GCSEs in 1986, all at A-C grades. 1 O level in 1986, at A-C grades, no CSEs.

Present in BCS 1996: No

135. CHRTID/SERIAL: 15240029

BCS 2000 Qualifications: 8 GCSEs, 4 in 1986 and 4 in 1987, all at A-C grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 10 O levels, 8 at A-C grades and 2 at lower grades; 4 CSEs, 3 at grade 1 and 1 at a lower grade; no GCSEs.

136. CHRTID/SERIAL: 3463026

BCS 2000 Qualifications: 3 GCSEs in 1986, 2 at A-C grades and 1 at a lower grade. 5 O levels in 1986, all at A-C grades; 4 CSEs in 1986, 3 at grade 1 and 1 at a lower grade.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 5 O levels at A-C grade; 4 CSEs, 3 at grade 1 and 1 at a lower grade; 3 GCSEs, 2 at A-C grades and 1 at a lower grade.

137. CHRTID/SERIAL: 8722069

BCS 2000 Qualifications: 8 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.

Present in BCS 1996: No

138. CHRTID/SERIAL: 59058

BCS 2000 Qualifications: 6 GCSEs (2 at A-C grade and 4 at lower grades), 5 in 1986 and 1 in 1987.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 7 CSEs, 2 at grade 1 and 5 at other grades; 6 O levels, 3 at A-C grades and 3 at other grades; 1 GCSE at A-C grade.

139. CHRTID/SERIAL: 6565089

BCS 2000 Qualifications: 2 GCSEs in 1986, both at A-C grades; no O levels; 3 CSEs in 1986, 1 at grade 1 and 2 at other grades.

Present in BCS 1996: No

140. CHRTID/SERIAL: 83035

BCS 2000 Qualifications: 6 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at 16; 6 CSEs, with 2 at grade 1 and 4 at other grades; no O levels and no GCSEs.

141. CHRTID/SERIAL: 6912091

BCS 2000 Qualifications: 9 GCSEs in 1986; 4 at A-C grades and 5 at other grades; No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 3 CSEs, 1 at grade 1 and 2 at lower grades; 4 O levels at A-C grades; 1 GCSE at A-C grade.

142. CHRTID/SERIAL: 10235038

BCS 2000 Qualifications: 6 GCSEs in 1986, 5 at A-C grades and 1 at a lower grade. No O levels; 5 CSEs, all at grade 1, 2 in 1991 and 3 in 1992.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 8 CSEs, 1 at grade 1; 7 at lower grades; 5 GCSEs, 3 at A-C grades and 2 at lower grades.

143. CHRTID/SERIAL: 11719051

BCS 2000 Qualifications: 6 GCSEs, 3 in 1986, 1 in 1992, 2 in 1993; summary data suggests 9 GCSEs, 2 at A-C grades and 7 at lower grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 6 CSEs, all below grade 1; 2 O levels at A-C grades; 4 GCSEs, 2 at A-C grades and 2 at lower grades.

144. CHRTID/SERIAL: 9515042

BCS 2000 Qualifications: 6 GCSEs in 1986, all at A-C grades.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 1 CSE at grade 1; 5 O levels at A-C grades; 4 GCSEs at below grade C.

145. CHRTID/SERIAL: 10695004

BCS 2000 Qualifications: 1 GCSE in 1986, at below grade C; 6 O levels in 1986, 4 at A-C grades, 2 at lower grades; 2 CSEs in 1986, 1 at grade 1, and 1 at other grade.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16; 2 CSEs, 1 at grade and 1 at lower grade; 5 O levels, 4 at A-C grade, 1 at other grade; 1 GCSE at below grade C.

146. CHRTID/SERIAL: 8599075

BCS 2000 Qualifications: 4 GCSEs in 1987, all at A-C grades; 2 O levels in 1986, 1 at A-C grades, and 1 at lower grade; 5 CSEs in 1986, 2 at grade 1 and 3 at lower grades.

Present in BCS 1996: No

147. CHRTID/SERIAL: 9560047

BCS 2000 Qualifications: 9 GCSEs, 7 in 1986 and 2 in 1987; 6 at A-C grades and 3 at lower grades. No O levels and no CSEs.

Present in BCS 1996: No

148. CHRTID/SERIAL: 2988063

BCS 2000 Qualifications: 7 GCSEs in 1986, all at below grade C. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 7 CSEs, all at below grade 1. No O levels and no GCSEs.

149. CHRTID/SERIAL: 11925081

BCS 2000 Qualifications: 8 GCSEs in 1986, 6 at A-C grades and 2 at other grades.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 5 O levels at A-C grades, and 2 CSEs at grade 1. No GCSEs.

150. CHRTID/SERIAL: 189064

BCS 2000 Qualifications: 9 GCSEs in 1986, 4 at A-C grades and 5 at lower grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 6 CSEs, 3 at grade 1 and 3 at lower grades; 1 O level at A-C grade; no GCSEs.

151. CHRTID/SERIAL: 11862083

BCS 2000 Qualifications: 8 GCSEs in 1986, 4 at A-C grades and 4 at lower grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16, 5 CSEs, 1 at grade 1 and 4 at lower grades; 5 O levels, 1 at A-C and 4 at lower grades; no GCSEs.

152. CHRTID/SERIAL: 16526038

BCS 2000 Qualifications: 2 GCSEs in 1986, both at A-C grades. No O levels and no CSEs.

Present in BCS 1996: No

153. CHRTID/SERIAL: 6135051

BCS 2000 Qualifications: 9 GCSEs in 1986, 5 at A-C grades and 4 at other grades.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 9 O levels, 5 at A-C grades, and 4 at other grades. No GCSEs.

154. CHRTID/SERIAL: 12418022

BCS 2000 Qualifications: 9 GCSEs in 1986, 8 at A-C grades and 1 at other grade.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16; 7 CSEs, 1 at grade 1 and 6 at lower grades; 1 GCSE at A-C grade.

155. CHRTID/SERIAL: 2961084

BCS 2000 Qualifications: 5 GCSEs in 1986, all at A-C grades; 7 O levels in 1985, 2 at A-C grades and 5 at lower grades. No CSEs.

Present in BCS 1996: No

156. CHRTID/SERIAL: 3563003

BCS 2000 Qualifications: 5 GCSEs in 1986, all at A-C grades; No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16; 8 O levels all at A-C grades; No GCSEs or CSEs.

157. CHRTID/SERIAL: 11633075

BCS 2000 Qualifications: 7 GCSEs (5 at A-C grades and 2 at lower grades), 5 in 1986 and 2 in 1989. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16, 2 CSEs at below grade 1; 7 O levels, 2 at A-C grades and 5 at lower grades. No GCSEs.

158. CHRTID/SERIAL: 12218068

BCS 2000 Qualifications: 9 GCSEs in 1986, 8 at A-C grade and 1 at lower grade. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 8 O levels at A-C grades; no GCSEs or CSEs.

159. CHRTID/SERIAL: 2985037

BCS 2000 Qualifications: 5 GCSEs in 1986, 1 at A-C grade and 4 at lower grades. 2 O levels in 1986, both at A-C grades, 1 CSE in 1986 at grade 1.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at 17; 1 CSE grade 1; 6 O levels, 3 at A-C, 3 at other grades.

160. CHRTID/SERIAL: 2986012

BCS 2000 Qualifications: 5 GCSEs in 1987, all at below C grade; 1 O level in 1987, at A-C grade; 8 CSEs all at below grade 1, 5 in 1986 and 3 in 1987.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 7 CSEs at below grade 1; 4 O levels, 1 at grade A-C, 3 at below C grade; no GCSEs.

161. CHRTID/SERIAL: 3406095

BCS 2000 Qualifications: 3 GCSEs in 1986, all at A-C grades; 4 CSEs in 1986, all below grade 1; No O levels.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16; 5 CSEs at below grade 1, no GCSEs.

162. CHRTID/SERIAL: 4233021

BCS 2000 Qualifications: 6 GCSEs in 1986, 3 at A-C and 3 at other grades. No O levels and no CSEs.

Present in BCS 1996: No

163. CHRTID/SERIAL: 8911072

BCS 2000 Qualifications: 5 GCSEs in 1986, 4 at A-C grades and 1 at other grade; no O levels; 1 CSE in 1986, at below grade 1.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 3 CSEs at below grade 1; 4 GCSEs at A-C grades.

164. CHRTID/SERIAL: 7465090

BCS 2000 Qualifications: 5 GCSEs in 1986, 1 at A-C grade and 4 at other grades. 5 O levels in 1986, 1 at A-C grade and 4 at lower grades; 3 CSEs in 1986, 1 at grade 1 and 2 at lower grades.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16; no information on qualifications.

165. CHRTID/SERIAL: 8935025

BCS 2000 Qualifications: 4 GCSEs in 1986, 2 at A-C grades and 2 at lower grades; No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 3 CSEs all at below grade 1; 4 O levels, 1 at A-C grade, 3 at lower grades; no GCSEs.

166. CHRTID/SERIAL: 4482082

BCS 2000 Qualifications: 2 GCSEs, both at A-C grade, 1 in 1987 and 1 in 1989. 3 O levels in 1986, all at A-C grades; 5 CSEs in 1986, 2 at grade 1 and 3 at lower grades.

Present in BCS 1996: No

167. CHRTID/SERIAL: 9735049

BCS 2000 Qualifications: 2 GCSEs in 1986, both at A-C grades; No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16; No O levels and no CSEs; 2 GCSEs at A-C grades.

168. CHRTID/SERIAL: 15607060

BCS 2000 Qualifications: 5 GCSEs in 1987, 3 at A-C grades and 2 at lower grades; 2 O levels in 1986, both at A-C grades; 1 CSE in 1986 at grade 1.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 4 CSEs at below grade 1; 4 O levels, 1 at A-C grade and 3 at other grades; 1 GCSE at A-C grade.

169. CHRTID/SERIAL: 10367095

BCS 2000 Qualifications: 3 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and f/t education at age 16; no information on qualifications.

170. CHRTID/SERIAL: 13324099

BCS 2000 Qualifications: 7 GCSEs in 1986, 4 at A-C grades and 3 at lower grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16; 2 CSEs at below grade 1; 4 O levels at A-C grades; 1 GCSE at below grade C.

171. CHRTID/SERIAL: 10214010

BCS 2000 Qualifications: 8 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 8 O levels at A-C grades; no CSEs and no GCSEs.

172. CHRTID/SERIAL: 6443008

BCS 2000 Qualifications: 11 GCSEs (5 at A-C grades and 6 at other grades), 7 in 1986, 2 in 1987, 2 in 1991.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16; 7 CSEs, 1 at grade 1, and 6 at lower grades; 2 O levels at A-C; 1 GCSE at A-C grade.

173. CHRTID/SERIAL: 10583000

BCS 2000 Qualifications: 7 GCSEs in 1986, 4 at A-C grades and 3 at lower grades; no O levels; 3 CSEs in 1986, all at below grade 1.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 3 CSEs, 1 at grade 1, 2 at lower grades; 3 O levels at A-C grades. No GCSEs.

174. CHRTID/SERIAL: 10287070

BCS 2000 Qualifications: 10 GCSEs in 1986, all at A-C grades; no O levels; 5 CSEs in 1986, 2 at grade 1 and 3 at lower grades.

