# Millennium Cohort Study

## Technical Report on Response

Third Edition

June 2010

Edited by
**Sosthenes C. Ketende**

With contributions from John McDonald, Heather Joshi, Lisa Calderwood and the MCS team



**Centre for Longitudinal Studies**
Institute of Education, University of London

The Centre for Longitudinal Studies (CLS) was one of five centres that comprise the Bedford Group for Lifecourse and Statistical Studies at the Institute of Education, University of London. It is currently a department within the Faculty of Policy and Society of the Institute of Education. The department houses an ESRC Resource Centre devoted to the collection, management and analysis of large-scale longitudinal data. It is the home of three internationally-renowned birth cohort studies: the 1958 National Child Development Study (NCDS), the 1970 British Cohort Study (BCS) and the Millennium Cohort Study (MCS).

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, or the consortium of government departments which also contribute to the costs of the Millennium Cohort Study. All errors and omissions remain those of the authors.

# Contents

## Acknowledgments

## Preface to the Third Edition

This is the third edition of the technical report on response in the Millennium Cohort Study as it progresses from sweep 4 onwards. This edition includes material on sweep 2, 3 and 4 and covers the evolution of the population and sample after sweep 1, allowing for the families who were included in the study for the first time at sweep 2. It presents response rates for sweeps 2, 3 and 4. It also covers the correlates of different types of unit non-response at sweep 3 and 4, and gives references to published articles detailing correlates of different types of non-response at sweep 2. This report does not cover item non-response, because response is very high in nearly all items except income. Family income missing data and new derived variables will be covered in a separate paper. Sections 1 to 8 covers material in the first  and second edition of this series with minor changes mainly typo corrections. Sections 9 to 12 covers the new materials.

# 1. Introduction

1.1　The longitudinal population for the Millennium Cohort Study is defined in Plewis (2004, para. 2.1) as:

> 'all children born between 1 September 2000 and 31 August 2001 (for England and Wales), and between 23 November 2000 and 11 January 2002 (for Scotland and Northern Ireland), alive and living in the UK at age nine months, eligible to receive Child Benefit at that age, and for as long as they remain living in the UK at the time of sampling'.

1.2　The longitudinal population declines with time to the extent that children die or permanently emigrate from the UK.

1.3　The longitudinal target samples for the second and subsequent sweeps of the Millennium Cohort Study (MCS) consist of those children in the 398 selected UK wards who were eligible to be included in MCS1 and who did not die or permanently emigrate between the first and subsequent sweeps of MCS.

1.4　Sections 2 and 3 of this document contains unchanged materials as appeared in the first edition. These are mainly about response at sweep 2. The new Section 4 is an additional section providing references to published articles dealing with prediction of response at sweep 2. Sections 5 through 8 cover sweep 3 material. Sections 8 through 11 cover new sweep 4 material.


# 2. The sample for sweep 2 (MCS2)

2.1　The issued sample at MCS2 differed from the longitudinal target sample because:

(1)　It *excluded* all children who were not issued to the field in MCS1 (see Plewis, 2004, Table 7.1).
(2)　It *included* only those children from the issued sample at MCS1 who were classified as productive then (see Plewis, 2004, Table 7.2).
(3)　It could have *included* children who, by virtue of death or emigration, turned out to be ineligible.

2.2　The sample for MCS1 was constructed from Child Benefit records and, in principle, included all eligible children aged nine months living in the selected electoral wards. The Child Benefit address records are not, however, up to date partly because, for many parents, Child Benefit is paid directly into a bank account and so the Inland Revenue (– now HMRC - who administered Child

Benefit) has no need to contact these parents on a regular basis. Another reason could be the delays in establishing a benefit record for recent immigrants. Thus, some children who had moved into the sampled areas after their initial registration on the Child Benefit system were not picked up for MCS1 at age nine months. Some mitigation of this problem was achieved for about half the sweep one sample by carrying out an extra scan of the Child Benefit records (see Plewis, 2004, para. 6.4). These so-called *new mover families* had somewhat different characteristics from the rest of the sample. They were more likely to be on means-tested benefits and had other characteristics that marked them as more disadvantaged than the rest (Plewis, 2004, para. 10.5 and Appendix 3). A second analysis (para. 10.6) indicated that time at current address was also related to socio-demographic variables.

2.3   The rhythms of the Child Benefit system prevented a full repair of the sample at sweep one; it was clear that those families who had moved out of the sampled areas before the cohort child was nine months old had not been replaced in sufficient numbers by those moving in. Consequently, Inland Revenue were funded to rescan the Child Benefit register to pick up eligible children who, as a result of late notifications to the system, were indeed found to have been living in a sampled ward at age nine months. This repair was, however, restricted to England. The families so found are referred to here as *'new families'*: they are part of the target samples for each sweep but were not part of the issued sample at MCS1.

