National Child Development Study and 1970 British Cohort Study Technical Report:

Changes in the NCDS and BCS70 Populations and Samples over Time

1st Edition

October 2004

By

Ian Plewis, Lisa Calderwood, Denise Hawkes and Gad Nathan

CLS

Centre for Longitudinal Studies
Bedford Group for Lifecourse and Statistical Studies
Institute of Education, University of London

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

The Centre for Longitudinal Studies (CLS) is one of five centres that comprise the Bedford Group for Lifecourse and Statistical Studies (www.ioe.ac.uk/bedfordgroup). CLS is devoted to the collection, management and analysis of large-scale longitudinal data. It has responsibility for Britain's internationally renowned birth cohort studies, the National Child Development Study (1958 cohort) and the 1970 British Cohort Study, and leads the consortium conducting the ESRC's Millennium Cohort Study.

The views expressed in this work are those of the authors and do not necessarily reflect the views of the Economic and Social Research Council. All errors and omissions remain those of the authors.

Contents

Acknowledgements

Preface to the First Edition

This document is the first edition of what is intended to be an expanding account of the NCDS and BCS70 populations and samples as they change over time. The second edition will update the material presented here to include the two sweeps (NCDS, sweep 7 and BCS70, sweep 6) currently in the field. This edition covers the complete sweeps of NCDS up to sweep 6 (when the cohort members were age 42 in the year 2000) and BCS70 up to sweep 5 (when the cohort members were age 30 in the year 2000). It does not, however, yet include any material on studies of sub-samples of the two cohorts nor about the NCDS sample of children of the cohort members.

Response data files for the two studies, containing survey outcome codes for each cohort member at each sweep, were used to construct the tables in this report. These files have been deposited at the UK Data Archive.

1. Introduction

1.1 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 17000 babies born in Great Britain during one particular week in 1958. Since this first study the whole cohort had, by 2000, been surveyed on six 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), 1991 (age 33) and 1999/2000 (age 41/42). 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 5000 children. There have also been surveys of sub-samples of the cohort, the most recent occurring in 1996 (age 37) when information was collected on the basic skills of a 10% sample of cohort members.

1.2 The 1970 British Cohort Study (BCS70) was designed along similar lines to the NCDS, surveying over 17000 babies born in Great Britain in a week in 1970. Since the birth survey there had, by 2000, been five 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), 1996 (age 26), and 1999/2000 (age 29/30). As in NCDS, sub-samples have been studied at various ages: for example at age 21, paralleling the NCDS survey at age 37, a 10% sample was assessed for difficulties with basic skills.

1.3 From their original focus on the circumstances and outcomes of birth, the two cohort studies have broadened in scope to map many aspects of the health, education and social development of their subjects as they passed through childhood and adolescence. In the more recent 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.

1.4 In common with all longitudinal studies, the sizes of the NCDS and BCS70 samples decline as the cohorts age. This report describes the extent and patterns of these declines and the reasons for them. It draws on earlier work (see Nathan, 1999 and references therein) but presents, for the first time, a coherent series of tables that set out all that can be known about response patterns up to the year 2000. Much of the data presented here can be regarded as definitive but some, because of the nature of continuing longitudinal studies, will change as the cohorts continue to mature and hence further information about the cohort members is obtained.

1.5 The report is structured as follows. Sections 2 and 3 define the populations and target samples for the two studies, drawing out the important distinction between longitudinal (Section 2) and cross-sectional (Section 3) populations. The declines in the longitudinal *target* samples over time are then set out in Section 4. Section 5 discusses the extent to which the cohort members were measured at the target ages. Section 6 shows how, as a result of non-

response, the *observed* samples are always smaller than the target samples. Estimated response rates at each sweep are also set out in this section. More detailed data about response patterns, in particular the way in which cohort members move in and out of the samples as they get older are provided in Section 7. Section 8 focuses on response patterns for domains of interest (education, health etc.) and how these vary from the overall patterns of response. Section 9 describes the samples as if they were cross-sectional samples at each age. Section 10 sums up the results.

1.6 This report is purely descriptive in its accounts of the samples. Questions about the correlates of sample loss and non-response, and the effects that these losses might have on substantive analyses are answered elsewhere (Hawkes and Plewis, 2004).

1.7 It is crucial to have a clear definition of the population in all quantitative social research. This is relatively straightforward for a cross-sectional study but more complicated for a longitudinal design. It is the definition of the longitudinal population that is of most interest for NCDS and BCS70 but, because these studies are sometimes used just cross-sectionally, we also define the populations at each contact age. We refer to contacts as 'sweeps' and adopt the convention of labelling sweeps sequentially from zero (at birth) upwards.

2. Longitudinal Populations and Target Samples

NCDS

2.1 The longitudinal population for inference is all children born (alive or dead) in Great Britain in 1958, until they die or permanently emigrate from Great Britain. 'In 1958' is a little restrictive and could be replaced by something a little less precise such as births in Great Britain in the late 1950s, say.

2.2 The longitudinal target sample is all children born (alive or dead) in Great Britain in a specific week in March 1958, until they die or permanently emigrate from Great Britain.

2.3 Note that the longitudinal population and target sample are restricted to births in Great Britain even though information was obtained in some later sweeps about subjects born outside GB. This highlights a difference between the longitudinal and cross-sectional populations (see Table 3.1 and also Endnote 1).

2.4 Any inferences from the longitudinal target sample to the broader population for inference must rest on the untestable but perhaps not unreasonable assumption that births in a week in March 1958 can, for many purposes, be regarded as a random sample of all births in the period of interest.

BCS70

2.5 The longitudinal population for inference is all children born (alive or dead) in Great Britain after the 24th week of gestation in 1970, until they die or

permanently emigrate from Great Britain. Again, as for the NCDS, 'in 1970' might be replaced by, for example, 'around 1970'.

2.6   The longitudinal target sample is all children born (alive or dead) in Great Britain after the 24th week of gestation in a specific week in April 1970, until they die or permanently emigrate from Great Britain.

2.7   These are, in all their important elements, the same as the NCDS definitions. BCS70 shares with NCDS the potential difficulty of making inferences from a single week's births to a wider population. Note that births in Northern Ireland, which were included in the birth survey, are excluded from these definitions because these births were not retained in the longitudinal sample at later sweeps.

3.   Cross-sectional Populations and Target Samples

NCDS

3.1   The cross-sectional populations and target samples for inference at sweep t (t = 0,1 …) are all children born anywhere in 1958 (population) or in a specific week in March (target sample), and living in Great Britain at sweep t (including stillbirths at sweep 0), that is at target ages 0, 7, 11, 16, 23, 33, 42 ...

3.2   Note that the sizes of the cross-sectional populations and target samples vary with t, not necessarily monotonically (i.e. consistently) downwards, because subjects not born in GB are included in them whereas temporary emigrants from GB are excluded. The first of these two sub-groups is excluded from the longitudinal population whereas the second is included. Immigrants to GB were, in principle, included in NCDS up to age 16 (sweep 3) but not thereafter. Thus, to the extent that immigration continued after age 16 (in 1974), the cross-sectional samples have a built-in bias by excluding post-16 immigrants entirely.

BCS70

3.3   The cross-sectional populations and target samples for inference at sweep t (t = 0, 1…) are all children born anywhere in 1970 (population) or in a specific week in April 1970 (target sample), living in Great Britain at sweep t (including stillbirths at sweep 0), that is at target ages 0, 5, 10, 16, 26, 30 ... (see Endnote 2). The previous remarks about the NCDS cross-sectional populations also apply to BCS70.

3.4   We can summarise the definitions used in this section and the previous one, in terms of which groups are included and excluded from the target populations, in Table 3.1.