Present in BCS 1996: Yes

BCS 1996 Qualifications: claims to have left school and full-time education at 15; no information on qualifications.

175. CHRTID/SERIAL: 9721072

BCS 2000 Qualifications: 6 GCSEs in 1986, 4 at A-C grades and 2 at lower grades. No O levels and no CSEs.

Present in BCS 1996: No

176. CHRTID/SERIAL: 9459071

BCS 2000 Qualifications: 3 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: (left school and f/t education at 16), 4 CSEs at below grade 1; no O levels or GCSEs.

177. CHRTID/SERIAL: 170010

BCS 2000 Qualifications: 3 GCSEs in 1987, 1 at A-C grade, 2 at lower grades. No O levels and no CSEs.

Present in BCS 1996: No

178. CHRTID/SERIAL: 5701006

BCS 2000 Qualifications: 2 GCSEs in 1986, both at A-C grades; no O levels; 4 CSEs in 1986, 1 at grade 1 and 3 at lower grades.

Present in BCS 1996: No

179. CHRTID/SERIAL: 174011

BCS 2000 Qualifications: 5 GCSEs in 1986, 3 at A-C grades and 2 at lower grades.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16; 3 CSEs at below grade 1; no O levels or GCSEs.

180. CHRTID/SERIAL: 10778055

BCS 2000 Qualifications: 2 GCSEs in 1987, both at A-C grades. No O levels, 6 CSEs in 1986, all at below grade 1.

Present in BCS 1996: No

181. CHRTID/SERIAL: 8561044
BCS 2000 Qualifications: 5 GCSEs in 1986, 2 at A-C grades and 3 at lower grades. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: left school and full-time education at age 16. 5 CSEs at below grade 1; 2 O levels at below grade C. No GCSEs.

182. CHRTID/SERIAL: 8471043
BCS 2000 Qualifications: 5 GCSEs in 1986, 4 at A-C grades, 1 at a lower grade. No O levels and no CSEs.
Present in BCS 1996: No

183. CHRTID/SERIAL: 5533031
BCS 2000 Qualifications: 5 GCSEs in 1986, 1 at A-C grade, 4 at lower grades; 1 O level in 1986 at A-C grade; 1 CSE in 1986 at grade 1.
Present in BCS 1996: No

184. CHRTID/SERIAL: 6072053
BCS 2000 Qualifications: 2 GCSEs, both below grade C, 1 in 1986, 1 in 1987. 5 O levels in 1986, 3 at A-C grades, 2 at lower grades; no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 4 CSEs at below grade 1; 6 O levels, 4 at A-C, 2 at other grades; no GCSEs.

185. CHRTID/SERIAL: 12158096
BCS 2000 Qualifications: 1 GCSE in 1986, at A-C grades; 5 O levels in 1986, 2 at A-C grades and 3 at lower grades; no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: No information on qualifications.

186. CHRTID/SERIAL: 8701041
BCS 2000 Qualifications: 10 GCSEs in 1986, 6 at A-C grades and 4 at lower grades.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 5 CSEs, 1 at grade 1 and 4 at other grades; 3 O levels at A-C grades; no GCSEs.

187. CHRTID/SERIAL: 4493033
BCS 2000 Qualifications: 5 GCSEs, all at A-C grades, 2 in 1987 and 3 in 1988. 2 O levels in 1986, both at A-C grades; no CSEs.
Present in BCS 1996: No

188. CHRTID/SERIAL: 4529028
BCS 2000 Qualifications: 8 GCSEs in 1986, 5 at A-C grades and 3 at lower grades. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: left school and full-time education at 16; no information on qualifications.

189. CHRTID/SERIAL: 4532078
BCS 2000 Qualifications: 8 GCSEs in 1986, 3 at A-C and 5 at other grades. No O levels and no CSEs.
Present in BCS 1996: Yes
BCS 1996 Qualifications: 4 CSEs, 1 at grade 1 and 3 at other grades; 1 O level at A-C grade; no GCSEs.

190. CHRTID/SERIAL: 13050095
BCS 2000 Qualifications: 4 GCSEs in 1986, all at below C grade; No O levels and no CSEs.
Present in BCS 1996: Yes

BCS 1996 Qualifications: 5 O levels at below grade C; GCSE below grade C coded as –3.

191. CHRTID/SERIAL: 1443017

BCS 2000 Qualifications: 3 GCSEs in 1986, all at A-C grades; 4 O levels in 1986, 2 at A-C grades and 2 at other grades.

Present in BCS 1996: No

192. CHRTID/SERIAL: 12326071

BCS 2000 Qualifications: 9 GCSEs in 1986, 7 at A-C grades, 2 at other grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 10 O levels, 7 at A-C grades, 3 at lower grades; no GCSEs.

193. CHRTID/SERIAL: 1437042

BCS 2000 Qualifications: 6 GCSEs in 1986, all at A-C grades. No O levels and no CSEs.

Present in BCS 1996: No

194. CHRTID/SERIAL: 1639047

BCS 2000 Qualifications: 7 GCSEs in 1986, 1 at A-C, 6 at other grades; No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 8 GCSEs at A-C grades; No O levels and no CSEs.

195. CHRTID/SERIAL: 10678078

BCS 2000 Qualifications: 7 GCSEs in 1986, 5 at A-C and 2 at other grades. No O levels. 2 CSEs in 1986, both at grade 1.

Present in BCS 1996: No

196. CHRTID/SERIAL: 2567026

BCS 2000 Qualifications: 6 GCSEs in 1986, 4 at A-C grades and 2 at other grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 6 GCSEs at below grade C, No O levels and no CSEs.

197. CHRTID/SERIAL: 13648084

BCS 2000 Qualifications: 2 GCSEs in 1985, both at A-C grades. No O levels and no CSEs.

Present in BCS 1996: No

198. CHRTID/SERIAL: 5609031

BCS 2000 Qualifications: 5 GCSEs in 1986, 3 at A-C and 2 at other grades. No O levels; 3 CSEs in 1986, all at below grade 1.

Present in BCS 1996: Yes

BCS 1996 Qualifications: left school and full-time education at age 16; 4 CSEs, all below grade 1. No GCSEs.

199. CHRTID/SERIAL: 12041190

BCS 2000 Qualifications: 6 GCSEs in 1986, 4 at A-C grades and 2 at other grades. No O levels and no CSEs.

Present in BCS 1996: Yes

BCS 1996 Qualifications: 5 CSEs, 2 at grade 1 and 3 at other grades; 3 O levels, 2 at A-C grades and 1 at other grade; 1 GCSE at A-C grade.

200. CHRTID/SERIAL: 509063

BCS 2000 Qualifications: 4 GCSEs in 1986, all at A-C grades; No O levels and no CSEs.

Present in BCS 1996: No

Report on GCSE date errors, Appendix 1. (Andrew Jenkins)

From: Andrew Jenkins [mailto:A.Jenkins@sta02.ioe.ac.uk]
Sent: Thursday, May 24, 2001 5:15 PM
To: Peter Shepherd (E-mail)
Subject: cleaning up ncds

Peter,

Here is a word file containing the actions which I've taken to produce an NCDS 6 dataset which is clean of erroneous dates for GCSEs. All those reporting pre-1988 GCSEs have been altered in the light of the information contained in NCDS 5 and the Exams file. I've tried to adopt a sensible solution for each case.

In the few cases where there is literally no data, we have simply recoded to CSEs.

I've done this by going into the dataset and changing the variables that need to be changed.

I can e-mail you the altered version of NCDS 6 education data in an spss file, or just the 94 cases which have been changed if that would be useful.

regards

Andrew

Dr Andrew Jenkins
Research Officer, Centre for the Economics of Education
Mathematical Sciences Group
Institute of Education
University of London
20 Bedford Way
London
WC1H 0AL

email a.jenkins@ioe.ac.uk
Tel 020 7612 6663

APPENDIX 7: Highest Qualification

Documentation relating to highest qualification produced by the following is reproduced below:

1. Andrew Jenkins (Institute of Education) and Gerald Makepeace (University of Cardiff)
2. Samantha Parsons, Centre for Longitudinal Studies

Queries about the contents of this appendix should be directed to the User Support Group (cohort@cls.ioe.ac.uk)

1. Andrew Jenkins and Gerald Makepeace

The table below shows how educational qualifications obtained by NCDS cohort members have been mapped into the national qualifications framework.

Level of Qualification	General (Academic)	Vocationally-related (Applied)	Occupational (Vocational)
5	Higher Degree		NVQ level 5 PGCE
4	Degree HE Diploma	BTEC Higher Certificate/Diploma HNC/HND	NVQ level 4 Professional degree level qualifications Nursing/paramedic Other teacher training qualification City & Guilds Part 4/Career Ext/Full Tech RSA Higher Diploma
3	A level AS levels Scottish Highers Scottish Cert of 6 th Year Studies	Advanced GNVQ BTEC National Diploma ONC/OND	NVQ level 3 City & Guilds Part 3/Final/Advanced Craft RSA Advanced Diploma Pitmans level 3
2	GCSE grade A*-C O levels grade A-C O levels grade D-E CSE grade 1 Scottish standard grades 1-3 Scottish lower or ordinary grades	Intermediate GNVQ BTEC First Certificate BTEC First Diploma	NVQ level 2 Apprenticeships City & Guilds Part 2/Craft/Intermediate City & Guilds Part 1/Other RSA First Diploma Pitmans level 2
1	GCSE grade D-G CSEs grades 2-5 Scottish standard grades 4-5 Other Scottish school qualification	Foundation GNVQ Other GNVQ	NVQ level 1 Other NVQ Units towards NVQ RSA Cert/Other Pitmans level 1 Other vocational qualifications HGV

SPSS syntax

***** DESCRIPTION OF PROGRAM *****

- * This program derives academic and qualifications using the 2000 data only.
- * It defines a set of binary variables showing whether an individual has a particular qualification.
- * These variables could then be used to devise various typologies for training .
- * We also derive the highest qualification below.
- * All variables based on data from the 2000 surveys end in 00.
- * The interpretation of the results differs across observations.

***** DEFINITION OF VARIABLES *****

- * There are 2 relevant questions.
- * The first asks 'Have you obtained any of the qualifications on this card since xx'.
- * If the respondent has the qualification, the second question asks 'How many did you get?'
eg 'How many GCE O-levels did you get ..?' 'How many City and Guilds quals have you obtained ..'
- * Where possible, we have defined the variables using the responses to the questions asking how many qualifications of a particular type the respondent has.
- * Often the respondent says none. We could interpret this as follows. The respondent said 'I have the qualification' when asked the first question but on reflection realised that they did not have it. This interpretation is strengthened by the fact that the information on all types of qualification is collected first.
- * Suppose the Cohort Member (CM) says I have a GCSE, a GCE and a CSE.
- * When the interviewer goes through the number of qualifications (one qualification at a time) the CM realises that the GCSE was GCE and one of the CCEs was a GCE.
- * The CM may know that he/she has a qualification but does not know how many O-levels, CSEs etc he/she has. The CM may respond 98, 99 but 0 seems unlikely if they have the qualification.
- * 0 is a genuine response. The feeder question for certain qualifications refers to all grades while the 'how many questions' refer to different ranges of grades. Consider somebody who has 8 GCSEs all at grade A. They will say they have a gcse in response to the feeder question, but then answer 8 for the number of grade A-Cs (edgcse1) and 0 to for the number of grade D-Es (edgcse2) .
- * The code below is easily amended to reflect the programmer's preferences.
- * Typically the response 'I have none of those qualification' (number equals 0) and the missing values for number of qualifications (8,9,98,99) have to be changed.
- * We have interpreted 8 and 98 as 'Do not know' and 9 and 99 as 'Not Applicable'.
- * We have set 8 and 98 equal to 1 (ie denoting possession of the qualification) suggesting that these values reflect an inability to remember the exact number of qualifications obtained.
- * We have set 9 and 99 equal to 0 suggesting that the respondent has considered the case further and decided that he/she does not have the qualification.
- * The missing value dummies set 9 for edquals2 and vocqual equal to 0.
- * Not Applicable is an odd response but it seems very close to saying 'No' I do not have any of the quals.