2.4   Table 2.1 shows how the 'new families' sample changed from the point at which it was provided to CLS by the Inland Revenue to being issued to the field. Some comparisons are drawn with results from MCS1.

Table 2.1:   From Child Benefit Sample to Issued Sample by Stratum ('new families'; England only)

|  | Advantaged | Disadvantaged | Min. Ethnic. | Total |
|---|---|---|---|---|
| Child Benefit sample (1) | 511 | 748 | 516 | 1775 |
| Exclusions, CLS | 48 | 28 | 28 | 104 |
| Child Benefit sample (2) | 463 | 720 | 488 | 1671 |
| Exclusions, IR | 35 (7.6%) | 101 (14%) | 70 (14%) | 206 (12%) |
| Opt-outs | 20 (4.3%) | 38 (5.3%) | 18 (3.7%) | 76 (4.5%) |
| **Issued sample (to field)** | **408 (88%)** | **581 (81%)** | **400 (82%)** | **1389 (83%)** |

<u>Notes on Table 2.1</u>

1) *Exclusions, CLS*
   These were mainly concentrated in waves 1 to 3 (n = 35) and in wave 13 (n = 53) and were cases that had already been issued in MCS1. See Plewis (2004, para. 6.4) for a description of 'wave'.

2) *Exclusions, Inland Revenue*
   As in MCS1 (see Table 7.1 in Plewis, 2004) but with an additional group – families already involved in a survey for the national evaluation of Sure Start (n = 76). The exclusion rate for the 'new families' sample shown in Table 2.1 is much higher than for the MCS1 sample for all strata – not wholly accounted for by the omission of the 'Sure Start' survey families - but the opt-out rate is lower.

3) There were 245 fewer families issued from waves 8 to 13 than for waves 1 to 7. This was to be expected because MCS1 did include the 'new movers' from wave 8 onwards, 179 of whom were issued in England. The 'new movers' and the 'new families' can reasonably be considered together as a single category in MCS2, at least for England.

2.5 There were originally 18553 productive families in MCS1 (Plewis, 2004, Table 7.2) but one case was subsequently withdrawn as it was discovered to have been invalid. The number of 'new families' was reduced from 1671 found by the Inland Revenue to the 1389 families that were issued to the field in MCS2 (through the exclusions and opt-outs shown in Table 2.1). Therefore, the issued sample for MCS2 was 19941: 18552 were productive families in MCS1 and 1389 were 'new families' (although 71 of the MCS1 productive families were not, in fact, issued to the field for various reasons such as death, emigration and refusal).

2.6 Data collection for MCS2 was carried out between September 2003 and April 2005 for England and Wales and between December 2003 and April 2005 for Scotland and Northern Ireland.

2.7 Table 2.2 gives the cohort child's age when the interview with the main respondent was carried out. It shows that 78% of the interviews were completed the cohort members were within the target window of 36 to 39 months of age, 11% just before that window opened and a further 11% after it closed, in a few cases quite substantially later, even beyond the fourth birthday.

**Table 2.2: Distribution of cohort member's age at MCS2**

| Age (Months) | n | % |
|---|---|---|
| 31-34 | 10 | 0.063 |
| 35 | 1756 | 11 |
| 36 | 6802 | 43 |
| 37 | 3294 | 21 |
| 38 | 1506 | 9.5 |
| 39 | 731 | 4.6 |
| 40 | 410 | 2.6 |
| 41 | 267 | 1.7 |
| 42 | 179 | 1.1 |
| 43 | 158 | 1.0 |
| 44 | 140 | 0.89 |
| 45 | 149 | 0.94 |
| 46 | 104 | 0.66 |
| 47 | 102 | 0.65 |
| 48-54 | 191 | 1.2 |
| **Total number of children** | **15799** | **100** |

Note: Interview date is missing for 9 cases.

# 3. Response at sweep 2

3.1  Table 3.1 shows that 78% (15590/19941) of the issued sample at MCS2 were productive. The refusal rates (REF1) are lower in Wales and Scotland than in England and, notably, than in Northern Ireland. The rates for England are somewhat higher because of the high refusal rates for 'new families' (Table 3.3).

3.2  Tables 3.2 and 3.3 have the same layout as Table 3.1 except that they are confined to England; they separate the response for the MCS1 productives from that of the 'new families'. The eligibility rates for 'new families' are lower than for the MCS1 productives but similar to those obtained at MCS1 for England (Plewis, 2004, Table 7.2).