Table 3.1:   Summary of the population definitions

	Born in GB during reference period			Born abroad during reference period and immigrated to GB before sweep *t*	
	Living in GB at sweep *t*	Dead or emigrated at sweep *t*		Living in GB at sweep *t*	Dead or emigrated at sweep *t*
		Temporary emigrant	Dead or permanent emigrant		Permanent or temporary emigrant
Longitudinal population – for period [0,*t*]	Included	Included	Excluded	Excluded	Excluded
Cross-sectional population - at sweep *t*	Included	Excluded	Excluded	Included	Excluded

4.   The Evolution of the Longitudinal Target Samples

NCDS

4.1   As shown in Table 4.1, the longitudinal target sample consists of 17634 babies at birth and then declines (monotonically by definition) to 15324 adults by sweep 6 at target age 42 (see endnote 3).

Table 4.1:   NCDS longitudinal target sample, sweeps 0 to 6.

SWEEP (AGE)	0 (0)	1 (7)	2 (11)	3 (16)	4 (23)	5 (33)	6 (42)
Target sample	17634 (100%)	16500 (93.6%)	16253 (92.2%)	16068 (91.1%)	15885 (90.1%)	15567 (88.3%)	15451 (87.6%)
Permanent emigrants	0	322 (1.8%)	552 (3.1%)	705 (4.0%)	869 (4.9%)	1090 (6.2%)	1190 (6.7%)
Deaths	0	812 (4.6%)	829 (4.7%)	861 (4.9%)	880 (5.0%)	977 (5.5%)	993 (5.6%)
Total	**17634**	**17634**	**17634**	**17634**	**17634**	**17634**	**17634**

Notes

1. Included in the total (n = 17634) are 219 cases coded at sweep 3 as born in GB (the only sweep which included a question about country of birth) but with no data at sweep 0. Out of these 219 cases, 163 joined the study at sweep 1, 23 at sweep 2 and 33 at sweep 3. It is possible (but not ascertainable) that there are other cases, mistakenly treated as immigrants at sweeps 1 and 2, for whom country of birth was not coded at sweep 3 and who were actually born in GB and therefore should be in the longitudinal target sample.

2. The deaths at sweep 1 include 388 stillbirths, 224 early neonatal deaths and 200 deaths after seven days.
3. Permanent emigrants are cohort members who leave the country and do not return, as confirmed by checking NHS records. The notion of permanence becomes less secure for later sweeps and it is possible that some cases will be reassigned at later sweeps. It is also likely that this method of defining permanent emigrants underestimates the true number (Hattersley and Creeser, 1995).
4. It is likely that deaths and permanent emigrants are underestimated from sweep 4 onwards (and the target sample therefore overestimated) because some of the refusals (see Table 6.1) would have fallen into these categories if contact had been maintained with them.
5. Percentages are based on the total (17634).
6. See endnote 3 for the evolution of the target sample from sweep 1.

BCS70

4.2   The longitudinal target sample consists of 17287 babies and then declines (monotonically by definition) to 16253 adults by sweep 5 at target age 30 (Table 4.2).

Table 4.2:   BCS70 longitudinal target sample, sweeps 0 to 5: available data.

SWEEP (AGE)	0 (0)	1 (5)	2 (10)	3 (16)	4 (26)	5 (30)
Target sample	17287 (100%)	16720 (96.7%)	16700 (96.6%)	16690 (96.5%)	16545 (95.7%)	16253 (94.0%)
Permanent emigrants	0	0	0	0	45 (0.2%)	287 (1.6%)
Deaths	0	567 (3.3%)	587 (3.4%)	597 (3.5%)	697 (4.0%)	747 (4.3%)
Total	**17287**	**17287**	**17287**	**17287**	**17287**	**17287**

Notes

1. Included in the total (n = 17287) are 715 cases coded at sweeps 2 or 3 as born in GB (the only sweeps which included a question about country of birth) but with no data at sweep 0. Out of these 715 cases, 248 joined the study at sweep 1, 427 at sweep 2 and 40 at sweep 3. It is possible (but not ascertainable) that there are other cases, mistakenly treated as immigrants at sweep 1, for whom country of birth was not coded at sweeps 2 and 3 and who were actually born in GB and therefore should be in the longitudinal target sample. The total also includes one still-born case, not included in the birth survey but identified as eligible from the death certificate.
2. The deaths at sweep 1 include 235 stillbirths, 173 early neonatal deaths and 159 deaths after seven days.
3. There is no information on permanent emigrants before sweep 4. Consequently, the target sample sizes after sweep 0 are too high. An alternative, albeit estimated, longitudinal target sample is given in Table 4.3.
4. Percentages are based on the total (17287).

Table 4.3:   BCS70 longitudinal target sample, sweeps 0 to 5: emigration estimated.

SWEEP (AGE)	0 (0)	1 (5)	2 (10)	3 (16)	4 (26)	5 (30)
Target sample	17287	16461 (95.2%)	16181 (93.6%)	15999 (92.5%)	15726 (91.0%)	15503 (89.7%)
Permanent emigrants	0	259 (1.5%)	519 (3%)	691 (4%)	864 (5%)	1037 (6%)
Deaths	0	567 (3.3%)	587 (3.4%)	597 (3.5%)	697 (4.0%)	747 (4.3%)
Total	**17287**	**17287**	**17287**	**17287**	**17287**	**17287**

4.3   The estimated permanent emigration rates in Table 4.3 are essentially those found for NCDS and so the decline in the estimated BCS70 longitudinal target sample is a little slower than for NCDS as death rates in infancy were lower. It is likely that the target sample estimated in Table 4.3 reflects more accurately the true situation than Table 4.2 does but it does depend on the assumption that permanent emigration rates did not change substantially between the two cohorts. It is likely that deaths and permanent emigrants are underestimated from sweep 4 onwards (and the target sample therefore overestimated) because some of the refusals (see Table 6.3) would have fallen into these categories if contact had been maintained with them.

5.   Target Ages and Measurement Ages

5.1   At each sweep for both studies, there is variability around the target age in the age at which cohort members were actually measured as well as some variability from instrument to instrument for individual cohort members.

5.2   At NCDS sweep 1, for 3% of the sample the parental questionnaire was administered when the cohort member was age eight rather than age seven; at sweep 2, 1.3% of the cohort members were 12 or over at the time the parental questionnaire was administered; at sweep 3, 30% were only age 15 rather than 16 but very few were more than 16; at sweep 4, 1.1% of the cohort members were measured at age 22 rather than at 23 but none were more than 23; at sweep 5, 0.5% were age 32 rather than 33 but none were more than 33 and at sweep 6, 56% were age 41 and 44% were age 42.

5.3   At BCS70 sweep 1, the parental questionnaire was administered when the cohort member was age four rather than age five in 5% of cases but in only a handful of cases was the cohort member age six; at sweep 2, 19% of cohort members were under age 10 when the parental questionnaire was administered but again only a handful were over 10; at sweep 3, 7% of cohort members were age 17 at the time of parent interview. There are no data about cohort members' age at data collection for sweep 4; at sweep 5, 68% were age 29 and 32% age 30.

5.4   The intervals between data collections were (on average in years) seven, four, five, seven, ten and nine for NCDS, and five, five, six, ten and four for BCS70.

In other words, data collection was somewhat more intensive for BCS70 than for NCDS up to the year 2000.

6. Non-response: the Difference between Longitudinal Target and Observed Samples

NCDS

6.1   Table 6.1 sets out the differences between the *longitudinal* target and observed samples at each sweep. (Table 9.2 gives the corresponding data for the *cross-sectional* target and observed samples and Table 10.1 compares the two kinds of observed samples.) It shows how the observed sample declines with age, both in absolute terms and as a fraction of the target sample (taken from Table 4.1) up to sweep 5. There is a sharp decline between sweeps 3 and 4 that corresponds to (and might be caused by) the change in respondent – from a parent or carer to the cohort members themselves as adults age 23. Most of this decline in the observed sample is accounted for by 'non-response: other'

Table 6.1:   NCDS longitudinal target and observed samples, sweeps 0 to 6.