***** INTERPRETATION OF VARIABLES *****

- * The BCS70 qualifications are collected at ages 26 and 30. The age 16 questions refer to future examinations (ie the survey took place before the exams were taken).

- * This program uses the age 30 data (from the 2000 survey). The routing questions ask 'Have you obtained any of the qualifications on this card since April 1986'. Since most school exams were taken in the summer, it is therefore a summary of 'have you ever had this qualification'. It could perhaps be improved by comparing it to the age 26 data but there are differences in the form of the question and the method of collecting the data.
- * The age 26 survey was a postal survey with a smaller sample than the age 30 sample and there is less detail in the questions asked at 26.
- * NCDS has regularly collected information on qualifications. The questions in 2000 asked 'whether the qualification has been obtained since the age of 33' for individuals in NCDS V and 'whether the qualification has been obtained since the age of 23' for individuals not in NCDS V.
- * It is therefore a good summary of 'have you got this qualification since you were x years old'.

***** EXAMPLE AND AN IMPLICATION OF OUR APPROACH *****

- * Consider the following code for the variable odeg00=1 if he/she has any other degree level qualifications .
- * We generate this from the answer to 'how many other degree level qualifications do you have?' (numothdg) using the following code.
- * recode numothdg (missing=0) (0=0) (98=1)(99=0) (1 thru hi=1) into odeg00 .
- * variable labels odeg00 "Other degree level quals such as graduate membership of professional institutes".
- * A crosstabs of numothdg by odeg00 shows that there 99 zeros for numothdeg that give odeg00=0 and 1 value of 98 that gives odeg00=1.
- * We could have defined the alternative measure, ALTodeg, based on the question 'Do you have any of the following qualifications?' If the individual has 'other degree level qualifications', these are recorded as the value 9 in one of the variables edqtyp14 to edqtyp26. The code for this variable is below.
- * compute ALTodeg=0 .
- * if (any(9,edqtyp14,edqtyp15,edqtyp16,edqtyp17,edqtyp18,edqtyp19,edqtyp20,edqtyp21,edqtyp22,edqtyp23,edqtyp24,edqtyp25,edqtyp26)) ALTodeg=1 .
- * A crosstabs of ALTodeg by odeg00 shows that there are 99 respondents with ALTodeg not equal odeg00 corresponding to the 99 zeros for numothdeg . This is a large number since there are 963 observations with odeg00=1.
- * Many people have used the definition similar to ALTodeg rather than the present one.

***** MISSING VALUES *****

- * We begin by giving missing values the value 0. We later adjust the data for missing values.
- * Some analysis below shows that there are few genuinely missing values. The feed questions are 'Have you obtained any of the qualifications on this card?' with separate cards for educational and vocational questions. The answers to these questions are contained in edquals2 and vocqual.
- * (Be careful as there are variables with similar names for the proxy returns.) .
- * The possible answers to the feed questions are Yes, No, 8, 9, missing.
- * We treat an answer of Yes or No as saying that we have information on whether the interviewee has one or more of the qualifications.
- * The same 66 observations have missing values for both the educational and vocational questions.
- * A further 50 of the educational and 44 of the vocational answers are 8 or 9.

- * There are 28 values of 8 or 9 for both the educational and vocational questions.
- * The variables allmiss, acamiss and vocmiss capture the missing answers.
- * No respondents have 8 or 9 for the feeder question but subsequently report a qualification.
- * (Based on edqtyp14 and voctyp12 only).

***** PROXY INTERVIEWS *****

- * We have not picked up the observations (perhaps 26) given for proxy interviews.
- * Proxy returns occur when someone completed the form on behalf of the CM.
- * The variable identifying a proxy interview is IntWho in the Household Grid and Housing but we did not extract that information and it is not used below.
- * A measure (proxy=1) of whether the form is completed by a proxy is computed below .
- * From the data available to us at present, there are 26 interviews completed by proxy.
- * Be careful because many of the variables only refer to proxy interview
eg agelfte, edquals and vocquals are for proxies,
agelfte2, edquals2 and vocqual are for interviews completed by CMs.
- * edqtyp01 to edqtyp13, voctyp01 to voctyp11 are completed by proxies.
- * edqtyp14 to edqtyp26, voctyp12 to voctyp22 are completed by CMs .

***** GCSE *****

- * An apparently large proportion of the BCS70 cohort claim to have done GCSEs. They were one of the last cohorts to do GCE/CSE at school. This program does not resolve the issues raised by this although it seems clear that some respondents have not identified the exams that they took properly.

***** ADAPTING THE CODE *****

- * The code be could be simplified by merging the recode and other commands. For example, the opening code below could be written as .
- * recode edgcse1 edgcse2 (missing=0) (0=0) (98=1) (99=0) (1 thru hi =1) into gcsehi00 gcse00 .
- * variable labels csehi00 "Has CM obtained one or more CSEs at grade 1?"
- * / cse00 "Has CM obtained one or more CSEs at grades 2 to 5?".
- * (Be careful. Some variables use 9 as a missing value but 9 is a legitimate value for other variables.) .
- * The program is divided into sections because the PC on which it was developed could not handle the whole lot in one go.

***** PROGRAM STARTS *****

***** LOAD DATA *****

```
get file = 'e:\00\ncds-bcs70\qca data\data1.sav' .
*ghm at cls get file = 'c:\work\ghmtemp\ncds data\data1.sav' .
*pjd at cls get file = 'd:\datawork\qca data analysis\data1.sav' .
```

***** GCSE, O-LEVELS, CSE *****

```
recode edgcse1 (missing=0) (0=0) (98=1) (99=0) (1 thru hi =1) into gcsehi00 .
recode edgcse2 (missing=0) (0=0) (98=1) (99=0) (1 thru hi =1) into gcse00 .
```

* Make sure 98, 99 are before (1 thru hi) if you want them to be 0 or else they will be coded 1.
* In these cases, hi is 13 and 9 respectively.

variable labels gcsehi00 "Has CM obtained one or more GCSEs at grades A to C ?".
variable labels gcse00 "Has CM obtained one or more GCSEs at grades D to E ?".

recode edolev1 (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into olevhi00 .
recode edolev2 (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into olevlo00 .
* Hi is 15 and 11.

variable labels olevhi00 "Has CM obtained one or more O levels at grades A to C?".
variable labels olevlo00 "Has CM obtained one or more O levels at grades D to E?".

recode edcse1 (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into csehi00 .
recode edcse2 (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into cselo00 .
* Hi is 12 and 11.

variable labels csehi00 "Has CM obtained one or more CSEs at grade 1?".
variable labels cselo00 "Has CM obtained one or more CSEs at grades 2 to 5?".

***** AS AND A-LEVELS *****
recode edasl1 (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into aslhi00 .
recode edasl2 (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into asllo00 .
* Hi is 9 and 7.

variable labels aslhi00 "Has CM obtained one or more AS levels at grades A to C?".
variable labels asllo00 "Has CM obtained one or more AS levels at grades D to G?".

compute Naslev00=edasl1 .
recode Naslev00 (missing=0) (98=1) (99=0) .

variable labels Naslev00 "Number of AS level passes at grade A-C".

recode edgcasl1 (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into alevhi00 .
recode edgcasl2 (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into alevlo00 .

compute alev00=alevhi00 + alevlo00 .
recode alev00 (0=0) (1,2=1) .

variable labels alevhi00 "Has CM obtained one or more A levels at grades A to C?".
variable labels alevlo00 "Has CM obtained one or more A levels at grades D to E?".
variable labels alev00 "Has CM obtained any A level passes at any grade?".

compute t1=edgcasl1 .
compute t2=edgcasl2 .
recode t1 (missing=0) (98=1) (99=0) .
recode t2 (missing=0) (98=1) (99=0) .

* Someone with a 98 or 99 only gets 1 A-level.

compute Nalev00 = t1+t2 .

variable labels Nalev00 "Number of A level passes at any grade".

***** SCOTTISH SCHOOL QUALIFICATIONS *****
*freq edscot1 edscot2 edscot3 edscot4 edscot5 edscot6 .

* The Scottish qualifications are in edscot1 to edscot6 although edscot5 and edscot6 only have missing values.
 * This is not obvious from documentation.

```
compute scotlo00=0 .
compute scot200=0 .
compute scot300=0 .
compute scot400=0 .
compute scot500=0 .
compute scot600=0 .
```

```
variable labels scotlo00 "Has CM obtained Scottish SCE standard grades 4-5 or equivalent?".
variable labels scot200 "Has CM obtained Scottish SCE standard grades 1-3 or equivalent?".
variable labels scot300 "Has CM obtained Scottish SUPE-SLC lower or ordinary grade?".
variable labels scot400 "Has CM obtained Scottish SCE-SUPE-SLC higher grade or equivalent?".
variable labels scot500 "Has CM obtained Scottish Certificate of 6th Year Studies?".
variable labels scot600 "Has CM obtained other Scottish school qualification?".
```

* Scottish higher equivalent to A level in IFS definitions.

```
if ( any(1,edscot1,edscot2,edscot3,edscot4,edscot5,edscot6) ) scotlo00=1 .
if ( any(2,edscot1,edscot2,edscot3,edscot4,edscot5,edscot6) ) scot200=1 .
if ( any(3,edscot1,edscot2,edscot3,edscot4,edscot5,edscot6) ) scot300=1 .
if ( any(4,edscot1,edscot2,edscot3,edscot4,edscot5,edscot6) ) scot400=1 .
if ( any(5,edscot1,edscot2,edscot3,edscot4,edscot5,edscot6) ) scot500=1 .
if ( any(6,edscot1,edscot2,edscot3,edscot4,edscot5,edscot6) ) scot600=1 .
```

***** DEGREES *****

* Is 9 the missing value for numdeg.

```
recode numdeg (missing=0) (0=0) (9=0) (1 thru hi=1) into deg00 .
* Hi is 4.
```

```
variable labels deg00 "Has CM obtained a degree?".
```

```
recode numothdg (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into odeg00 .
* hi = 6.
```

```
variable labels odeg00 "Other degree level quals such as graduate membership of professional institutes".
```

```
compute ALTodeg=0 .
```

```
if (any(9,edqtyp14,edqtyp15,edqtyp16,edqtyp17,edqtyp18,edqtyp19,edqtyp20,edqtyp21,
edqtyp22,edqtyp23,edqtyp24,edqtyp25,edqtyp26) ) ALTodeg=1 .
```

```
*crosstabs ALTodeg by odeg00 .
```

```
*crosstabs numothdg by odeg00 .
```

```
recode numhghdg (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into hdeg00 .
*hi = 3, no 98 or 99.
```

```
variable labels hdeg00 "Has CM obtained a higher degree?".
```

***** NURSING, DIPLOMAS, PGCE *****

*There are 3 values for edqtyp21, all the values for edqtyp22 to edqtyp26 are missing.

```
compute dip00=0 .
```

```
if (any(7,edqtyp14,edqtyp15,edqtyp16,edqtyp17,edqtyp18,edqtyp19,edqtyp20,edqtyp21,
edqtyp22,edqtyp23,edqtyp24,edqtyp25,edqtyp26)
```

```
or (numotht > 0 & numotht < 15)) dip00=1 .
```

variable labels dip00 "Diploma-Certificate-Teacher Training Qual (not pgce) in Higher Ed".
* 357 have 1 teacher training qualification. 1 person has 12.

recode numparam (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into nurse00 .
*hi=15 No 98 or 99.

variable labels nurse00 "Has CM obtained any nursing or paramedic qualifications?".

compute pgce00=0 .
If (any(12,edqtyp14,edqtyp15,edqtyp16,edqtyp17,edqtyp18,edqtyp19,
edqtyp20,edqtyp21,edqtyp22,edqtyp23,edqtyp24,edqtyp25,edqtyp26)) pgce00=1

variable labels pgce00 "Has CM obtained a PGCE or other postgraduate certificate in HE?".

execute .