**Table 3.1: From Issued Sample to Final Sample by Stratum and Country: MCS2**

| | England | | | | Wales | | | Scotland | | | N. Ireland | | | UK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Ad. | Disad. | Eth. | Total | Ad. | Disad | Total | Ad. | Disad. | Total | Ad. | Disad. | Total | All |
| Issued sample | 5025 | 5103 | 2794 | 12922 | 832 | 1928 | 2760 | 1145 | 1191 | 2336 | 723 | 1200 | 1923 | **19941** |
| Ineligible | 101 | 55 | 39 | 195 | 8 | 8 | 16 | 22 | 11 | 33 | 5 | 6 | 11 | **255** |
| ELIG | 97.8% | 98.8% | 98.5% | 98.4% | 99.0% | 99.6% | 99.4% | 98.0% | 99.0% | 98.5% | 99.3% | 99.5% | 99.4% | **98.6%** |
| Uncertain eligibility | 155 | 295 | 178 | 628 | 19 | 80 | 99 | 28 | 58 | 86 | 11 | 44 | 55 | **868** |
| Unproductive | 560 | 873 | 616 | 2049 | 113 | 271 | 384 | 163 | 240 | 403 | 121 | 271 | 392 | **3228** |
| REF1 | 8% | 10% | 13% | 10% | 9% | 8% | 8% | 9% | 8% | 9% | 14% | 17% | 16% | **10%** |
| **Productive** | **4209** | **3880** | **1961** | **10050** | **692** | **1569** | **2261** | **932** | **882** | **1814** | **586** | **879** | **1465** | **15590** |

Notes on Tables 3.1 to 3.3

*Issued sample*
See para. 2.5. For Wales, Scotland and Northern Ireland these are the productive families from MCS1 (Plewis, 2004, Table 7.2).

*Ineligible*
Child deaths (n = 16); emigrants (n = 169); failed eligibility (n = 70; 'new families' only).

*ELIG*
ELIG is the eligibility rate of the issued sample. This is the ratio of cases known or estimated to be eligible to all issued cases as defined by Lynn et al. (2001). The estimated eligibility rates in Tables 3.2 and 3.3 are used for cases with uncertain eligibility.

*Uncertain eligibility*
This includes untraced movers out of the MCS1 productives (who might have died or emigrated) and 'new families' not found in the field.

*Unproductive*
This includes three sub-groups of cases with their UK sizes in brackets:

- (i) Non-contact (1070).
- (ii) Refusal (2002).
- (iii) Other non-response (156).

*REF1*
REF1 is the refusal rate in the field. This is the ratio of refusals to all issued cases known or estimated to be eligible as defined by Lynn et al. (2001). Estimated eligibility rates are used for cases with uncertain eligibility as explained above.

*Productive*
All families with some data (from at least one instrument - Main, partner, proxy, British Ability Scales (BAS), Bracken Basic Concept Scale, height, weight) other than data carried forward from MCS1.

**Table 3.2: From Issued Sample to Final Sample by Stratum: MCS1 productives, England.**

| | England | | | |
|---|---|---|---|---|
| | **Ad.** | **Disad.** | **Eth.** | **Total** |
| Issued sample | 4617 | 4522 | 2394 | 11533 |
| Ineligible | 62 | 26 | 19 | 107 |
| ELIG | 98.6% | 99.4% | 99.2% | 99.0% |
| Uncertain eligibility | 103 | 221 | 123 | 447 |
| Unproductive | 454 | 679 | 488 | 1621 |
| REF1 | 7% | 9% | 12% | 9% |
| **Productive** | **3998** | **3596** | **1764** | **9358** |

**Table 3.3:   From Issued Sample to Final Sample by Stratum: 'new families', England.**

| | England | | | |
|---|---|---|---|---|
| | **Ad.** | **Disad.** | **Eth.** | **Total** |
| Issued sample | 408 | 581 | 400 | 1389 |
| Ineligible | 39 | 29 | 20 | 88 |
| ELIG | 89% | 94% | 94% | 92% |
| Uncertain eligibility | 52 | 74 | 55 | 181 |
| Unproductive | 106 | 194 | 128 | 428 |
| REF1 | 21% | 21% | 19% | 20% |
| **Productive** | **211** | **284** | **197** | **692** |

3.3   Table 3.4 documents the decline from an initial sample of 28927 at MCS1 (including 'new families') to 15590 productive families at MCS2.

**Table 3.4: From Initial Sample, MCS1 to Final Sample, MCS2 by Stratum and Country**

| | England | | | | Wales | | | Scotland | | | NI | | | UK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Ad. | Disad. | Eth. | Total | Ad. | Disad. | Total | Ad. | Disad. | Total | Ad. | Disad. | Total | All |
| Initial Sample | 6859 | 7406 | 4400 | 18665 | 1076 | 2787 | 3863 | 1581 | 1750 | 3331 | 1109 | 1959 | 3068 | **28927** |
| Ineligible | 288 | 248 | 142 | 678 | 27 | 85 | 112 | 80 | 75 | 155 | 24 | 46 | 70 | **1015** |
| Uncertain Eligibility | 545 | 1048 | 695 | 2288 | 72 | 307 | 379 | 115 | 203 | 318 | 134 | 328 | 462 | **3447** |
| Unproductive | 1817 | 2230 | 1602 | 5649 | 285 | 826 | 1111 | 454 | 590 | 1044 | 365 | 706 | 1071 | **8875** |
| **Productive, MCS2** | **4209** | **3880** | **1961** | **10050** | **692** | **1569** | **2261** | **932** | **882** | **1814** | **586** | **879** | **1465** | **15590** |

Notes on Table 3.4

*Initial sample*
The sum of the initial sample from Table 7.3 in Plewis (2004) plus, for England, the Child Benefit sample ('New families') from Table 2.1.