SWEEP (AGE)	0 (0)	1 (7)	2 (11)	3 (16)	4 (23)	5 (33)	6 (42)
Observed sample	17415 (98.8%)	15051 (91.2%)	14757 (90.8%)	13917 (86.6%)	12044 (75.8%)	10986 (70.6%)	10979 (71.1%)
Non-response: refusal	0	80 (0.5%)	783 (4.8%)	1114 (6.9%)	1130 (7.1%)	1735 (11.1%)	2043 (13.2%)
Non-response: other	219 (1.2%)	1178 (7.1%)	491 (3.0%)	708 (4.4%)	1705 (10.7%)	1100 (7.1%)	308 (2.0%)
Uncertain eligibility	0	191 (1.2%)	222 (1.4%)	329 (2.0%)	1006 (6.3%)	1746 (11.2%)	2121 (13.7%)
Target sample	**17634 (100%)**	**16500 (100%)**	**16253 (100%)**	**16068 (100%)**	**15885 (100%)**	**15567 (100%)**	**15451 (100%)**

Notes

1. The observed sample consists of cohort members with at least some data at that sweep. The percentage is not necessarily the response rate – see Table 6.2.
2. Non-response: refusal. Refusals were re-issued up to sweep 4 but not thereafter. The apparent increase in the refusal rate could also be explained by more accurate recording of survey outcomes thereafter.
3. Non-response: other. This group consists of cases for which there are no data (either a response or a refusal/other contact) for the current sweep but some data (either a response or a refusal/other contact) at a later sweep (which means that they are still in the longitudinal population at this sweep), and 'temporary emigrants' who were known to be abroad then but who returned to GB later. The cases in this category at sweep 6 were traced but not interviewed.
4. Uncertain eligibility. These are cases for whom eligibility is unknown because there are no data (either a response or a refusal/other contact) both in this sweep

and in any later sweep. These cases will include cases who were not traced (and were therefore eligible) and also cases who had died or permanently emigrated (and were therefore not eligible). At sweeps 5 and 6 in particular, this group includes cases likely to be reallocated to 'non-response: other' as we may find evidence in the future that indicates that they were eligible at these sweeps. For example, they may be traced and take part in a subsequent sweep. This is less likely though possible for cases in this group from earlier sweeps.

6.2   Table 6.2 sets out the response rates by sweep. Note that the calculation of the response rate depends on the assumptions made about the eligibility of those cases labelled 'uncertain eligibility' in Table 6.1 (American Association for Public Opinion Research, 2000). Table 6.2 shows that response rates decline with age, gently during the ages of childhood but with a marked shift downwards as the cohort members become adults. The response rates, however defined, are slightly higher for sweep 6 than for sweep 5.

Table 6.2:   NCDS longitudinal response rates, sweeps 0 to 6.

SWEEP (AGE)	0 (0)	1 (7)	2 (11)	3 (16)	4 (23)	5 (33)	6 (42)
Eligibility rate (ER)	100%	93.5%	92.1%	91.0%	89.5%	87.0%	86.0%
Response rate – estimated (RR)	98.8%	91.3%	90.9%	86.8%	76.3%	71.6%	73.0%
Response rate - upper limit (URR)	98.8%	92.3%	92.0%	88.4%	81.0%	79.5%	84.3%
Response rate – lower limit (LRR)	98.8%	91.2%	90.8%	86.6%	75.8%	70.6%	71.4%

Notes

1. The eligibility rate is defined as:

    ER = (Observed sample + Non-response)/(Observed sample + Non-response + Not eligible)

    The 'not eligible' group comprises deaths and permanent emigrants.

2. The alternative response rates in the last three rows are defined as:

    RR = Observed sample/ (Observed sample + Non-response + (ER * Uncertain eligibility)). This assumes that those with unknown eligibility are divided between non-response and not eligible in the same ratio as those with known eligibility.

    URR: assumes an ER=0, i.e. that none on those with uncertain eligibility are eligible - unlikely in this case especially for sweep 6.

    LRR: assume that ER=1 i.e. that all of those with uncertain eligibility are eligible - not unreasonable as most of these are likely to be untraced cases rather than deaths/permanent emigrants.

BCS70

6.3   The response pattern for BCS70 in Table 6.3 is rather different from that of
        NCDS in Table 6.1. We see marked fluctuations in the observed sample as a
        fraction of the total sample, being relatively low at sweeps 1, 3 and 4.
        Moreover, there are more cases with uncertain eligibility in BCS70 than in
        NCDS. (Table 9.6 gives the sizes of the cross-sectional target and observed
        samples and Table 10.2 compares the two kinds of observed samples.)

**Table 6.3:   BCS70 estimated longitudinal target sample and observed sample,
                sweeps 0 to 5.**

SWEEP (AGE)	0 (0)	1 (5)	2 (10)	3 (16)	4 (26)	5 (30)
Observed sample	16571 (95.9%)	12981 (78.9%)	14350 (88.7%)	11206 (70.0%)	8654 (55.0%)	10833 (69.9%)
Non-response	716 (4.1%)	2812 (17.0%)	1108 (6.9%)	3293 (20.6%)	4765 (30.3%)	1833 (11.8%)
Uncertain eligibility	0	668 (4.1%)	723 (4.5%)	1500 (9.4%)	2307 (14.7%)	2837 (18.3%)
Target sample (estimated)	17287 (100%)	16461 (100%)	16181 (100%)	15999 (100%)	15726 (100%)	15503 (100%)

Notes

1. The target sample is based on the estimated numbers from Table 4.3.
2. There was not enough information from earlier sweeps (0-3) to divide non-
   response into the two sub-categories used for NCDS.
3. The observed sample at sweep 1 includes 42 cases born in GB but living (and
   interviewed) abroad at the time of interview who subsequently returned to GB.
4. Uncertain eligibility. These are cases for whom eligibility is unknown because
   there are no data (either a response or a refusal/other contact) both in this sweep
   and in any later sweep. These cases will include cases who were not traced (and
   were therefore eligible) and also cases who had died or permanently emigrated
   (and were therefore not eligible). At sweep 5, this group includes cases likely to
   be reallocated to 'non-response' at subsequent sweeps. Generally, the rates are
   higher than for NCDS, possibly because record keeping was less accurate (see
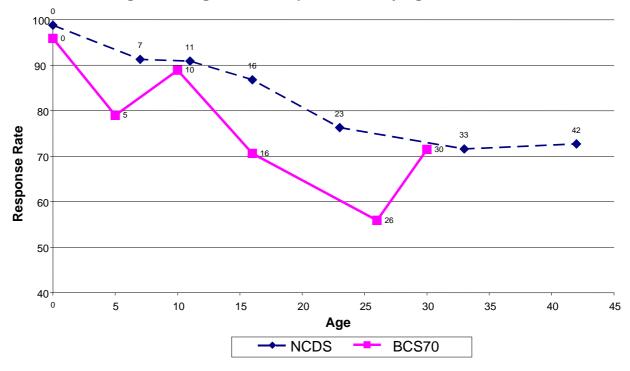   note 2 above).

Table 6.4:   BCS70 longitudinal response rates, sweeps 0 to 5.

SWEEP (AGE)	0 (0)	1 (5)	2 (10)	3 (16)	4 (26)	5 (30)
Eligibility rate (ER)	100%	95.0%	93.3%	91.8%	89.6%	87.7%
Response rate - estimated (RR)	95.9%	79.0%	88.9%	70.6%	55.9%	71.5%
Response rate – upper limit (URR)	95.9%	82.2%	92.8%	77.3%	64.5%	85.5%
Response rate – lower limit (LRR)	95.9%	78.9%	88.7%	70.0%	55.0%	69.9%

Note

For definitions of eligibility and response rates see notes after Table 6.2.