***** FIRST BLOCK FOR RUNNING PROGRAM *****

***** ONC or OND AND HNC or HND *****
recode voconc (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into oncond00 .
* hi=7 .

variable labels oncond00 "Does CM have ONC or OND?".

recode vohnc (missing=0) (0=0) (9=0) (98=1) (99=0) (1 thru hi=1) into hnchnd00 .
* hi=4. One value of 9.

variable labels hnchnd00 "Does CM have HNC or HND?".

**** BTECS ****.

compute btec100=0 .
compute btec200=0 .

variable labels btec100 "Does CM have a BTEC etc First-General Certificate?".
variable labels btec200 "Does CM have a BTEC etc First-General Diploma?".

compute btec300=0 .
compute btec400=0 .

variable labels btec300 "Does CM have a BTEC etc National Certificate Diploma?".
variable labels btec400 "Does CM have a BTEC etc Higher Certificate Diploma?".

compute btec500=0 .

variable labels btec500 "Does CM have other BTEC qualification?".

if (any(1,btecle1,btecle2,btecle3,btecle4,btecle5,btecle6,btecle7,btecle8,btecle9,
btecle10,btecle11,btecle12,btecle13,btecle14,btecle15)) btec100=1 .

if (any(2,btecle1,btecle2,btecle3,btecle4,btecle5,btecle6,btecle7,btecle8,btecle9,
btecle10,btecle11,btecle12,btecle13,btecle14,btecle15)) btec200=1 .

if (any(3,btecle1,btecle2,btecle3,btecle4,btecle5,btecle6,btecle7,btecle8,btecle9,
btecle10,btecle11,btecle12,btecle13,btecle14,btecle15)) btec300=1 .

if (any(4,btecle1,btecle2,btecle3,btecle4,btecle5,btecle6,btecle7,btecle8,btecle9,
btecle10,btecle11,btecle12,btecle13,btecle14,btecle15)) btec400=1 .

if (any(5,btecle9,btecle10,btecle11,btecle12,btecle13,btecle14,btecle15)) btec500=1 .

- * Only one observation in each of btecle9 to btecle15. All have first or general certificate.
- * Some values equal to 9 eg 89 observations with 9 for btecle9.

***** CITY AND GUILDS *****

compute city100=0 .
compute city200=0 .
compute city300=0 .
compute city400=0 .
compute city500=0 .

variable labels city100 "Does CM have a City and Guilds Part 1 " .
variable labels city200 "Does CM have a City and Guilds Part2 or Craft or Intermediate" .
variable labels city300 "Does CM have a City and Guilds Part3 or Final or Advanced Craft" .
variable labels city400 "Does CM have a City and Guilds Part4 or Career Extension or Full technological" .
variable labels city500 "Does CM have a City and Guilds Other C&G qualification" .

If (any(1,citylev,citylev2,citylev3,citylev4,citylev5,citylev6,citylev7,citylev8,citylev9,citylev10,citylev11,citylev12,citylev13,citylev14,citylev15)) city100=1 .

If (any(2,citylev,citylev2,citylev3,citylev4,citylev5,citylev6,citylev7,citylev8,citylev9,citylev10,citylev11,citylev12,citylev13,citylev14,citylev15)) city200=1 .

If any(3,citylev,citylev2,citylev3,citylev4,citylev5,citylev6,citylev7,citylev8,citylev9,citylev10,citylev11,citylev12,citylev13,citylev14,citylev15) city300=1 .

If any(4,citylev,citylev2,citylev3,citylev4,citylev5,citylev6,citylev7,citylev8,citylev9,citylev10,citylev11,citylev12,citylev13,citylev14,citylev15) city400=1 .

If any(5,citylev,citylev2,citylev3,citylev4,citylev5,citylev6,citylev7,citylev8,citylev9,citylev10,citylev11,citylev12,citylev13,citylev14,citylev15) city500=1 .

**** RSA *****

compute rsa100=0 .
compute rsa200=0 .
compute rsa300=0 .
compute rsa400=0 .
compute rsa500=0 .

variable labels rsa100 "Does CM have RSA certificate?".
variable labels rsa200 "Does CM have RSA First Diploma?".
variable labels rsa300 "Does CM have RSA Advanced Diploma or Certificate?".
variable labels rsa400 "Does CM have RSA Higher Diploma" .
variable labels rsa500 "Does CM have other RSA qualification?".

If (any(1,rsalev,rsalev2,rsalev3,rsalev4,rsalev5,rsalev6,rsalev7,rsalev8,rsalev9,rsalev10,rsalev11,rsalev12)) rsa100=1 .

If (any(2,rsalev,rsalev2,rsalev3,rsalev4,rsalev5,rsalev6,rsalev7,rsalev8,rsalev9,rsalev10,rsalev11,rsalev12)) rsa200=1 .

If (any(3,rsalev,rsalev2,rsalev3,rsalev4,rsalev5,rsalev6,rsalev7,rsalev8,rsalev9,rsalev10,

```

    rsalev11,rsalev12)) rsa300=1 .
If (any(4,rsalev,rsalev2,rsalev3,rsalev4,rsalev5,rsalev6,rsalev7,rsalev8,rsalev9,rsalev10,
    rsalev11,rsalev12)) rsa400=1 .
If (any(5,rsalev,rsalev2,rsalev3,rsalev4,rsalev5,rsalev6,rsalev7,rsalev8,rsalev9,rsalev10,
    rsalev11,rsalev12)) rsa500=1 .

```

* rsalev10 to rsalev12 only have 1 observation each. All values are for first diplomas (2).

**** PITMANS *****.

* Pitmans are only in the 2000 data.

```

compute pit100=0 .
compute pit200=0 .
compute pit300=0 .
compute pit400=0 .

```

```

variable labels pit100 "Does CM have a Pitmans level 1?".
variable labels pit200 "Does CM have a Pitmans level 2?".
variable labels pit300 "Does CM have a Pitmans level 3?".
variable labels pit400 "Does CM have other Pitmans qualification?".

```

```

If (any(1,pitlev,pitlev2,pitlev3,pitlev4,pitlev5,pitlev6,pitlev7)) pit100=1 .
If (any(2,pitlev,pitlev2,pitlev3,pitlev4,pitlev5,pitlev6,pitlev7)) pit200=1 .
If (any(3,pitlev,pitlev2,pitlev3,pitlev4,pitlev5,pitlev6,pitlev7)) pit300=1 .
If (any(4,pitlev,pitlev2,pitlev3,pitlev4,pitlev5,pitlev6,pitlev7)) pit400=1 .

```

* pitlev7 has one observation with a value of level 2.

***** NVQ *****.

*nvqs (current sweep only) .

```

compute nvq100=0 .
compute nvq200=0 .
compute nvq300=0 .
compute nvq400=0 .
compute nvq500=0 .
compute nvq600=0 .
compute nvq700=0 .
compute nvq800=0 .

```

```

variable labels nvq100 "Does CM have an NVQ at Level 1?".
variable labels nvq200 "Does CM have an NVQ at Level 2?".
variable labels nvq300 "Does CM have an NVQ at Level 3?".
variable labels nvq400 "Does CM have an NVQ at Level 4?".
variable labels nvq500 "Does CM have an NVQ at Level 5?".
variable labels nvq600 "Does CM have an NVQ at Level 6?".
variable labels nvq700 "Does CM have Trusts towards NVQ-SVQ?".
variable labels nvq800 "Does CM have other NVQ?".

```

```

If (any(1,nvqlev,nvqlev2,nvqlev3,nvqlev4,nvqlev5,nvqlev6,nvqlev7,nvqlev8,nvqlev9,nvqlev10))
nvq100=1 .
If (any(2,nvqlev,nvqlev2,nvqlev3,nvqlev4,nvqlev5,nvqlev6,nvqlev7,nvqlev8,nvqlev9,nvqlev10))
nvq200=1 .
If (any(3,nvqlev,nvqlev2,nvqlev3,nvqlev4,nvqlev5,nvqlev6,nvqlev7,nvqlev8,nvqlev9,nvqlev10))
nvq300=1 .
If (any(4,nvqlev,nvqlev2,nvqlev3,nvqlev4,nvqlev5,nvqlev6,nvqlev7,nvqlev8,nvqlev9,nvqlev10))
nvq400=1 .

```

```

If (any(5,nvqlev,nvqlev2,nvqlev3,nvqlev4,nvqlev5,nvqlev6,nvqlev7,nvqlev8,nvqlev9,nvqlev10))
nvq500=1 .
If (any(6,nvqlev,nvqlev2,nvqlev3,nvqlev4,nvqlev5,nvqlev6,nvqlev7,nvqlev8,nvqlev9,nvqlev10))
nvq600=1 .
If (any(7,nvqlev,nvqlev2,nvqlev3,nvqlev4,nvqlev5,nvqlev6,nvqlev7,nvqlev8,nvqlev9,nvqlev10))
nvq700=1 .
If (any(8,nvqlev,nvqlev2,nvqlev3,nvqlev4,nvqlev5,nvqlev6,nvqlev7,nvqlev8,nvqlev9,nvqlev10))
nvq800=1 .

```

*nvqlev7 to nvqlev10 have no values.

***** GNVQ *****

* gnvqs (current sweep only)

```

compute gnvq100=0 .
compute gnvq200=0 .
compute gnvq300=0 .
compute gnvq400=0 .

```

```

variable labels gnvq100 "Does CM have GNVQ Foundation?".
variable labels gnvq200 "Does CM have GNVQ Intermediate?".
variable labels gnvq300 "Does CM have GNVQ Advanced?".
variable labels gnvq400 "Does CM have other GNVQ qualification?".

```

```

If (any(1,gnvlev,gnvlev2,gnvlev3,gnvlev4,gnvlev5)) gnvq100=1 .
If (any(2,gnvlev,gnvlev2,gnvlev3,gnvlev4,gnvlev5)) gnvq200=1 .
If (any(3,gnvlev,gnvlev2,gnvlev3,gnvlev4,gnvlev5)) gnvq300=1 .
If (any(4,gnvlev,gnvlev2,gnvlev3,gnvlev4,gnvlev5)) gnvq400=1 .