*Ineligible*
From Table 7.3 (Plewis, 2004) and Table 3.1.

*Uncertain eligibility*
From Table 7.3 (Plewis, 2004), Table 3.1 and the exclusions (Table 3.3) from the 'new families' sample.

*Unproductive*
From Table 7.3 (Plewis, 2004), Table 3.1 and the opt-outs (Table 3.3) from the 'new families' sample.

*Productive*
As in Table 3.1.

**Table 3.5:      Response Rates (%) by Stratum and Country**

| | England | | | | Wales | | | Scotland | | | NI | | | UK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Ad. | Disad. | Eth. | Total | Ad. | Disad. | Total | Ad. | Disad. | Total | Ad. | Disad. | Total | All |
| RR1 | 65 | 56 | 48 | 58 | 67 | 60 | 62 | 63 | 54 | 59 | 56 | 48 | 51 | **58** |
| RR2 | 86 | 77 | 71 | 79 | 84 | 82 | 82 | 83 | 75 | 79 | 82 | 74 | 77 | **79** |
| RR2 (W) | 86 | 77 | 71 | 81 | 84 | 82 | 83 | 83 | 75 | 80 | 82 | 74 | 79 | **81** |
| CON | 94 | 87 | 85 | 90 | 93 | 90 | 91 | 93 | 87 | 90 | 97 | 92 | 94 | **90** |
| COOP | 91 | 88 | 84 | **88** | 90 | 90 | **90** | 89 | 86 | **87** | 84 | 80 | **82** | **88** |

Notes on Table 3.5

*RR1*      This is the *overall* response rate to the study at MCS2 defined as the ratio of productive cases to all cases in the initial sample for the study known or estimated to be eligible (Lynn et al., 2001). An overall eligibility rate of 74% was used.

*RR2*      This is the response rate to the study at MCS2, based on Tables 3.1 to 3.3 and defined as  the ratio of productive cases to all cases known or estimated to be eligible in the MCS longitudinal sample.

RR2W   This is the weighted version of RR2, allowing for varying selection probabilities across strata (see Plewis, 2004, Table 5.5). It differs from RR2 only for the four countries and the UK as a whole.

CON     This is the contact rate – the proportion of all cases in which a household member was reached by the interviewer with whom there was contact in person.

COOP    This is the cooperation rate – the number of productive cases as a proportion of cases who were contacted during the fieldwork period

3.4 Table 3.5 gives a series of response rates like those in Table 7.4 in Plewis (2004). It shows that the overall response rate across the two sweeps (RR1) is 58% for the UK as a whole, ranging from 48% in the English minority ethnic wards and the disadvantaged wards in Northern Ireland to 67% in the advantaged wards in Wales.

3.5 The response rates (RR2) and contact rates (CON) for MCS2 were slightly lower than those achieved in MCS1 but the cooperation rates (COOP) were higher.

# 4. Predictors of non-response at sweep 1 and 2

4.1 Variables used for the prediction of response at sweep 1 and 2 were used as a starting point in the prediction of response at sweep 3 (see Section 7.6). A majority of these variables were found to still be good predictors of non-response at sweep 3.

4.2  Although sweep 1 was the first sweep of the MCS, using variables from the Child Benefit system measured at individual level; we were able to predict losses from the child benefit sample using a series of logistic regression models. Predictors of non-response at sweep 1 from these analyses were presented in Plewis (2007b), available from the CLS website (http://www.cls.ioe.ac.uk/studies.asp?section=00010002000100040006 ).

4.3 Predictors of non-response at sweep 2 have been presented in two published papers. These are Plewis (2007a) and Plewis et al. (2008). As stated in 4.1 above, most of these variables are still predictors of non-response at sweep 3 in Table 7.1.

# 5.  Response at sweep 3 (MCS 3)

5.1 Table 5.1 shows that 79.2% (15246/19244) of MCS families who have participated in at least one previous survey were productive at sweep 3. The refusal rates (REF1) range from 12% in England to 14% in Northern Ireland. These rates are slightly higher in all UK countries from sweep 2 levels except in Northern Ireland where the rate has decreased from 16% at sweep 2.