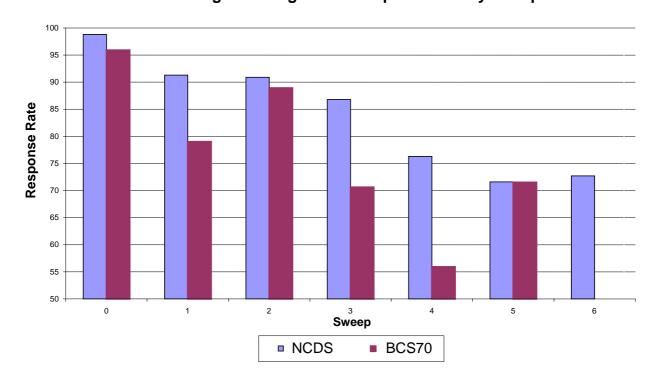**Fig 6.1: Longitudinal response rate by age of cohort member**

Fig 6.2: Longitudinal response rate by sweep

6.4   Figures 6.1 and 6.2 compare the longitudinal response rates for the two studies by plotting them by age of cohort member (Fig. 6.1) and by sweep (Fig. 6.2). The rates are generally higher for NCDS but appear to be converging in adulthood: the response rate for BCS70 at sweep 5 (age 30) is close to the corresponding response rate in NCDS (at age 33). BCS70 sweep 4 was conducted by mail and cohort members were not traced once the survey had started, which probably explains why the response rate was low then.

7.   Patterns of Response for the Longitudinal Target Samples

7.1   Response patterns of cohort members in the longitudinal target samples are described in detail in this section. The ways in which cohort members move in and out of the observed samples over time are first described prospectively by looking at responses at later sweeps for the observed samples from each sweep. This tells us what percentage of eligible respondents from an earlier sweep took part at each of the later sweeps. Response patterns are also examined retrospectively for the members of the observed sample at each sweep. This tells us what percentage of respondents from later sweeps took part at each of the earlier sweeps.

Response from each sweep at later sweeps

NCDS

7.2   Table 7.1 shows the relationship between the observed sample at a baseline sweep and the target and observed samples at all later sweeps, conditional on being in the observed sample at the particular baseline sweep.

7.3 The numbers in the observed samples for each of the baseline sweeps are given in the second column and the cumulative effect of the loss of these respondents (through death and permanent emigration) from the target samples at later sweeps is shown. For example, of the 17415 cohort members in the observed sample at sweep 0, 1745 had died or permanently emigrated by sweep 4 leaving 15670 who were in the target sample at this sweep. The longitudinal target sample at sweep 4 is, from Table 4.1, 15885. In Table 7.1, the target samples, conditional on being observed in sweeps 0 to 3, fall from 15670 (sweep 0) to 13762 (sweep 3). The differences between the target samples and observed samples at later sweeps are also shown. Thus only 11889 out of the 15670 members of the sweep 0 observed sample who are in the target sample at sweep 4 are actually observed. This proportion is also expressed as a conditional response rate i.e. 76.3% of the sweep 0 observed sample who are in the sweep 4 target sample were observed at sweep 4.

7.4 Table 7.1 shows that attrition due to death and emigration increases over time (by definition) and is highest from the sweep 0 observed sample due to higher death rates around birth and in infancy. In general, attrition due to non-response is low in the non-adult sweeps (1-3) and increases at the adult sweeps (4-6). For example, the response rates of the sweep 0 observed sample is over 86% at sweeps 1, 2 and 3 falling to around 73% at sweeps 4, 5 and 6. Attrition due to non-response appears to have stabilised somewhat in adulthood. For example, of the observed sample from sweep 4 who are in the target sample at sweeps 5 and 6, 80.2% and 80.0% respectively were interviewed at these later sweeps.

Table 7.1: NCDS response in later sweeps

From Sweep to Sweep	Observed Sample at Baseline Sweep	Deaths/ Emigrants	Target Sample at Later Sweep	Non-response	Uncertain Eligibility	Observed Sample at Later Sweep	Eligibility Rate (ER)	Condit. Response Rate - Estimated (RR)	Condit. Response Rate - Upper Limit (URR)	Condit. Response Rate - Lower Limit (LRR)
0-1	17415	1134	16281	1205	188	14888	93.4%	91.5%	92.5%	91.4%
0-2	17415	1381	16034	1247	213	14574	92.0%	91.0%	92.1%	90.9%
0-3	17415	1566	15849	1851	297	13701	90.9%	86.6%	88.1%	86.5%
0-4	17415	1745	15670	2949	832	11889	89.5%	76.3%	80.1%	75.9%
0-5	17415	2059	15356	3014	1503	10839	87.1%	71.5%	78.2%	70.6%
0-6	17415	2174	15241	2023	2390	10828	85.5%	72.7%	84.3%	71.1%
1-2	15051	244	14807	965	25	13817	98.4%	93.3%	93.5%	93.3%
1-3	15051	402	14649	1625	89	12935	97.3%	88.3%	88.8%	88.3%
1-4	15051	566	14485	2682	562	11241	96.1%	77.7%	80.7%	77.6%
1-5	15051	856	14195	2762	1188	10245	93.8%	72.6%	78.8%	72.2%
1-6	15051	966	14085	1831	2006	10248	92.6%	73.5%	84.8%	72.8%
2-3	14757	169	14588	1432	76	13080	98.9%	89.7%	90.1%	89.7%
2-4	14757	331	14426	2573	535	11318	97.7%	78.5%	81.5%	78.5%
2-5	14757	624	14133	2676	1170	10287	95.4%	73.1%	79.4%	72.8%
2-6	14757	737	14020	1763	1989	10268	94.2%	73.8%	85.4%	73.2%
3-4	13917	155	13762	2275	507	10980	98.8%	79.8%	82.8%	79.8%
3-5	13917	446	13471	2413	1110	9948	96.5%	74.1%	80.5%	73.9%
3-6	13917	548	13369	1548	1911	9910	95.4%	74.6%	86.5%	74.1%
4-5	12044	265	11779	1770	572	9437	97.7%	80.2%	84.2%	80.1%
4-6	12044	351	11693	1128	1247	9318	96.8%	80.0%	89.2%	79.7%
5-6	10986	71	10915	511	896	9508	99.3%	87.2%	94.9%	87.1%

For definitions of eligibility and response rates see notes after Table 6.2.

7.5   Table 7.2 gives a summary of response at all later sweeps for the observed sample at any given sweep. It should be noted that the reason for non-participation at a later sweep might be because the cohort member has died or permanently emigrated and these cases are not distinguished in this table. The table shows, as expected, that the percentage of the observed sample for any given sweep observed in all later sweeps decreases as the number of later sweeps increases. Only 39.4% of the observed sample at sweep 0 took part in all later sweeps and 7.9% were not observed in any later sweep.

Table 7.2:   NCDS summary of response at later sweeps

SWEEP	0	1	2	3	4	5
Observed Sample	17415	15051	14757	13917	12044	10986
Observed in all later sweeps	6863 (39.4%)	6946 (46.2%)	7302 (49.5%)	7645 (54.8%)	8309 (68.9%)	9508 (86.6%)
Observed in all but one later sweep	3635 (20.9%)	3329 (22.1%)	3286 (22.3%)	2923 (21.0%)	2137 (17.7%)	
Observed in all but two later sweeps	2563 (14.7%)	2298 (15.3%)	2234 (15.1%)	2057 (14.8%)		
Observed in all but three later sweeps	1676 (9.6%)	1471 (9.8%)	1419 (9.6%)			
Observed in all but four later sweeps	775 (4.5%)	604 (4.0%)				
Observed in all but five later sweeps	536 (3.1%)					
Not observed in any later sweep	1367 (7.9%)	403 (2.7%)	516 (3.5%)	1292 (9.3%)	1598 (13.3%)	1478 (13.5%)

BCS70

7.6   As for NCDS, Table 7.3 shows the relationship between the observed sample at any given baseline sweep and the target and observed samples at all later sweeps, conditional on being in the observed sample at the baseline sweep.