```

*gnvlev4 and gnvlev5 have no values.

*** OTHER QUALIFICATIONS *****

```

recode vocappr (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into appren00 .
* hi is 7.

```

```

variable labels appren00 "Has CM obtained any recognised trade apprentice qualifications?".

```

```

recode vochgv (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into hgv00 .
* hi is 3.

```

```

variable labels hgv00 "Has CM obtained at least one hgv qualification?".

```

```

recode vocoth (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into othvoc00 .
* hi is 15.

```

```

variable labels othvoc00 "Has CM obtained at least one other vocational qualification?".

```

```

*****
***** MISSING VALUE ANALYSIS *****
*****

```

```

* edquals2, vocqual 1 ='Yes' 2='No' .

```

* We argue that if initial feeder questions (edquals2, vocqual) are answered 'properly' then we know whether the individual has a qualification or not. If the answer is missing or 'do not know' then we do not know. The code below sets all the educational qualifications to missing if edquals is not answered Yes or No or NA and all the vocational qualifications to missing if vocqual is not answered Yes or No or 'Not Applicable'.

recode edquals2 (2=0) (1=0) (8=1) (9=0) (missing=2) into acamiss.

variable label acamiss 'Dummy for missing value for do you have an educational qualification'.
value label acamiss 0 'not missing' 1 '8 (do not know)' 2 'missing'.

recode vocqual (1=0) (2=0) (8=1) (9=0) (missing=2) into vocmiss.

* vocmiss adjusted for nurse00 (nurse00 appears as an educational qualification) below.

variable label vocmiss 'Dummy for missing value for do you have a vocational qualification'.
value label vocmiss 0 'not missing' 1 '8 (do not know)' 2 'missing'.

do if (vocmiss eq 0 and acamiss eq 0) .

compute allmiss =0 .

else if (vocmiss eq 1 and acamiss eq 0) .

compute allmiss =1 .

else if (vocmiss eq 0 and acamiss eq 1) .

compute allmiss =2 .

else if (vocmiss eq 1 and acamiss eq 1) .

compute allmiss =3 .

else if (vocmiss eq 2 and acamiss eq 2) .

compute allmiss =4 .

end if .

variable label allmiss 'Missing values, Do you have any of these qualifications?'.
value label allmiss 0 'not missing' 1 '8 vocat quals, not missing for educ quals'

2 'vocat quals not missing, 8 for educ quals' 3 '8 for both vocat and educ quals'

4 'missing for both vocational & educational quals'.

***** Set missing values *****.

do if (acamiss gt 0) .

compute gcsehi00 =-9 .

compute gcselo00 =-9 .

compute olevhi00 = -9 .

compute olevlo00 = -9 .

compute csehi00 = -9 .

compute cselo00 = -9 .

compute aslhi00 = -9 .

compute asllo00 = -9 .

compute naslev00 = -9 .

compute alevhi00 = -9 .

compute alevlo00 = -9 .

compute alev00 = -9 .

compute nalev00 = -9 .

compute scotlo00 = -9 .

compute scot200 = -9 .

compute scot300 = -9 .

compute scot400 = -9 .

compute scot500 = -9 .

compute scot600 = -9 .

compute dip00 = -9 .

compute deg00 = -9 .

compute odeg00 = -9 .

compute dip00 = -9 .

compute hdeg00 = -9 .

compute nurse00 = -9 .

compute pgce00 = -9 .

end if .

if (nurse00 eq 1) vocmiss eq 0 .

* nurse is a vocational qualification under some typologies eg nvq but educational here.

```
do if (vocmiss gt 0) .
  compute oncond00 =-9 .
  compute hnchnd00 =-9 .
  compute btec100=-9 .
  compute btec200=-9 .
  compute btec300=-9 .
  compute btec400=-9 .
  compute btec500=-9 .
  compute city100 =-9 .
  compute city200 =-9 .
  compute city300 =-9 .
  compute city400 =-9 .
  compute city500 = -9 .
  compute rsa100 = -9 .
  compute rsa200 =-9 .
  compute rsa300 =-9 .
  compute rsa400 =-9 .
  compute rsa500 =-9 .
  compute pit100 =-9 .
  compute pit200 =-9 .
  compute pit300 =-9 .
  compute pit400 =-9 .
  compute nvq100 =-9 .
  compute nvq200 =-9 .
  compute nvq300 =-9 .
  compute nvq400 =-9 .
  compute nvq500 =-9 .
  compute nvq600 =-9 .
  compute nvq700 =-9 .
  compute nvq800 =-9 .
  compute gnvq100 =-9 .
  compute gnvq200 =-9 .
  compute gnvq300 =-9 .
  compute gnvq400 =-9 .
  compute appren00=-9 .
  compute hgv00 =-9 .
  compute othvoc00 =-9 .
end if .
```

missing value gcsehi00 to pgce00 oncond00 to othvoc00 (-9) .

```
compute proxy = 0 .
if (edquals gt 0) proxy =1 .
* if (vocquals gt 0) proxy1 =1 not needed as proxy = proxy1.
```

variable label proxy 'Was interview completed by a proxy'.

* Acamiss=vocmiss=1 when proxy=1 so proxy 'accounts' for 26 of the missing values.

```
compute acamiss1=0 .
compute vocmiss1=0 .
```

```
if (missing(edqtyp14) eq 1 and missing(edqtyp15) eq 1 and missing(edqtyp16) eq 1 and
  missing(edqtyp17) eq 1 and missing(edqtyp18) eq 1 and missing(edqtyp19) eq 1 and
  missing(edqtyp20) eq 1 and missing(edqtyp21) eq 1 and missing(edqtyp22) eq 1 and
```

missing(edqtyp23) eq 1 and missing(edqtyp24) eq 1 and missing(edqtyp25) eq 1 and
missing(edqtyp26) eq 1) acamiss1=1 .

if (missing(voctyp12) eq 1 and missing(voctyp13) eq 1 and missing(voctyp14) eq 1 and
missing(voctyp15) eq 1 and missing(voctyp16) eq 1 and missing(voctyp17) eq 1 and
missing(voctyp18) eq 1 and missing(voctyp19) eq 1 and missing(voctyp20) eq 1 and
missing(voctyp21) eq 1 and missing(voctyp22) eq 1) vocmiss1=1 .

***** GOVERNMENT TRAINING SCHEMES *****

recode numyts (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into yts00 .
* hi is 15.

variable labels yts00 "Has CM done a YTS course?".

recode numgov (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into gov00 .
* hi is 15.

variable labels gov00 "Has CM done another gov course (including New Deal)?".

recode numap (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into modapp00 .
* hi is 3.

variable labels modapp00 "Has CM done a modern apprenticeship?".

execute .

***** SECOND BLOCK WHEN RUNNING PROGRAM *****

***** OTHER COURSES *****

recode numfqual (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into nfail00.
*Hi =11.

variable labels nfail00 "Number of courses started where CM did not obtain the qualification".

recode numac (missing=0) (0=0) (98=1) (99=0) (1 thru hi=1) into access00 .
* hi is 10.

variable labels access00 "Has CM done an access course?".

compute workc00=numwrktr .
recode workc00 (missing=0) (0=0) (98=1) (99=0) .
* max is 30.

variable labels workc00 "Number of other work related courses".

***** NON WORK RELATED COURSES *****

compute leis00=numleis .
recode leis00 (missing=0) (0=0) (98=1) (99=0) .
* max is 15.

variable labels leis00 "Number of courses for interest or leisure".

**** READING, WRITING AND MATHS COURSES ***** .

recode numread (missing=0) (0=0) (98=1) (99=0) (1=1) (2=2) (3 thru hi=3) into read00 .
* max is 11.

variable labels read00 "Number of courses to improve reading".

recode numwrite (missing=0) (0=0) (98=1) (99=0) (1=1) (2=2) (3 thru hi=3) into write00 .
* max is 11.

variable labels write00 "Number of courses to improve writing".
* creative writing ?.

recode nummaths (missing=0) (0=0) (98=1) (99=0) (1=1) (2=2) (3 thru hi=3) into maths00 .
* max is 15.

variable labels maths00 "Number of courses to improve maths".
value labels read00 write00 maths00 3 "3 or more" .

*** THERE IS A LOT ABOUT CURRENT COURSES THAT WE HAVE NOT USED ***** .

***** DUMMIES FOR GROUPS OF QUALIFICATIONS *****

***** HAS CM OBTAINED A PARTICULAR ACADEMIC QUALIFICATION? ***** .

```
do if (acamiss gt 0) .  
  compute Aca000=-9 .  
  compute Aca100=-9 .  
  compute Aca200=-9 .  
  compute Aca300=-9 .  
  compute Aca400=-9 .  
  compute Aca500=-9 .  
  compute Aca600=-9 .  
  compute Aca700=-9 .  
  compute Aca800=-9 .
```

```
else .  
  compute Aca000=1 .  
  compute Aca100=0 .  
  compute Aca200=0 .  
  compute Aca300=0 .  
  compute Aca400=0 .  
  compute Aca500=0 .  
  compute Aca600=0 .  
  compute Aca700=0 .  
  compute Aca800=0 .
```

```
end if .
```

missing values Aca000 to Aca800 (-9) .

* AcaX00 dummy variables variables showing whether CM has one of a group of qualifications.

variable labels Aca000 "1 if CM has no academic qualifications".

variable labels Aca100 "1 if CM has bad O-levels ".

variable labels Aca200 "1 if CM has CSE 2-5, Other Scots ".

variable labels Aca300 "1 if CM has good O-levels ".

variable labels Aca400 "1 if CM has 1 A-level or more than 1 AS at grade A-C ".

variable labels Aca500 "1 if CM has 2 or more A-levels".

variable labels Aca600 "1 if CM has diploma "
variable labels Aca700 "1 if CM has degree, PGCE, other degree level "
variable labels Aca800 "1 if CM has higher degree".

value labels aca000 0 "At least one academic qualification" 1 "No academic qualifications".

* Bad O-levels

GCSEs D to E, GCEs D to E, Scottish SCE standard grades 4-5' .

* Good O-levels

CSE grade 1, GCSE grades A-C, GCE grades A-C, Scots standard grades 1-3, Scots lower or ordinary.

* 2 or more A-LEVELS

2 or more A-levels, Scottish higher grade or Certificate of 6th Year Studies .

```
do if (hdeg00 =1) .  
    compute Aca800=1 .  
    compute Aca000=0 .  
end if .
```

```
do if (any(1,deg00,odeg00,pgce00)) .  
    compute Aca700=1 .  
    compute Aca000=0 .  
end if .
```

```
do if (dip00 eq 1) .  
    compute Aca600=1 .  
    compute Aca000=0 .  
end if .
```

```
do if (Nalev00 gt 1 or scot400 eq 1 or scot500 eq 1) .  
    compute Aca500=1 .  
    compute Aca000=0 .  
end if .
```

```
do if (Nalev00 eq 1 or Naslev00 gt 1) .  
    compute Aca400=1 .  
    compute Aca000=0 .  
end if .
```

```
do if (any(1,gcsehi00,olevhi00,csehi00,scot200,scot300)) .  
    compute Aca300=1 .  
    compute Aca000=0 .  
end if .
```

```
do if (cselo00 eq 1 or scot600 eq 1) .  
    compute Aca200=1 .  
    compute Aca000=0 .  
end if .
```

```
do if (any(1,gcselo00,olevlo00,scotlo00)) .  
    compute Aca100=1 .  
    compute Aca000=0 .  
end if .
```

***** HAS CM OBTAINED A PARTICULAR VOCATIONAL QUALIFICATION? *****.

```
do if (Vocmiss gt 0) .  
    compute Voc000=-9 .  
    compute Voc100=-9 .
```

```

        compute Voc200=-9 .
        compute Voc300=-9 .
        compute Voc400=-9 .
        compute Voc500=-9 .
else .
        compute Voc000=1 .
        compute Voc100=0 .
        compute Voc200=0 .
        compute Voc300=0 .
        compute Voc400=0 .
        compute Voc500=0 .
end if .