5.2 Table 5.2 shows the decline from the initial sample of 28,927, which includes the 'new families' to 15,246 productive families at sweep 3. This is only 344 fewer productive families than sweep 2

5.3 The 19,244 'issued sample' in Table 5.1 was not the actually number issued to the field. The number issued to the field was 18,526 or 718 cases fewer ( see Calderwood et al.,2008)

Table 5.1: From Issued Sample to Final Sample by Country and Stratum: MCS3

| | England | | | | Wales | | | Scotland | | | Northern Ireland | | | UK |
| | Adv. | Dis. | Ethn. | Total. | Adv | Dis. | Total | Adv. | Dis. | Total | Adv. | Dis. | Total | All |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Issued sample** | **4828** | **4806** | **2591** | **12225** | **832** | **1928** | **2760** | **1145** | **1191** | **2336** | **723** | **1200** | **1923** | **19244** |
| Ineligible | 112 | 52 | 28 | 192 | 11 | 18 | 29 | 39 | 20 | 59 | 12 | 8 | 20 | 300 |
| ELIG (%) | 97.6 | 98.9 | 98.9 | 98.4 | 98.7 | 99.0 | 98.9 | 96.5 | 98.3 | 97.4 | 98.3 | 99.3 | 98.9 | 98.4 |
| **Uncertain eligibility** | 72 | 160 | 135 | 367 | 11 | 53 | 64 | 30 | 42 | 72 | 11 | 33 | 44 | 547 |
| **Unproductive** | 575 | 835 | 539 | 1949 | 141 | 345 | 486 | 159 | 232 | 391 | 106 | 219 | 325 | 3151 |
| REF1 (%) | 10 | 12 | 14 | 12 | 13 | 13 | 13 | 12 | 14 | 13 | 13 | 14 | 14 | 12 |
| **Productive** | **4069** | **3759** | **1889** | **9717** | **669** | **1512** | **2181** | **917** | **897** | **1814** | **594** | **940** | **1534** | **15246** |

Notes on tables 5.1

*Adv.:*　　　　　　　　Advantaged ward, *Dis:* Disadvantaged ward , *Ethn:* Ethnic minority ward

*Issued sample:*　　　These are the productive families from sweep 1plus productive 'new families' at sweep 2.

*Ineligible :*　　　　　Child deaths (n=18) and permanent emigrants (n=282)

*Uncertain eligibility*: Untraced movers: 547 cases, among the 'new families' and sweep 1 productives (who might have died
　　　　　　　　　　or emigrated)

*Unproductive:*　　　*T*hree sub-groups of cases with their UK sizes in brackets: (i) Non-contact (63).(ii) Refusal  (2798).
　　　　　　　　　　(iii) Other non-response (290).

*REF1*　　　　　　　　Refusal rate in the field. This is the ratio of refusals to all issued cases known or estimated to be eligible
　　　　　　　　　　as defined by Lynn et al. (2001). Estimated eligibility rates are used for cases with uncertain eligibility as
　　　　　　　　　　explained above.

*Productive:*　　　　All families with some data (from at least one instrument – Main- interview or self completion, partner,
　　　　　　　　　　proxy and child measurements) other than data carried forward from MCS3.

15

Table 5.2: From Initial Samples to Final Sample at MCS3, by Country and Stratum

| | England | | | | Wales | | | Scotland | | | Northern Ireland | | | UK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Adv. | Dis. | Ethn. | Total | Adv | Dis. | Total | Adv | Dis. | Total | Adv. | Dis. | Total | All |
| **Initial Sample** | 6859 | 7506 | 4400 | 18665 | 1076 | 2787 | 3863 | 1581 | 1750 | 3331 | 1109 | 1959 | 3068 | 28927 |
| **Ineligible** | 299 | 245 | 131 | 675 | 30 | 95 | 125 | 97 | 84 | 181 | 31 | 48 | 79 | 1060 |
| **Uncertain Eligibility** | 462 | 913 | 652 | 2027 | 64 | 280 | 344 | 117 | 187 | 304 | 134 | 317 | 451 | 3126 |
| **Unproductive** | 2029 | 2489 | 1728 | 6246 | 313 | 900 | 1213 | 450 | 582 | 1032 | 350 | 654 | 1004 | 9495 |
| **REF2 (%)** | 24.4 | 24.7 | 27.0 | 25.1 | 23.7 | 25.6 | 25.2 | 26.3 | 25.1 | 25.7 | 31.2 | 30.9 | 31.0 | 25.8 |
| **Productive** | 4069 | 3759 | 1889 | 9717 | 669 | 1512 | 2181 | 917 | 897 | 1814 | 594 | 940 | 1534 | 15246 |

Notes on Table 5.2

*Adv:* Advantaged ward. *Dis:* .Disadvantaged ward. *Ethn:* .Ethnic minority ward.

*Initial sample:* The sum of the initial sample from Table 7.3 in Plewis (2007b) plus, for England, the 'New families' from Table 2.1.

*Ineligible* From Table 7.3 and Table 4.1 (Plewis, 2007b).

*Uncertain eligibility From* Table 7.3 and Table 4.1 (Plewis, 2007b).