7.7   In contrast with NCDS, attrition due to non-response fluctuates at each later sweep and is consistent with the pattern of response rates in Table 6.4. These fluctuations are related to variability in the baseline response rate – lower

baseline response rates, for example, tend to lead to higher response rates at subsequent sweeps.

Table 7.3:   BCS70 response in later sweeps

From Sweep to Sweep	Observed Sample at Baseline Sweep	Deaths/ Emigrants (est.)	Target Sample at Later Sweep (est.)	Non-response	Uncertain Eligibility (est.)	Observed Sample at Later Sweep	Eligibility Rate (ER)	Condit. Response Rate - Estimated (RR)	Condit. Response Rate - Upper Limit (URR)	Condit. Response Rate - Lower Limit (LRR)
0-1	16571	864	15707	2345	629	12733	94.6%	812%	84.4%	81.1%
0-2	16571	1116	15455	1063	712	13680	93.0%	88.8%	92.8%	88.5%
0-3	16571	1258	15313	3189	1450	10674	91.7%	70.3%	77.0%	69.7%
0-4	16571	1521	15050	4569	2187	8294	89.4%	56.0%	64.5%	55.1%
0-5	16571	1765	14806	1762	2659	10385	87.3%	71.8%	85.5%	70.1%
1-2	12981	212	12769	654	123	11992	98.4%	93.9%	94.8%	93.9%
1-3	12981	338	12643	2472	650	9521	97.3%	75.4%	79.4%	75.3%
1-4	12981	550	12431	3825	1275	7331	95.3%	59.3%	65.7%	59.0%
1-5	12981	764	12217	1421	1656	9140	93.3%	75.5%	86.5%	74.8%
2-3	14350	137	14123	2801	822	10590	99.0%	74.6%	79.1%	74.5%
2-4	14350	376	13974	4314	1567	8093	97.1%	58.1%	65.2%	57.9%
2-5	14350	619	13731	1580	1996	10155	95.0%	74.5%	86.5%	74.0%
3-4	11206	179	11027	3196	901	6930	98.3%	67.3%	68.4%	62.8%
3-5	11206	379	10827	1088	1194	8545	96.2%	85.5%	88.7%	78.9%
4-5	8654	170	8484	577	522	7385	97.9%	90.1%	92.8%	87.0%

For definitions of eligibility and response rates see notes after Table 6.2.

7.8   Table 7.4 gives a summary of response at all later sweeps for the observed sample at any given sweep. It should be noted that the reason for non-participation at a later sweep may be because the cohort member has died or permanently emigrated and these cases are not distinguished in this table. The table shows, as expected, that the percentage of the observed sample for any given sweep observed in all later sweeps decreases as the number of later sweeps increases. Only 30.1% of the observed sample at sweep 0 took part in all later sweeps and 9.5% were not observed in any later sweep.

Table 7.4:   BCS70 summary of response at later sweeps

SWEEP	0	1	2	3	4
Observed Sample	16571	12981	14350	11206	8654
Observed in all later sweeps	4984 (30.1%)	5101 (39.3%)	5793 (40.4%)	6071 (54.2%)	7385 (85.3%)
Observed in all but one later sweep	4295 (25.9%)	3805 (29.3%)	4314 (30.1%)	3333 (29.7%)	
Observed in all but two later sweeps	3099 (18.7%)	2491 (19.2%)	2831 (19.7%)		
Observed in all but three later sweeps	1745 (10.5%)	1183 (9.1%)			
Observed in all but four later sweeps	879 (5.3%)				
Not observed in any later sweep	1569 (9.5%)	401 (3.1%)	1412 (9.8%)	1802 (16.1%)	1269 (14.7%)

Earlier response status of observed sample at each sweep

   NCDS

7.9   Table 7.5 summarises the response at earlier sweeps by showing the total number of earlier sweeps that the observed sample at each sweep has participated in. As these tables are retrospective, there is no uncertainty about the eligibility at an earlier sweep of members of the observed sample at a subsequent sweep. This is because if they are in the observed sample at a given time they cannot have died or permanently emigrated at an earlier time.

7.10   The percentage of the observed sample that participated in all earlier waves declines as the number of earlier sweeps increases. For example, 81.1% of the observed sample at sweep 4 took part in all earlier sweeps compared with 62.5% of the observed sample at sweep 6.

7.11   The table shows that, even at the latest sweep (sweep 6), the clear majority (62.5%) of the observed sample had taken part in all earlier sweeps and only just less than 12% of the observed sample at sweep 6 had missed more than one earlier sweep. But it also illustrates an important point about the structure

of the NCDS sample: members of the longitudinal target sample can and do move in and out of the observed sample over time.

Table 7.5:   NCDS summary of response in earlier sweeps

SWEEP	1	2	3	4	5	6
Observed Sample	15051	14757	13917	12044	10986	10979
Observed in all earlier sweeps	14888 (98.9%)	13657 (92.6%)	12164 (87.4%)	9768 (81.1%)	7760 (70.6%)	6863 (62.5%)
Observed in all but one earlier sweep		1077 (7.3%)	1501 (10.8%)	1857 (15.4%)	2450 (22.3%)	2821 (25.7%)
Observed in all but two earlier sweeps			219 (1.6%)	366 (3.0%	621 (5.7%)	976 (8.9%)
Observed in all but three earlier sweeps				53 (0.4%)	140 (1.3%)	263 (2.4%)
Observed in all but four earlier sweeps					15 (0.1%)	48 (0.4%)
Observed in all but five earlier sweeps						8 (0.1%)
Not observed in any earlier sweeps	163 (1.1%)	23 (0.2%)	33 (0.2%)	0	0	0

BCS70

7.12   As for NCDS, Table 7.6 summarises the response at earlier sweeps for the observed sample at each sweep. Again, there is considerable movement in and out of the target sample over time.

7.13   The table shows that at the latest sweep (sweep 5) only just under half (46.0%) of the observed sample had taken part in all earlier sweeps. However, about 80% of the observed sample at sweep 5 had missed none or only one earlier sweep and a total of 95% had taken part in three or more out of five earlier sweeps.

Table 7.6:   BCS70 summary of response in earlier sweeps

SWEEP	1	2	3	4	5
Observed Sample	12981	14350	11206	8564	10833
Observed in all earlier sweeps	12733 (98.1%)	11749 (81.9%)	8932 (79.7%)	5657 (65.4%)	4984 (46.0%)
Observed in all but one earlier sweep		2174 (15.2%)	1755 (15.7%)	2178 (25.2%)	3739 (34.5%)
Observed in all but two earlier sweeps			479 (4.3%)	667 (7.7%)	1587 (14.6%)
Observed in all but three earlier sweeps				152 (1.8%)	450 (4.2%)
Observed in all but four earlier sweeps					73 (0.7%)
Not observed in any earlier sweeps	248 (1.9%)	427 (3.0%)	40 (0.4%)	0	0

8.   Response Patterns for Domains of Interest

8.1   This section deals with domain response, a topic that falls between unit non-response (the main focus of this paper) and item non-response (not covered here).

8.2   Up to this point, a cohort member is regarded as part of the observed sample at any sweep even if their response is less than complete. It is, for example, possible for data to be missing for one part of the schedule especially as, during the years of childhood, data were obtained from different sources (parents, teachers and medical personnel). The measured variables were therefore split into six domains of interest:

(i) Family background and financial situation of parents, covering the childhood years from birth to sixteen.
(ii) Health, covering the entire lifecourse.
(iii) Housing and neighbourhood, covering the entire lifecourse (although not at sweep 0 for BCS70).
(iv) Education and training, covering all sweeps except sweep 0.
(v) Attitudes and family life, covering the adult (i.e. post-16) sweeps.
(vi) Employment and financial situation, covering the adult (i.e. post-16) sweeps.