```

missing values Voc000 to Voc500 (-9) .

* VocX00 dummy variables variables showing whether CM has one of a group of qualifications.

variable labels Voc000 "1 if CM has no vocational qualifications".
variable labels Voc100 "1 if CM has vocational qualification equivalent to NVQ1".
variable labels Voc200 "1 if CM has vocational qualification equivalent to NVQ2".
variable labels Voc300 "1 if CM has vocational qualification equivalent to NVQ3".
variable labels Voc400 "1 if CM has vocational qualification equivalent to NVQ4".
variable labels Voc500 "1 if CM has vocational qualification equivalent to NVQ5".

value labels voc000 0 "At least one vocational qualification" 1 "No vocational qualifications".

```

*          100 'C&G Opo & Other, RSA 1, HGV, Other tech or bus qualification'
          200 'C&G Craft-Inter-Ord-Part1-Cant say, RSA2, Insig,JIB NJC etc'
          300 'C&G Advanced-Final-Part II or III, ONC-OND, RSA 3 '
          400 'Nursing, BTEC HNC-HND, HNC-HND, C&G Full Tech (FTC)'
          500 'NVQ awarded at level 5'.

```

```

do if (any(1,nvq500,nvq600)) .
        compute Voc500=1 .
        compute Voc000=0 .
end if .

```

```

do if (any(1,nurse00,hnchnd00,btec400,rsa400,nvq400)) .
        compute Voc400=1 .
        compute Voc000=0 .
end if .

```

```

do if (any(1,city200,city300,city400,oncond00,btec300,rsa300,pit300,nvq300,gnvq300)) .
        compute Voc300=1 .
        compute Voc000=0 .
end if .

```

```

do if (any(1,city100,btec200,rsa200, pit200,appren00,nvq200,gnvq200)) .
        compute Voc200=1 .
        compute Voc000=0 .
end if .

```

```

do if (any(1,city500,btec100,rsa100,pit100,nvq100,gnvq100,hgv00,othvoc00,rsa500,
        btec500,pit400,nvq700,nvq800,gnvq400)) .
        compute Voc100=1 .
        compute Voc000=0 .
end if .

```

 ***** HIGHEST QUALIFICATIONS *****

- * This code computes the highest academic qualification, the highest vocational qualification and the highest NVQ level recorded in the 2000 survey.
- * Previous published work has used slightly different definitions of these variables so you may need to adapt the code below.
- * Check bad O-levels, many descriptions give them NVQ level of 0
- * All vocational stuff needs checking especially pitmans btec rsa .

***** HIGHEST ACADEMIC QUALIFICATION *****

```
do if (acamiss gt 0) .
  compute Hiaca00=-9 .

else if (Aca800=1) .
  compute Hiaca00=8 .

else if (Aca700=1) .
  compute Hiaca00=7 .

else if (Aca600=1) .
  compute Hiaca00=6 .

else if (Aca500=1) .
  compute Hiaca00=5 .

else if (Aca400=1) .
  compute Hiaca00=4 .

else if (Aca300=1) .
  compute Hiaca00=3 .

else if (Aca200=1) .
  compute Hiaca00=2 .

else if (Aca100=1) .
  compute Hiaca00=1 .

else if (acamiss eq 0) .
  compute Hiaca00=0 .

end if .
```

variable label Hiaca00 'Highest academic qualification recorded in 2000 survey'.

value labels Hiaca00
 -9 Missing
 0 'None'
 1 'Bad O-levels'
 2 'CSE 2-5, other Scottish school qualification'
 3 'Good O levels'
 4 '1 A level or more than 1 AS level at grade A-C'
 5 '2 or more A-levels'
 6 'Diploma'
 7 'Degree, PGCE, other degree level'

8 'Higher degree' .

* Notice diploma=degree .

recode Hiaca00 (-9=-9) (0=0) (1=1) (2=1)(3=2)(4=2)(5=3)(6=4)(7=4)(8=5) into nvqaca00 .

variable label nvqaca00 'Highest NVQ level from an academic qualification in 2000 survey'.

recode Hiaca00 (-9=-9) (0=0) (1=1) (2=1)(3=2)(4=2)(5=3)(6=4)(7=5)(8=6) into ghmaca00 .

variable label ghmaca00 'Highest academic qualification in 2000 survey ghm measure see also IFS'.

value labels ghmaca00 0 'None' 1 'Bad O levels, CSE 2-5' 2 'Good O-levels, 1 A-level'
3 '2 or more A-levels' 4 'Sub-degree' 5 'Degree' 6 'Higher Degree' .

***** HIGHEST VOCATIONAL QUALIFICATION *****

do if (vocmiss gt 0) .
compute HiVoc00=-9 .

else if (Voc500 eq 1) .
compute HiVoc00=5 .

else if (Voc400 eq 1) .
compute HiVoc00=4 .

else if (Voc300 eq 1) .
compute HiVoc00=3 .

else if (Voc200 eq 1) .
compute HiVoc00=2 .

else if (Voc100 eq 1) .
compute HiVoc00=1 .

else if (vocmiss eq 0) .
compute HiVoc00=0 .

end if .

variable label HiVoc00 'Highest vocational qualification recorded in 2000 survey'.

* HiVoc00 is also the highest NVQ level from a vocational qualification in 2000 survey.

value labels HiVoc00

-9 Missing

0 'None'

1 'C&G Opo & Other, RSA 1, HGV, Other tech or bus qualification'

2 'C&G Craft-Inter-Ord-Part1-Cant say, RSA2, Insig,JIB NJC etc'

3 'C&G Advanced-Final-Part II or III, ONC-OND, RSA 3'

4 'Nursing, BTEC HNC-HND, HNC-HND, C&G Full Tech (FTC)'

5 'NVQ awarded at level 5'.

***** Highest NVQ whatever type *****

compute HiNVQ00=max(nvqaca00,HiVoc00) .

variable label HiNVQ00 'Highest NVQ level (whether academic or vocational)' .

missing value Hiaca00 nvqaca00 ghmaca00 HiVoc00 HiNVQ00 (-9) .

***** OTHER VARIABLES *****

```
do if (sample eq 1) .
  compute group = 1.
else if (sample eq 2 and dmpart eq 1) .
  compute group = 2.
else if (sample eq 2 and dmpart eq 2) .
  compute group = 3.
end if .
```

variable label group "3 categories - BCS70 and NCDS by whether present in 1991 survey" .
value label group 1 "BCS70" 2 "NCDS present in 1991" 3 "NCDS not present in 1991" .

recode dmsex (1=0) (2=1) into female .
value labels female 0 "male" 1 "female" .

***** SAVE DATA *****

save outfile = 'e:\00\ncds-bcs70\qca data\all2000.sav'
/ drop ALToDeg t1 t2 vocmiss1 acamiss1 .

***** THIRD BLOCK WHEN RUNNING PROGRAM *****

***** FINISH *****

2. Samantha Parsons

**JUNE 2000

**MERGING NCDS5 QUALIFICATION DATA WITH 2000 DATA.

**CHANGED 2000 TO MATCH NCDS33.

**simple way of recoding highest academic/vocational qualifications - to match NCDS at 33 hqual vars.

**ghmaca00 - levels of academic qualifications: ranges from 0 to 6: with 5 = degree and 6 = higher degree.

.

**hivoc00 - levels of vocational NVQ qualifications: ranges from 0 to 5: with 5 = NVQ level 5.

**variable HINVQ00 combines the information from the two variables above with a range of 0 to 5, BUT level 5 represents ghmaca00 level 6 and hivoc00 level 5 only. Surely it should be ghmaca00 level 5 AND 6 and hivoc00 level 5?.

**made such a variable .

```
compute hq2ka = ghmaca00.  
if (hivoc00 > ghmaca00) hq2ka = hivoc00.
```

```
recode hq2ka (5,6=5).
```

** Highest qualification at NCDS5 (age 33) – NB: Uses NCDS5 variables n501513 to n501541

```
do rep x=n501441 to n501469/y=n501513 to n501541.  
if x=37 and range (y,1,36) x=y.  
if x=37 and missing (y) or y=0 y=37.  
end rep.  
compute hqual33=-2.  
do rep x=n501441 to n501469.  
if any (x,37) hqual33=0.  
if (any (x,10,25,1) and hqual33 < 1) hqual33=1.  
if (any (x,20,19,18,17,13,14,12,11,7,6,4,3,2) and hqual33 < 2) hqual33=2  
if (any (x,23,21,15,09,08,05) and hqual33 < 3) hqual33=3.  
if (any (x,30,29,28,27,26,24,22,16) and hqual33 < 4) hqual33=4.  
if (range (x,31,33) and hqual33 < 5) hqual33=5.  
end rep.  
var labels hqual33 "Highest qual gained at age 33 – based on NCDS5 variables ".  
value labels hqual33 0 'No qualification' 1 'CSE 2-5/equiv NVQ1' 2 'O Level/equiv NVQ2' 3 'A  
Level/equiv NVQ3' 4 'Higher qual NVQ4' 5 'Degree/higher NVQ5/6' -1 'No information'.  
formats hqual33 f2.0).  
recode hqual33 (-2=-1) missing values hqual33 (-1).
```

**DIRTY highest qualification variable combining information of NCDS cohort at age 33.

```
compute hq332k = hq2ka.  
if (hq2ka = 0 and hqual33 >=0) hq332k = hqual33.  
do if (hq2ka > 0 and hqual33 >= 0).  
if (hqual33 < hq2ka) hq332k = hq2ka.  
if (hqual33 = hq2ka) hq332k = hq2ka.  
if (hqual33 > hq2ka) hq332k = hqual33.
```

end if.

variable labels hq332k 'highest NVQ qualification - academic or vocational - recoding 2000 to match NCDS5'.