*REF2* is the overall refusal rate, based on the initial sample and defined in the same way as REF1 using the two eligibility rates given previously.

*Unproductive* From Table 7.3 and Table 4.1 (Plewis, 2007b).

*Productive* As in Table 5.1.

5.4    Table 5.3 gives a series of response rates like those in Table 7.4 in Plewis (2004). It shows that the overall response rate across the two sweeps (RR1) is 58% for the UK as a whole, ranging from 46% in the English minority ethnic wards to 65% in the advantaged wards in Wales.

5.5    The response rates (RR2) for sweep 3 were slightly higher than those achieved in sweep 2 although the productive sample is lower. This is probably due to the denominator at sweep 3 being lower than at sweep 2 because of the exclusion of unproductive 'new families' from the denominator at sweep 3. The contact rates (CON) were also higher at sweep 3 than sweep 1 but the cooperation rates (COOP) were lower.

Table 5.3: Sweep 3 Response rates (%) by County and Stratum

| | England | | | | Wales | | | Scotland | | | Northern Ireland | | | UK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Adv. | Dis. | Ethn. | Total | Adv. | Dis. | Total | Adv. | Dis. | Total | Adv. | Dis. | Total | All |
| RR1 | 63 | 54 | 46 | 56 | 65 | 58 | 60 | 63 | 55 | 59 | 57 | 51 | 53 | 58 |
| RR1 (W) | 64 | 55 | 46 | 59 | 66 | 58 | 63 | 62 | 53 | 60 | 54 | 46 | 55 | 59 |
| RR2 | 86 | 79 | 74 | 81 | 82 | 79 | 80 | 83 | 77 | 80 | 84 | 79 | 81 | 81 |
| RR2 (W) | 86 | 79 | 74 | 83 | 82 | 79 | 81 | 83 | 77 | 81 | 84 | 79 | 82 | 82 |
| CON | 97 | 93 | 91 | 94 | 96 | 93 | 94 | 96 | 93 | 95 | 98 | 95 | 96 | 94 |
| COOP | 89 | 85 | 81 | 86 | 85 | 85 | 85 | 86 | 83 | 84 | 86 | 83 | 84 | 85 |

<u>Notes on Table 5.3</u>

*RR1*    This is the *overall* response rate to the study at sweep 3 defined as the ratio of productive cases to all cases in the initial sample for the study known or estimated to be eligible (Lynn et al., 2001). An overall eligibility rate of 74% was used.

*RR2*    This is the response rate to the study for sweep 3, based on Tables 5.1 and defined as the ratio of productive cases to all cases known or estimated to be eligible in the MCS longitudinal sample.

*RR1W*    This is the weighted version of RR1.

*RR2W*    This is the weighted version of RR2, allowing for varying selection probabilities across strata (see Plewis, 2007b, Table 5.5). It differs from RR2 only for the four countries and the UK as a whole.

*CON*    This is the contact rate – the proportion of all cases in which a household member was reached by the interviewer with whom there was contact in person.

*COOP*    This is the cooperation rate – the number of productive cases as a proportion of cases who were contacted during the fieldwork period.

# 6. MCS longitudinal response up to sweep 3

6.1    Table 6.1 which is reproduced from the sweep 3 data user guide (Hansen and Joshi, 2008), shows the MCS longitudinal participation up to sweep 3 out of all the 19,244 families who have ever participated the study, either in sweep 1 and the new families who took part in sweep 2.

Table 6.1. Longitudinal perspective of the MCS productive sample

| Response Description | MCS Sweep Response Pattern | | | MCS1 and 'new families' productives | Breakdown by Country at MCS1 | | | |
|---|---|---|---|---|---|---|---|---|
| | Sweep 1 | Sweep 2 | Sweep 3 | | England | Wales | Scotland | N. Ireland |
| Productive at all sweeps | Y | Y | Y | **13234** | **8314** | **2002** | **1596** | **1322** |
| Productive at sweep 1 and 2 but not 3 | Y | Y | X | **1664** | 1044 | 259 | 218 | 143 |
| Productive at sweep 1 and 3 but not 2 | Y | X | Y | **1444** | **835** | **179** | **218** | **212** |
| Productive at sweep 1 only | Y | X | X | **2210** | 1340 | 320 | 304 | 246 |
| New families: Productive at sweep 2 and 3 | X | Y | Y | **568** | **568** | NA | NA | NA |
| New families: Productive at sweep 2 only | X | Y | X | **124** | 124 | NA | NA | NA |
| **MCS cohort (MCS1 productive+ Productive New families)** | **18552** | **15590** | **15246** | **19244** | **12225** | **2760** | **2336** | **1923** |

*Y*=productive, *X*=non-productive, *NA*=not applicable

6.2    The first row shows that 13,234 families or 68.8% of the total, responded on all three occasions. If we include also the 568 families    in the fifth row who joined at MCS2 and also completed MCS3, there are 13,802 families altogether who have participated in all sweeps for which they were eligible- about 72 percent of the MCS cohort 'ever participated sample'.