8.3   The variables for each domain were chosen to be as similar across the sweeps as possible and are referred to as 'key variables'. Moreover, similar variables were chosen for each of the domains for both NCDS and BCS70 although the different questionnaires meant that priority was given to consistency across sweeps within the studies rather than between the two studies themselves. See Appendix 1 for details of the key variables contained in these domains. The figures, although influenced by the choice of key variables, nevertheless give a general indication of coverage at each sweep.

NCDS

8.4   Table 8.1 sets out the response patterns for each of the six domains. It shows the patterns for having answered 'any' or 'all' of the questions relating to the selected key variables, conditional on being in the longitudinal observed sample. For any questions answered, the response rate is generally over 90% with just two cases - education and housing at sweep 3 - when the response falls to around 80%.

8.5   The response rate is, as expected, lower for all questions answered. Around half have a rate over 90% to a domain in a sweep. In three cases - health and attitudes in sweep 4 and attitudes in sweep 5 - the rate is less than 50%.

Table 8.1:   NCDS domain response, sweeps 0 to 6, longitudinal observed samples.

SWEEP	0	1	2	3	4	5	6
Longitudinal observed sample	17415	15051	14757	13917	12044	10986	10979
Family Situation							
Any	17413	15044	14496	13432			
	(99.9%)	(99.9%)	(98.2%)	(96.5%)			
All	17413	13303	11969	8717			
	(99.9%)	(88.4%)	(81.1%)	(62.6%)			
Health							
Any	17374	14327	13609	12524	12044	10985	10979
	(99.8%)	(95.2%)	(92.2%)	(90.0%)	(100%)	(99.9%)	(100%)
All	16115	13589	11730	8670	5317	10899	10948
	(92.5%)	(90.3%)	(79.5%)	(62.3%)	(44.1%)	(99.2%)	(99.7%)
Housing							
Any	16920	14260	13320	11140	12042	10985	10924
	(97.2%)	(94.7%)	(90.3%)	(80.0%)	(100%)	(100%)	(99.5%)
All	16920	14175	13299	11042	11728	10946	10924
	(97.2%)	(94.2%)	(90.1%)	(79.3%)	(97.4%)	(99.6%)	(99.5%)
Education							
Any		14629	13607	11470	12043	10911	10948
		(97.2%)	(92.2%)	(82.4%)	(100%)	(99.3%)	(99.7%)
All		14482	13603	11406	7342	9358	10948
		(96.2%)	(92.2%)	(82.0%)	(61.0%)	(85.2%)	(99.7%)
Attitudes							
Any					12041	10959	10948
					(100%)	(99.8%)	(99.7%)
All					3058	991	10948
					(25.4%)	(9.0%)	(99.7%)
Employment							
Any					12039	10959	10948
					(100%)	(99.8%)	(99.7%)
All					12039	10790	6574
					(100%)	(98.2%)	(59.9%)

Notes

1. "Any" refers to a response to at least one of the key questions in the domain. "All" refers to a response to all the key questions in the domain.
2. Percentages based on the longitudinal observed samples.
3. Blanks indicate that the variables in the domain were not asked in the sweep.

BCS70

8.6   Table 8.2 sets out the response patterns for each of the six domains for BCS70. Once again it shows the pattern for having answered any and all of the key variable questions. For any questions answered, the response rate is generally over 90% but with six cases when the response is lower, notably education at sweeps 1 and 3. A teachers' strike hampered response at sweep 3.

8.7   The response rate is, as expected, lower for all questions answered. As for NCDS, around half have a rate over 90% to a domain in a sweep. There is a smaller difference between the 'all' and 'any' rates for BCS70 than for NCDS.

Table 8.2:   BCS70 domain response, sweeps 0 to 5, longitudinal observed samples.

SWEEP	0	1	2	3	4	5
Longitudinal observed sample	16571	12981	14350	11206	8654	10833
Family Situation						
Any	16558	12859	13317	9330		
	(99.9%)	(99.1%)	(92.8%)	(83.3%)		
All	10663	12859	8082	4060		
	(64.3%)	(99.1%)	(56.3%)	(36.2%)		
Health						
Any	16535	12969	14099	8369	8647	10781
	(99.8%)	(99.9%)	(98.3%)	(74.7%)	(99.9%)	(99.5%)
All	16535	12383	10039	5328	6986	10548
	(99.8%)	(95.4%)	(70.0%)	(47.5%)	(80.7%)	(97.4%)
Housing						
Any		12961	13336	9225	8589	10708
		(99.8%)	(92.9%)	(82.3%)	(99.2%)	(98.8%)
All		12782	13197	9175	8430	10706
		(98.5%)	(92.0%)	(81.9%)	(97.4%)	(98.8%)
Education						
Any		4997	11685	5616	8152	10778
		(38.5%)	(81.4%)	(50.1%)	(94.2%)	(99.5%)
All		4997	11685	5616	8152	10778
		(38.5%)	(81.4%)	(50.1%)	(94.2%)	(99.5%)
Attitudes						
Any					8648	10778
					(99.9%)	(99.5%)
All					7844	9460
					(90.6%)	(87.3%)
Employment						
Any					8552	10800
					(98.8%)	(99.7%)
All					8552	10788
					(98.8%)	(99.6%)

Notes

1. "Any" refers to a response to at least one of the key questions in the domain. "All" refers to a response to all the key questions in the domain.
2. Percentages based on the longitudinal observed samples.
3. Blanks indicate that the variables in the domain were not asked in the sweep.

9. Cross-sectional Response

NCDS

9.1 A total of 18563 cases have been involved in NCDS at least once. This comprises a cohort of 17634 babies born in GB, and an additional 929 children possibly born outside GB who were added to the survey between sweeps 1 and 3 as immigrants. These 929 children were discovered and added to the NCDS as the cohort was traced through school records up to the age of 16. Therefore, in addition to treating the NCDS as a longitudinal study, the separate sweeps can be used as cross-sectional datasets. As noted in Section 3, no new immigrants were added to the dataset after sweep 3 and so subsequent sweeps suffer from undercoverage. The target sample does not, therefore, decline over time. It fluctuates between sweeps 0 and 3 and then declines from sweep 4 onwards (Table 9.1).

Table 9.1:   NCDS cross-sectional target samples, sweeps 0 to 6.

SWEEP (AGE)	0 (0)	1 (7)	2 (11)	3 (16)	4 (23)	5 (33)	6 (42)
Target sample	17634 (100%)	16727 (92.8%)	16754 (91.6%)	16901 (91.0%)	16482 (88.8%)	16240 (87.4%)	16240 (87.5%)
- of which immigrants		380 (2.1%)	651 (3.6%)	929 (5.0%)	0	0	0
Deaths		812 (4.5%)	829 (4.5%)	862 (4.6%)	883 (4.8%)	986 (5.3%)	1002 (5.4%)
Emigrants		475 (2.6%)	702 (3.8%)	800 (4.3%)	1198 (6.6%)	1337 (7.2%)	1321 (7.1%)
CUMULATIVE TOTAL	17634	18014	18285	18563	18563	18563	18563

Note

The cumulative total is that of the target sample, deaths and emigrants (excluding the immigrants who are already included in the target sample).

9.2 Table 9.2 sets out the differences between the target and the observed cross-sectional samples at each sweep. The observed sample declines as the sweeps progress although more slowly than the longitudinal sample. This may be due to the inclusion of immigrants up to sweep 3.

Table 9.2:   NCDS cross-sectional target and observed samples, sweeps 0 to 6.