frequencies
variables=hq332k.

```
*****  
*  
**some qualifications coded differently at each sweep, so made individual qualification vars at age 33  
to re-create various highest qualification variables at 33 to match 2000.  
*****
```

```
*  
  
do if (n501441 >= 0).  
compute noquals = 0.  
compute cse2_5 = 0.  
compute cse1 = 0.  
compute oleva_c = 0.  
compute gcsea_c = 0.  
compute alev = 0.  
compute scotoa_c = 0.  
compute scots1_3 = 0.  
compute scothg = 0.  
compute scot6th = 0.  
compute rsa1 = 0.  
compute rsa2 = 0.  
compute rsa3 = 0.  
compute cg_op = 0.  
compute cg_cio1 = 0.  
compute cg_afp23 = 0.  
compute cg_ft = 0.  
compute cg_oth = 0.  
compute cg_dk = 0.  
compute cg_ia = 0.  
compute jibnjcct = 0.  
compute onchnd = 0.  
compute hnchnd = 0.  
compute bt_ngcd = 0.  
compute bt_hhncd = 0.  
compute othtbq = 0.  
compute fprofq = 0.  
compute pprofq = 0.  
compute nurse = 0.  
compute cert = 0.  
compute hdippgce = 0.  
compute degree = 0.  
compute pgraddip = 0.  
compute hdegree = 0.  
do repeat x = n501441 to n501469.  
if (x = 37) noquals = 1.  
if (x = 1) cse2_5 = 1.  
if (x = 2) cse1 = 1.  
if (x = 3) oleva_c = 1.  
if (x = 4) gcsea_c = 1.  
if (x = 5) alev = 1.  
if (x = 6) scotoa_c = 1.  
if (x = 7) scots1_3 = 1.
```

```

if (x = 8) scothg = 1.
if (x = 9) scot6th = 1.
if (x = 10) rsa1 = 1.
if (x = 11) rsa2 = 1.
if (x = 12) rsa3 = 1.
if (x = 13) cg_op = 1.
if (x = 14) cg_cio1 = 1.
if (x = 15) cg_afp23 = 1.
if (x = 16) cg_ft = 1.
if (x = 17) cg_oth = 1.
if (x = 18) cg_dk = 1.
if (x = 19) cg_ia = 1.
if (x = 20) jibnjcct = 1.
if (x = 21) onchnd = 1.
if (x = 22) hnchnd = 1.
if (x = 23) bt_ngcd = 1.
if (x = 24) bt_hhncd = 1.
if (x = 25) otthbq = 1.
if (x = 26) fprofq = 1.
if (x = 27) pprofq = 1.
if (x = 28) nurse = 1.
if (x = 29) cert = 1.
if (x = 30) hdippgce = 1.
if (x = 31) degree = 1.
if (x = 32) pgraddip = 1.
if (x = 33) hdegree = 1.
end repeat.
end if.

```

frequencies

```

variables=noquals cse2_5 cse1 oleva_c gcsea_c alev scotoa_c scots1_3 scothg scot6th rsa1 rsa2
rsa3 cg_op cg_cio1 cg_afp23
cg_ft cg_oth cg_dk cg_ia jibnjcct onchnd hnchnd bt_ngcd bt_hhncd otthbq fprofq pprofq nurse cert
hdippgce degree pgraddip hdegree .

```

```

*****
*****
**making highest qualification variables to best match 2000 data categories.
**a levels are treated as equivalent to 2 plus alevels in 2000 as th enumber of a levels not collected in
NCDS5.
*****
*****

```

```

*****
**highest academic qualification to match 8 category HIACA00.
**a levels in ncds5 are treated as equivalent to 2 plus alevels in 2000.
*****

```

```

compute hacad33b = 0.
if (any(1,cse2_5)) hacad33b = 2.
if (any(1,cse1,oleva_c,gcsea_c,scotoa_c,scots1_3)) hacad33b = 3.
if (any(1,alev,scothg,scot6th)) hacad33b = 5.
if (any(1,cert,hdippgce)) hacad33b = 6.
if (any(1,degree,pgraddip)) hacad33b = 7.
if (hdegree = 1) hacad33b = 8.
if (missing(noquals)) hacad33b = -1.
missing values hacad33b (-1).

```

variable labels hacad33b 'ncds33: highest academic qualification to match HIACA00'.
value labels hacad33b 0'no quals' 2'cse2-5' 3'good o levels' 5'a levels' 6'diploma' 7'degree PGCE'
8'higher degree'.

frequencies
variables=hacad33b.

**highest academic qualification variable - NVQ levels - to match NVQACA00.

recode hacad33b (0=0) (2=1) (3=2) (5=3) (6,7=4) (8=5) into hacnvq33.
variable labels hacnvq33 'ncds33: highest NVQ level qual from academic qualification to match
NVQACA00'.

frequencies
variables=hacnvq33.

**highest academic qualification variable - ghm measure to match GHMACA00.

recode hacad33b (0=0) (2=1) (3=2) (5=3) (6=4) (7=5) (8=6) into hacghm33.
variable labels hacghm33 'ncds33: highest academic qualification - ghm measure to match
GHMACA00'.
value labels hacghm33 0'no quals' 1'cse2-5' 2'o levels' 3'a levels' 4'sub degree' 5'degree' 6'higher
degree'.

frequencies
variables=hacghm33.

**highest vocational qualification at 33 - including nursing - to NVQ levels - to match HIVOC00'.

compute hvoc33 = 0.
if (any(1,rsa1,othtbq,cg_oth,cg_dk)) hvoc33 = 1.
if (any(1,rsa2,cg_op,cg_cio1,cg_ia,jibnjcct)) hvoc33 = 2.
if (any(1,rsa3,cg_afp23,onchnd,bt_ngcd)) hvoc33 = 3.
if (any(1,cg_ft,hnchnd,bt_hhncd,fprofq,pprofq,nurse)) hvoc33 = 4.
if (missing(noquals)) hvoc33 = -1.
missing values hvoc33 (-1).

variable labels hvoc33 'ncds33: highest vocational qualification - to match HIVOC00'.

frequencies
variables=hvoc33.

**highest academic or vocational qualification - to NVQ levels - to match HINVQ00.

compute hq33a = 0.
if (any(1,cse2_5,rsa1,othtbq,cg_oth,cg_dk)) hq33a = 1.
if (any(1,cse1,oleva_c,olevd_e,scotoa_c,scots1_3,rsa2,cg_op,cg_cio1,cg_ia,jibnjcct)) hq33a = 2.
if (any(1,alev,scothg,scot6th,rsa3,cg_afp23,onchnd,bt_ngcd)) hq33a = 3.

```
if (any(1,cg_ft,hnchnd,bt_hhncd,fprofq,pprofq,nurse,cert,hdippgce,degree,pgraddip)) hq33a = 4.
if (hdegree = 1) hq33a = 5.
if (missing(noquals)) hq33a = -1.
missing values hq33a (-1).
```

variable labels hq33a 'ncds33: highest NVQ qualification - academic or vocational - to match HINVQ00'.

```
frequencies
  variables=hq33a.
```

```
*****
**COMBINING NCDS33 AND 2000 DATA VARIABLES.
*****

*****
**combined highest academic qualification variables - 8 category.
*****
```

```
compute hqaca00 = HIACA00.
if (HIACA00 = 0 and hacad33b >=0) hqaca00 = hacad33b.
do if (HIACA00 > 0 and hacad33b >= 0).
if (hacad33b < HIACA00) hqaca00 = HIACA00.
if (hacad33b = HIACA00) hqaca00 = HIACA00.
if (hacad33b > HIACA00) hqaca00 = hacad33b.
end if.
```

variable labels hqaca00 'highest academic qualification - combining hacad33b and hiaca00'.
value labels hqaca00 0'no quals' 1'bad o levels' 2'cse2-5' 3'good o levels' 4'as levels or 1 a level' 5'a levels' 6'diploma' 7'degree PGCE' 8'higher degree'.

```
frequencies
  variables=hacad33b hiaca00 hqaca00.
```

```
*****
**combined highest academic qualification variables - NVQ levels.
*****
```

```
compute hqanvq00 = nvqaca00.
if (nvqaca00 = 0 and hacnvq33 >= 0) hqanvq00 = hacnvq33.
do if (nvqaca00 > 0 and hacnvq33 >= 0).
if (nvqaca00 > hacnvq33) hqanvq00 = nvqaca00.
if (nvqaca00 = hacnvq33) hqanvq00 = nvqaca00.
if (nvqaca00 < hacnvq33) hqanvq00 = hacnvq33.
end if.
```

variable labels hqanvq00 'highest NVQ level qual from academic qualification - combining hacnvq33 and nvqaca00'.

```
frequencies
  variables=hqanvq00 nvqaca00 hacnvq33 .
```

```
*****
**combined highest academic qualification variables - ghm measure .
*****
```

```

compute hqaghm00 = ghmaca00.
if (ghmaca00 = 0 and hacghm33 >= 0) hqaghm00 = hacghm33.
do if (ghmaca00 > 0 and hacnvq33 >= 0).
if (ghmaca00 > hacghm33) hqaghm00 = ghmaca00.
if (ghmaca00 = hacghm33) hqaghm00 = ghmaca00.
if (ghmaca00 < hacghm33) hqaghm00 = hacghm33.
end if.

```

variable labels hqaghm00 'highest academic qualification - ghm measure - combining hacghm33 and ghmaca00'.

value labels hqaghm00 0'no quals' 1'cse2-5' 2'o levels' 3'a levels' 4'sub degree' 5'degree' 6'higher degree'.

```

frequencies
  variables=hacghm33 ghmaca00 hqaghm00.

```

```

*****
**highest vocational qualification.
*****

```

```

compute hqvoc00 = hivoc00.
if (hivoc00 = 0 and hvoc33 >= 0) hqvoc00 = hvoc33.
do if (hivoc00 > 0 and hvoc33 >= 0).
if (hivoc00 > hvoc33) hqvoc00 = hivoc00.
if (hivoc00 = hvoc33) hqvoc00 = hivoc00.
if (hivoc00 < hvoc33) hqvoc00 = hvoc33.
end if.

```

variable labels hqvoc00 'highest vocational qualification - combining hvoc33 and hivoc00'.

```

frequencies
  variables=hvoc33 hivoc00 hqvoc00.

```

```

*****
**highest academic or vocational qualification - to NVQ levels.
*****

```

```

compute hqnvq00 = hinvq00.
if (hinvq00 = 0 and hq33a >= 0) hinvq00 = hq33a.
do if (hinvq00 > 0 and hq33a >= 0).
if (hinvq00 > hq33a) hqnvq00 = hinvq00.
if (hinvq00 = hq33a) hqnvq00 = hinvq00.
if (hinvq00 < hq33a) hqnvq00 = hq33a.
end if.

```

variable labels hqnvq00 'highest NVQ qualification - academic or vocational - combining hq33a and hinvq00'.

```

frequencies
  variables=hq33a hinvq00 hqnvq00.

```

APPENDIX 8: Attitude Scales

Documentation relating to the development of attitude scales produced by John Preston of the Centre for Longitudinal Studies is reproduced below:

Queries about the contents of this appendix should be directed to the User Support Group
(cohort@cls.ioe.ac.uk)

Attitude Scales for latest sweeps of NCDS and BCS70

Methodology

Attitude scales were constructed using the 50 items in the 'views' section of the self-completion questionnaire. Each survey was analysed separately, although the factors identified in each were, fortuitously, identical. Principal components analysis was employed using varimax rotation to orthogonal simple structure. The number of factors chosen was based upon those with Eigen values greater than one, followed by the 'scree' test to determine that point at which Eigen values leveled off. Factors were examined for interpretability and to ensure that at least three items had loadings on each of them.

Cronbach's alpha was employed as a test of internal consistency of the factors. In all cases, alpha scores were above 0.5 which indicates that it is acceptable to consider the items as forming an identifiable factor.

Factor loadings are shown in each table, values less than +/- 0.4 have been suppressed. For each factor, a derived variable has been created which is the mean of the scores for each item within the component. The name of this derived variable is displayed in bold type. Certain items on the factor were inverted prior to the calculation and this is indicated beneath each table.