6.3    The second row shows that 1,664 families (8.6% of the total) were productive at each of the first two sweeps and then appear to have dropped out.  The fourth  row, who have only so far been seen at the first sweep (2210) are,  like the first row displaying a pattern consistent with  a unilateral direction of survey loss or attrition,  so are  the 124 member of the new family sample who were recruited at MCS2 but who did not complete MCS3, shown in the sixth row.

6.4    However row three shows a pattern of change in the opposite direction:  There are 1,444 cases (7.5% of the total) which were productive at MCS3 after having missed sweep two. This pattern of 'non-monotonic' non-response  indicates that the non-respondents at sweep 3 may not be lost forever, and also offers a challenge to the conventional techniques for dealing with uni-directional, monotonic non-response.


## 7. Predicting response at sweep 3 for weight adjustment

7.1    Typically the non-response weight at the current sweep is the inverse of the predicted probability of responding based on a logistic regression model using data from previous sweep(s). However, at sweep 3 of the MCS there were 1,444 productive families who were unproductive at sweep 2. This resulted in the non-monotone response pattern where a respondent is present at sweep s-1, absent at sweep s and present at sweep s+1. The typical approach of using sweep 2 variables to predict response at sweep 3 would have left all these cases out of the estimation sample. In order to calculate non-response weights for all productive cases at sweep 3, multiple imputation (Rubin, 1987, 2004) was used to impute the required missing data at sweep 2 for the logistic regression model for the probability of responding. The imputation was only done for variables found in earlier  analyses to be related to non-response; see Plewis (2007a), Plewis et al (2008) and Hawkes and Plewis (2006). Predictor variables were imputed where missing due to item non-response or unit non-response in the case of the 1,444 families.

7.2    The **ice** (Royston, 2005) and **mim** (Galati, Royston and Carlin, 2007) commands of Stata version 10 (StataCorp, 2005)  were used for the imputation and analysis within the survey framework. **ice** is an implementation for Stata of methods of multivariate imputation of missing values under missing-at-random assumptions (Royston, 2004) which implements multiple imputations by chained equations.

7.3    The main aim was to impute missing values for sweep 2 and then analyse the data as if there was monotone non-response at sweep 3. In order to achieve this, we first identified possible non-response predictor variables based on our previous work (Plewis et al, 2008) and exploring the data for sweep 2 respondents by assuming a monotone non-response at sweep 3 i.e. ignoring the 1,444 families productive at sweep 3 but not sweep 2.  After the identification of the predictor variables, four imputation models were fitted using variables from different sweeps to predict missing values at sweep 2. In model 1, 2 and 4 below, similar variables such as the highest education attainment of the main respondent at sweep 1 was used as an independent variable to predict the highest education attainment of the main respondent at sweep 2 together with other variables i.e. the "similar" variables were the minimum requirement for the imputation of each sweep 2 variable. In model 3, other variables observed at sweep 2 such as type of accommodation, housing tenure, gender and age of the main respondent were used to predict the missing values of the highest education attainment of the main respondent. The four imputation models were as follows:
     1.  Sweep 1 variables predicting similar sweep 2 missing values

2. Sweep 3 variables predicting similar sweep 2 missing values
3. Other sweep 2 variables  predicting sweep 2 missing values
4. Sweep 1 and 3 variables predicting similar sweep 2 values

After imputation, four models were fitted using complete data from the above imputation and a fifth model using the same predictor variables but ignoring missing data (assuming monotone non-response) was also fitted. The distribution of the predicted non-response weights from each of the four  models using imputed data were then compared to the distribution of sampling and non-response weights in the previous sweeps as well as sweep 3 non-response weight  from the fifth model. Standard errors and the minimum and maximum values of the estimated weights by UK stratum were also compared with the equivalent estimates from previous sweeps.  It was clear from this comparison that the joint predictive power of sweep 1 and 3 variables (model 4) was better at predicting missing values of similar variables at sweep 2 than either sweep 1 or 3 variables separately (models 1 and 3).  Thus, sweep 3 non-response weights were generated using imputed data from model 4 above.

7.4 As a result of using multiple imputation, all 18526 sweep 3 issued cases were used in the logistic modeling of response. Missing values were imputed 10 times and a logistic model of responding at sweep 3 was estimated 10 times, once for each imputed dataset. This yields 10 estimated non-response weights at sweep 3 and the weights issued for sweep 3 are the average of the 10 weights.

7.5 Table 7.1 shows estimates of predictors of response from a logistic regression model. The non-response weights at sweep 3 are the inverse of predicted probabilities from this model.