SWEEP (AGE)	0 (0)	1 (7)	2 (11)	3 (16)	4 (23)	5 (33)	6 (42)
Observed sample	17415 (98.8%)	15425 (92.2%)	15337 (91.5%)	14647 (86.7%)	12537 (76.1%)	11407 (70.3%)	11419 (70.3%)
Non-response	219 (1.2%)	80 (0.5%)	797 (4.8%)	1152 (6.8%)	1161 (7.0%)	1806 (11.1%)	2135 (13.1%)
Uncertain eligibility		1222 (7.3%)	620 (3.7%)	1102 (6.5%)	2784 (16.9%)	3027 (18.6%)	2686 (16.5%)
Target Sample	**17634**	**16727**	**16754**	**16901**	**16482**	**16240**	**16240**

Notes

1. Percentages are of the target sample which excludes those who have emigrated (both permanently and temporarily) and those who have died, but includes those who have immigrated and have been found through schools.
2. Uncertain eligibility. This group is larger than the corresponding group in Table 6.1 as it includes apart from sweep 0, those labelled 'Non-response: other' in Table 6.1. The eligibility of this group is less certain for cross-sectional target samples than it is for longitudinal samples.

9.3   Table 9.3 sets out the response rates by sweep. As with the longitudinal sample the response rates decline with age, gently during the ages of childhood but with a marked shift downwards as the cohort members become adults.

Table 9.3:   NCDS cross-sectional response rates, sweeps 0 to 6.

SWEEP (AGE)	0 (0)	1 (7)	2 (11)	3 (16)	4 (23)	5 (33)	6 (42)
Eligibility rate (ER)	100%	92.3%	91.3%	90.5%	86.8%	85.1%	85.4%
Response rate - estimated (RR)	98.8%	92.7%	91.8%	87.2%	77.8%	72.3%	72.1%
Response rate – upper limit (URR)	100%	99.5%	95.1%	92.7%	91.5%	86.3%	84.3%
Response rate – lower limit (LRR)	98.8%	92.2%	91.5%	86.7%	76.1%	70.2%	70.3%

Note
For definitions of eligibility and response rates see notes after Table 6.2.

BCS70

9.4　A total of 18126 cases have been involved in BCS70 at least once. (This number excludes the 605 children born in Northern Ireland who did not move to GB up to sweep 3.) This comprises a cohort of 17287 babies born in GB, and an additional 839 children possibly born outside GB who were added to the survey between sweeps 1 and 3 as immigrants. As with NCDS, these 839 children were discovered and added to BCS70 as the cohort was traced through school records up to the age of 16. There were 127 children who were discovered at sweep 3 (age 16) but who were interviewed for the first time at sweep 4 (79 cases) or sweep 5 (48 cases). These cases are included in the non-response category for sweep 3 (and sweep 4 if applicable).

Table 9.4:　BCS70 cross-sectional target and observed samples, sweeps 0 to 5: available data

SWEEP (AGE)	0 (0)	1 (5)	2 (10)	3 (16)	4 (26)	5 (30)
Target sample	17287 (100%)	16810 (96.4%)	17275 (96.7%)	17529 (96.7%)	17329 (95.6%)	17050 (94.1%)
- of which immigrants		132 (0.8%)	575 (3.2%)	839 (4.7%)	0	0
Deaths		567 (3.3%)	587 (3.3%)	597 (3.3%)	700 (3.9%)	753 (4.2%)
Emigrants		42 (0.2%)	0	0	97 (0.5%)	323 (1.8%)
CUMULATIVE TOTAL	**17287**	**17419**	**17862**	**18126**	**18126**	**18126**

Notes

1.　The cumulative total is that of the target sample, deaths and emigrants (excluding the immigrants who are already included in the target sample).
2.　The immigrants include Northern Ireland immigrants to Great Britain.

Table 9.5:   BCS70 cross-sectional target and observed samples, sweeps 0 to 5: emigration estimated

SWEEP (AGE)	0 (0)	1 (5)	2 (10)	3 (16)	4 (26)	5 (30)
Target sample	17287 (100%)	16381 (94.0%)	16586 (92.9%)	16750 (92.4%)	16266 (89.7%)	16068 (88.6%)
- of which immigrants		132 (0.8%)	575 (3.2%)	839 (4.9%)	0	0
Deaths		567 (3.3%)	587 (3.3%)	597 (3.3%)	700 (3.9%)	753 (4.2%)
Emigrants		471 (2.7%)	689 (3.9%)	770 (4.2%)	1160 (6.4%)	1305 (7.2%)
CUMULATIVE TOTAL	**17287**	**17419**	**17862**	**18126**	**18126**	**18126**

Notes

1. The cumulative total is that of the target sample, deaths and emigrants (excluding the immigrants who are already included in the target sample).
2. The immigrants include Northern Ireland immigrants to Great Britain.

9.5   The emigration rates estimated in Table 9.5 are essentially those found for NCDS. It is likely that the estimated target sample in Table 9.5 more accurately reflects the true situation than Table 9.4 but it does depend on the assumption that emigration rates did not change substantially between the two cohorts.

9.6   Table 9.6 sets out the differences between the target and the observed samples at each sweep. In contrast with NCDS, BCS70 observed cross-sectional samples fluctuate across sweeps just as the longitudinal BCS70 samples do.

Table 9.6:   BCS70 cross-sectional target and observed samples, sweeps 0 to 5.

SWEEP (AGE)	0 (0)	1 (5)	2 (10)	3 (16)	4 (26)	5 (30)
Observed sample	16571 (95.9%)	13071 (79.8%)	14874 (89.7%)	11621 (69.4%)	9003 (55.3%)	11261 (70.1%)
Non-response	716 (4.1%)	0	0	0	146 (0.9%)	1939 (12.1%)
Uncertain eligibility	0	3310 (20.3%)	1712 (10.3%)	5129 (30.6%)	7117 (43.8%)	2868 (17.8%)
Target Sample (estimated)	**17287**	**16381**	**16586**	**16750**	**16266**	**16068**

1. Percentages are of total eligible which excludes those who have emigrated (both permanently and temporarily) and those who have died but includes those who have immigrated and have been found through the schools.
2. Uncertain eligibility. This group is larger than the corresponding group in Table 6.3 as it includes apart from sweep 0, those labelled 'Non-response' in sweeps 1 to 3 and the great majority of those labelled 'Non-response' in sweep 4. The eligibility of the non-response group is less certain for cross-sectional target samples than it is for longitudinal samples.

9.7   Table 9.7 sets out the response rates by sweep. As with the longitudinal sample the response rates follow the same pattern and fluctuate across the sweeps. The response rate is especially low in sweep 4 (see 6.4).

Table 9.7:   BCS70 cross-sectional response rates, sweeps 0 to 5.

SWEEP (AGE)	0 (0)	1 (5)	2 (10)	3 (16)	4 (26)	5 (30)
Eligibility rate (ER)	100%	92.6%	92.1%	89.4%	83.1%	86.0%
Response rate - estimated (RR)	95.9%	81.0%	90.4%	71.7%	59.8%	74.7%
Response rate – upper limit (URR)	95.9%	100%	100%	100%	98.4%	89.3%
Response rate – lower limit (LRR)	95.9%	79.8%	89.7%	69.4%	55.3%	72.8%

Note

For definitions of eligibility and response rates see notes after Table 6.2.

10.   Conclusion

10.1   As stated in the Introduction, this report is descriptive and not prescriptive. It provides users of both datasets with an account of the sizes of the different samples and how and why they have changed as the cohort members have aged. This information is, we believe, a crucial component of any analysis plans. The report does not, however, say anything about the changing characteristics of the samples over time (described by Hawkes and Plewis (2004) for NCDS), nor does it attempt to answer the important question of whether and how researchers should incorporate corrections for non-response into their analyses.

10.2   The distinction between the longitudinal and cross-sectional populations and samples is an important one and Tables 10.1 and 10.2 summarise the numbers of respondents for the two kinds of observed samples. The cross-

sectional observed sample is always larger than the corresponding longitudinal sample.