Factors identified

1. *Anti-racism*

TOLERACE	BCS70	NCDS
AR2 Wouldn't mind if family of different race moved next door	.826	.791
AR4 Wouldn't mind working with people from other races	.774	.745
AR3 Wouldn't mind kids going to school with other races	.755	.724
AR5 Would not want another race person as my boss	-.688	-.675
AR1 Mixed race marriage is OK	.676	.663

All scale items, with the exception of AR5 were inverted prior to construction of the scale which indicates the level of support for tolerant views related to other ethnic groups.

2. Left-Right

LEFTR	BCS70	NCDS
LR5 Ordinary people don't get a fair share of the nations wealth.	.651	.692
LR6 Government should redistribute income.	.648	.655
LR7 One law for the rich and one for the poor?	.612	.641
LR1 Big business benefits owners at the expense of workers.	.572	.605
LR2 Private schools should be abolished	.562	.583
LR3 Management get the better of employees	.500	.567

The magnitude of the score on LEFTR indicates the level of sympathy with right wing views in terms of redistribution of income, privatization and support for business and management.

3. Support for Authority

AUTHORIT	BCS70	NCDS
A4 Give law breakers stiffer sentences	.661	.675
A6 Schools teach children to authority	.613	.627
A2 Death penalty for some crimes	.566	.593
A5 Young people don't have respect for traditional values	.564	.560
A3 Censorship is needed to uphold morals	.481	.463
A1 The law should be obeyed even if it is wrong	.442	.441

All items were inverted prior to the construction of a scale of support for authority.

4. Support for traditional marital values

PROFAM	BCS70	NCDS
MOR4 Marriage is for life	.647	.718
MOR6 Alright for unmarried people to have kids	-.629	-.620
MOR3 Couples with kids should not separate	.592	.618
MOR1 Divorce is too easy to get these days	.533	.544
MOR5 Women should have the right to an abortion	-.479	-.519
MOR2 Married people happier than unmarried	.478	.430

Items MOR4, MOR3, MOR1 and MOR2 were inverted prior to the calculation of this scale which indicates the level of support for marriage and traditional family values.

5. Effect of children on quality of life

NOKIDS	BCS70	NCDS
C4 People with no kids are missing out.	.756	.731
C1 Unless you have kids, you will be lonely in your old age.	.690	.688
C2 Can have fulfilling life with no kids.	-.661	-.652

C2 was inverted to C2inv prior to the calculation of the scale. NOKIDS indicates the level of agreement with statements opposing the value of children in life experiences. The higher the value of NOKIDS, the greater the level of agreement with these statements.

5. Support for working Mothers

MUMWORK	BCS70	NCDS
WM3 Kids benefit if mum has job outside the home	.702	.716
WM4 A mother and family are happier if she goes out to work	.700	.713
WM1 Pre-school kids suffer if mum works	-.607	-.611
WM2 Family life suffers if mum is working full time	-.607	-.594
WM5 Dads job is to earn money; mums to stay home	-.408	<-0.4

WM3 and WM4 were inverted prior to the calculation of the scale. MUMWORK indicates agreement with Mothers working. Higher scores indicate that respondents believe that women should work. Note that in the PCA for NCDS, the factor loading for WM5 was of a scalar magnitude less than 0.4.

6. Political cynicism

POLITICAL CYNICISM (POLCYN)	BCS70	NCDS
PC2 No difference which political party is in power	.769	.769
PC1 No political party would benefit me	.703	.710
PC3 Politicians are in politics for their own benefit	.630	.676

Scores on POLCYN indicate agreement with statements regarding the futility and redundancy of the current political system. All scale items were inverted prior to construction of the scale.

7. Technophobia

INFORMATION TECHNOLOGY (ANTITECH)	BCS70	NCDS
IT2 Computers enrich the lives of users	.717	.739
IT4 Every family should have a computer	.693	.670
IT1 Computers at work are destroying peoples skills	-.563	-.534
IT5 Learning to use a computer is more trouble than it is worth	-.477	-.479

IT1 and IT5 were inverted prior to the creation of the attitude scale. The higher the score on ANTITECH, the greater the degree of opposition to information technology.

8. Environmentalism

ENVIRONMENTALISM (ENVIRON)	BCS70	NCDS
E3 The environment vs. economic growth	.722	.717
E2 Preserving the environment is most important.	.691	.705
E1 Problems in the environment are not that serious	-.640	-.589

E2 and E3 were inverted to E1inv prior to the construction of the attitude scale. The higher the score on ENVIRON, the greater the degree of individuals support for environmental issues.

9. Support for the work ethic

WORK ETHIC (WORKETH)	BCS70	NCDS
WE3 It is important to hang on to any job even if unhappy	.765	.777
WE2 If I didn't like a job I'd pack it in	-.756	-.758
WE1 Any job is better than being unemployed	.567	.551

WE1 and WE3 were inverted before construction of the scale which indicates an individuals support for the work ethic.

Alpha coefficients

SCALE	BCS70	NCDS
Anti-Racism (TOLERACE)	0.8222	0.8000
Left-Right (LEFTR)	0.6792	0.7313
Support for authority (AUTHORIT)	0.6176	0.6470
Support for traditional marital values (PROFAM)	0.6345	0.6620
Effect of children on quality of life (NOKIDS)	0.6232	0.6141
Support for working mothers (MUMHOME)	0.6782	0.6880
Political cynicism (POLCYN)	0.6526	0.6742
Technophobia (ANTITECH)	0.5732	0.5892
Environmentalism (ENVIRON)	0.5070	0.4664
Support for the Work ethic (WORKETH)	0.5469	0.5388

Items not included in scales

The following items were not included in the above scales and did not appear to form any clear factors of their own:-

- C3 Having children interferes with parents freedom
- L2 Knowing the right people helps more than qualifications in getting a job
- L4 Effort of getting qualifications is more trouble than it is worth
- L1 More likely to get better job if do education training
- IT5 Learning to use a computer is more trouble than it is worth
- LR4 Take out own health care, stop relying on the NHS

John Preston
21/04/2001

SPSS Code

```
MISSING VALUES lr1 lr2 lr3 lr4 lr5 lr6 lr7 ar1 ar2 ar3 ar4 ar5 e1 e2 e3 a1 a2 a3 a4 a5 a6 c1 c2 c3 c4  
pc1 pc2 pc3 l1 l2 l3 l4 mor1 mor2 mor3 mor4 mor5 mor6 wm1 wm2 wm3 wm4 wm5 it1 it2 it4 it5 we1  
we2 we3
```

```
("8." "9.").
```

```
COMPUTE tolerace = MEAN((6-ar2),(6-ar4),(6-ar3),ar5,(6-ar1)) .
```

```
EXECUTE .
```

```
COMPUTE leftr = MEAN(lr5,lr6,lr7,lr1,lr2,lr3) .
```

```
EXECUTE .
```

```
COMPUTE authorit = MEAN((6-a4),(6-a6),(6-a2),(6-a5),(6-a3),(6-a1)) .
```

```
EXECUTE .
```

```
COMPUTE profam = MEAN((6-mor4),mor6,(6-mor3),(6-mor1),mor5,(6-mor2)) .
```

```
EXECUTE .
```

```
COMPUTE nokids = MEAN(c4,c1,(6-c2)) .
```

```
EXECUTE .
```

```
COMPUTE mumwork = MEAN((6-wm3),(6-wm4),wm1,wm2,wm5) .
```

```
EXECUTE .
```

```
COMPUTE polcyn = MEAN((6-pc2),(6-pc1),(6-pc3)) .
```

```
EXECUTE .
```

```
COMPUTE antitech = MEAN(it2,it4,(6-it1),(6-it5)) .
```

```
EXECUTE .
```

```
COMPUTE environ = MEAN((6-e3),(6-e2),e1) .
```

```
EXECUTE .
```

```
COMPUTE worketh = MEAN((6-we1),we2,(6-we3)) .
```

```
EXECUTE .
```

```
SUMMARIZE
```

```
  /TABLES=ar2 ar4 ar3 ar5 ar1 tolerace
```

```
  /FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50
```

```
  /TITLE='Case Summaries'
```

```
  /MISSING=VARIABLE
```

```
  /CELLS=COUNT .
```

```
SUMMARIZE
```

```
  /TABLES=lr5 lr6 lr7 lr1 lr2 lr3 leftr
```

```
  /FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50
```

```
  /TITLE='Case Summaries'
```

```
  /MISSING=VARIABLE
```

```
  /CELLS=COUNT .
```

```
SUMMARIZE
```

```
  /TABLES=a4 a6 a2 a5 a3 a1 authorit
```

```
  /FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50
```

```
  /TITLE='Case Summaries'
```

```
  /MISSING=VARIABLE
```

```
  /CELLS=COUNT .
```

```
SUMMARIZE
```

```
  /TABLES=mor4 mor6 mor3 mor1 mor5 mor2 profam
```

```
  /FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50
```

```
  /TITLE='Case Summaries'
```

```
  /MISSING=VARIABLE
```

```
  /CELLS=COUNT .
```

```
SUMMARIZE
```

```
  /TABLES=c4 c1 c2 nokids
```

```
  /FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50
```

```
  /TITLE='Case Summaries'
```

```
  /MISSING=VARIABLE
```

```
  /CELLS=COUNT .
```

```
SUMMARIZE
```

```
  /TABLES=wm3 wm4 wm1 wm2 wm5 mumwork
```

```
  /FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50
```

```
/TITLE='Case Summaries'  
/MISSING=VARIABLE  
/CELLS=COUNT .  
SUMMARIZE  
/TABLES=pc2 pc1 pc3 polcyn  
/FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50  
/TITLE='Case Summaries'  
/MISSING=VARIABLE  
/CELLS=COUNT .  
SUMMARIZE  
/TABLES=it2 it4 it1 it5 antitech  
/FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50  
/TITLE='Case Summaries'  
/MISSING=VARIABLE  
/CELLS=COUNT .  
SUMMARIZE  
/TABLES=e3 e2 e1 environ  
/FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50  
/TITLE='Case Summaries'  
/MISSING=VARIABLE  
/CELLS=COUNT .  
SUMMARIZE  
/TABLES=we3 we2 we1 worketh  
/FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50  
/TITLE='Case Summaries'  
/MISSING=VARIABLE  
/CELLS=COUNT .
```

```
/TITLE='Case Summaries'  
/MISSING=VARIABLE  
/CELLS=COUNT .  
SUMMARIZE  
/TABLES=pc2 pc1 pc3 polcyn  
/FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50  
/TITLE='Case Summaries'  
/MISSING=VARIABLE  
/CELLS=COUNT .  
SUMMARIZE  
/TABLES=it2 it4 it1 it5 antitech  
/FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50  
/TITLE='Case Summaries'  
/MISSING=VARIABLE  
/CELLS=COUNT .  
SUMMARIZE  
/TABLES=e3 e2 e1 environ  
/FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50  
/TITLE='Case Summaries'  
/MISSING=VARIABLE  
/CELLS=COUNT .  
SUMMARIZE  
/TABLES=we3 we2 we1 worketh  
/FORMAT=VALIDLIST CASENUM TOTAL LIMIT=50  
/TITLE='Case Summaries'  
/MISSING=VARIABLE  
/CELLS=COUNT .
```