Table 10.1: NCDS longitudinal and cross-sectional observed samples, sweeps 0 to 6.

SWEEP (AGE)	0 (0)	1 (7)	2 (11)	3 (16)	4 (23)	5 (33)	6 (42)
Longitudinal observed sample	17415	15051	14757	13917	12044	10986	10979
Cross-sectional observed sample	17415	15425	15337	14647	12537	11407	11419

Table 10.2: BCS70 longitudinal and cross-sectional observed samples, sweeps 0 to 5.

SWEEP (AGE)	0 (0)	1 (5)	2 (10)	3 (16)	4 (26)	5 (30)
Longitudinal observed sample	16571	12981	14350	11206	8654	10833
Cross-sectional observed sample	16571	13071	14874	11621	9003	11261

10.3 The response rates in this report are necessarily estimates, relying as they do on assumptions about eligibility and also about losses from the target sample as a result of deaths and, especially, emigration.

10.4 The response rates are generally high during the childhood years and remain above 70% in adulthood. The rates for BCS70 are somewhat lower, notably for sweeps 1, 3 and 4. Sweeps 1 and 3 were handicapped by limited funding and a teachers' strike and sweep 4 was a postal survey for which it is more difficult to achieve high response. Nevertheless, by sweep 5, the response rate at age 30 for BCS70 (71.5%) was almost identical to that for NCDS at age 33 (71.6%).

10.5 A further complete sweep for NCDS is taking place in 2004 (age 46) as a telephone survey. At the same time the BCS70 cohort is being followed up at age 34 with a personal interview, together with a sample of the cohort members' children. These data – sweep 7 for NCDS and sweep 6 for BCS70 – should be available in 2005. This document will be updated in due course to reflect information gathered at these sweeps.

Endnotes

1.   Because the interval between birth and sweep 1 at age seven was so wide, and because NCDS was not originally set up as a longitudinal study, and also because data collection at age seven did not depend on there being data at birth, there is an argument for regarding NCDS as a longitudinal study that starts at age seven with the data collected at birth treated as retrospective data. The longitudinal target sample would then be defined differently:

> *all children born (alive or dead) in a specific week in March 1958 and living in Great Britain at age seven, until they died or permanently emigrated from Great Britain.*

2.   This definition does include immigrants to Great Britain after birth and before age 16 but not thereafter.

It would be possible to amend the definitions at sweep 0 to:

> *'The cross-sectional populations and target samples for inference at sweep 0 are all births (including stillbirths) in the UK in 1970 (population) or in a specific week in April 1970 (target sample).'*

In other words, births in Northern Ireland could be included here.

3.   The NCDS target sample declines more slowly if the study is deemed to start at sweep 1, essentially because the death rate is relatively high in infancy and then falls. The target sample at sweep 1 is then higher than in Table 4.1 because some immigrants (those arriving between birth and age seven and responding to the survey) are included (table available on request).

Appendix 1: Selected Key Variables

Actual variable names from the datasets are provided along with the sweep in which that variable occurs in parentheses.

Family background & financial situation of parents	
NCDS	**BCS70**
- number of children / family members n504 (0), n419 (1), n116 (2), n2360 (3) - parents wish child to stay past school leaving age n116 (1), n1137 (2), n2407 (3) - mother's interest in child schooling n44 (1), n852 (2), n2325 (3) - in care n132 (1), n1133 (2), n2379 (3)	- father's age left education a0010 (0), T7.1 (3) - mother's age left education a0009 (0), T7.2 (3) - father's employment status a0015 (0), c4.1a (2), T8.2 (3) - mother's employment status a0019 (0), e205 (1), c4.2a (2), T8.1 (3) - household income c9.1-c9.8 (2), oe2 (3)

Health	
NCDS	**BCS70**
- height (of cohort member except at birth then of mother of cohort member) n510 (0), n332 (1), n1510 (2), n1949 (3), n5752 (4), n504731 (5), height2 (6) - weight n574 (0), n337 (1), n1515 (2), n1953 (3), n5755 (4), n504734 (5), weight2 (6) - hospital – overnight (tonsils) n246 (1), n1353 (2), n2589 (3), n5854 (4), n504214 (5), hospital (6) - hospital – road accident n250 (1), n1286 (2), n2566 (3), n5819 (4), n504434 (5), accidan1 (6) - handicapped/disabled n401 (1), n1507 (2), n2018 (3), n4139 (4), n504121 (5), lsiany (6)	- height f102 (1), meb17 (2), rd2.1/ha1.2 (3), b960433-b960437 (4), height2 (5) - weight (birth weight at five) a0278 (0), e01a (1), meb19.1 (2), rd4.1/ha1.1 (3), b960439-b960435 (4), weight (5) - smoking status k051 (2), gh1 (3), b960632 (4), smoking (5)

Housing & neighbourhood	
NCDS	**BCS70**
- tenure n200 (1), n1152 (2), n2471 (3), n5029 (4), n502913 (5), tenure (6) - number of rooms (persons per room at birth) n512 (0), n201 (1), n1156 (2), n2476 (3), n5323 (4), n502947 (5), numrooms (6)	- tenure e220 (1), d2 (2), of3.1-of3.7 (3), b960421 (4), tenure2 (5) - number of rooms d5 (2), of4.1 (3), b960422 (4), numrooms (5)

Education & training	
NCDS	**BCS70**
- mathematics test n90 (1), n926 (2), n2930 (3) - reading test n92 (1), n923 (2), n2928 (3) - number of O levels (or equivalent vocational qualifications) n4655 (4), n501441 (5), edquals2 (6) - number of A levels (or equivalent qualifications) n4657 (4), n501513 (5)	- age left full time education b960131/b960132 (4), agelfte2 (5) - reading/spelling test c7a1-c7a100 (3) - reading/writing test f109 (1), tests10 (2)

Attitudes & family life	
NCDS	**BCS70**
- number of children n4925 (4), n502013 (5), everpreg (6) - living with partner n5031 (4), n501713 (5), marstat (6) - voting n5959 (4), n504635 (5), vote97 (6) - religion n5967 (4), n504651 (5), religion (6)	- law obeyed b960120 (4), a1 (5) - job better than unemployment b960121 (4), we1 (5) - death penalty b960123 (4), a2 (5) - political party no differences b960125 (4), pc2 (5) - divorce is too easy b960126 (4), mor1 (5) - job pack in b960127 (4), we2 (5) - marriage for life b960128 (4), mor4 (5) - politicians b960424 (4), pc3 (5) - censorship b960425 (4), a3 (5) - political no benefit b960427 (4), pc1 (5) - criminal sentencing b960428 (4), a4 (5) - couples with children b960429 (4), mor3 (5) - government redistribution b960430 (4), lr6 (5) - interest in politics b960661 (4), politint (5) - who voted for at last election vote97 (5) - who vote for b960662 (4), votenow (5) - religion b960663 (4), religion (5)

Employment & financial situation	
NCDS	**BCS70**
- ever employed/economic status n4143 (4), n500520 (5), econact (6) - ever unemployed n4715 (4) n500866 (5), activity (6) - on benefits n5513 (4), n503313 (5), bencod01 (6)	- economic status b960259-b960269 (4), econact (5) - year/month started current main activity cstartyear-cstartmo (5) - benefit receipt bencod01 (5)

References

American Association for Public Opinion Research (2000), *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys.* Ann Arbor, Michigan: AAPOR.

Hattersley, L. and Creeser, R. (1995), *Longitudinal Study 1971-1991.* OPCS Series LS No. 7. London: HMSO.

Hawkes, D. and Plewis, I. (2004), Modelling non-response in the National Child Development Study (submitted for publication).

Nathan, G. (1999), *A Review of Sample Attrition and Representativeness in Three Longitudinal Surveys.* Government Statistical Service Methodology Series No. 13.