Table 7.1: Estimates from a logit response model

| Predictor variable | Predictor variable level | Coef. | Std. Err. | P>\|t\| | MI.df |
|---|---|---|---|---|---|
| Age<br>(Ref: 20-30) | Under 20 | -0.161 | 0.098 | 0.104 | 50.5 |
| | 30-40 | 0.209 | 0.073 | 0.006 | 42.8 |
| | 40+ | 0.559 | 0.255 | 0.033 | 52.6 |
| Ethnicity<br>(Ref: White) | mixed | -0.416 | 0.159 | 0.011 | 61.2 |
| | Indian | -0.473 | 0.172 | 0.007 | 102.4 |
| | Pakistani and Bangladeshi | -0.311 | 0.143 | 0.033 | 79.6 |
| | Black or black British | -0.664 | 0.151 | 0 | 136.8 |
| | other ethnic group (inc. Chinese) | -0.508 | 0.206 | 0.015 | 119.3 |
| Tenure<br>(Ref: Mortgage) | Own | -0.525 | 0.144 | 0.001 | 32.6 |
| | Rent LA or HA | -0.198 | 0.092 | 0.037 | 42 |
| | Rent privately | -0.212 | 0.105 | 0.045 | 97 |
| | Other | -0.266 | 0.252 | 0.305 | 18.2 |
| Type of accommodation<br>(Ref: A flat or mansionette, studio, other) | House or bungalow | 0.356 | 0.095 | 0.001 | 46 |
| Highest education attainment<br>(Ref: o level   GCSE grades A-C) | Higher degree | 0.385 | 0.189 | 0.043 | 131.7 |
| | First degree | 0.361 | 0.108 | 0.001 | 105 |
| | Diplomas in higher education | 0.214 | 0.118 | 0.074 | 94.7 |
| | A  AS  S levels | 0.080 | 0.115 | 0.492 | 39.9 |
| | GCSE grades D-G | -0.204 | 0.093 | 0.032 | 71.1 |
| | Other academic qualifications | -0.131 | 0.193 | 0.498 | 52 |
| | None of these qualifications | -0.330 | 0.081 | 0 | 83 |
| Did main resp. Ever breast feed CM? | Breast fed CM | 0.363 | 0.055 | 0 | 328.7 |
| Family working status | Partner not in work or leave | -0.290 | 0.095 | 0.003 | 118.1 |
| | Single parent | -0.168 | 0.082 | 0.043 | 114.1 |
| When joined MCS | New family | -0.313 | 0.124 | 0.012 | 385.9 |
| Income item non-response | Missing income data | -0.285 | 0.094 | 0.005 | 26.1 |
| Constant | _cons | 1.616 | 0.171 | 0 | 200.3 |

Model stats: Multiple-imputation estimates (svy: logit) , Number of Imputations= 10,  Minimum obs (N) =   18526,
Minimum dof =    18.2, LA or HA =Local authority or Housing association,  MI.df=  multiple imputation degrees of freedom.

7.6    Table 7.2 shows estimates of predictors of different types of response from a multinomial logistic regression model. This model shows for example that, while families renting from local authority or housing associations are more likely to not be contacted, those renting from private landlords are more likely to untraced, or other types of unproductive outcome.

Table 7.2: Multinomial model: MCS3 multinomial-logit response model

| Predictor variable | Levels of predictor variable | Refusal vs Productive | | | | No contact vs Productive | | | | Other unproductive, ineligible & untraced vs Productive | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Coef. | Std. Err. | P>\|t\| | MI.df | Coef. | Std. Err. | P>\|t\| | MI.df | Coef. | Std. Err. | P>\|t\| | MI.df |
| Age (Ref: 20-30) | Under 20 | 0.122 | 0.128 | 0.345 | 58.6 | 0.137 | 0.205 | 0.508 | 54.8 | 0.242 | 0.157 | 0.128 | 64.5 |
| | 30-40 | -0.156 | 0.090 | 0.092 | 42.3 | -0.381 | 0.151 | 0.013 | 108.3 | -0.232 | 0.118 | 0.054 | 57.9 |
| | 40+ | -0.404 | 0.288 | 0.164 | 76.0 | -1.185 | 0.594 | 0.052 | 46.7 | -0.669 | 0.460 | 0.153 | 39.2 |
| Ethnicity (Ref: White) | mixed | 0.246 | 0.226 | 0.283 | 35.9 | 0.042 | 0.404 | 0.919 | 43.2 | 0.801 | 0.209 | <0.001 | 100.8 |
| | Indian | 0.528 | 0.212 | 0.014 | 134.1 | 0.358 | 0.380 | 0.35 | 79.8 | 0.402 | 0.305 | 0.191 | 82.3 |
| | Pakistani and Bangladeshi | 0.247 | 0.203 | 0.231 | 38.6 | 0.430 | 0.233 | 0.067 | 185.3 | 0.358 | 0.204 | 0.082 | 158.1 |
| | Black or black British | 0.700 | 0.198 | 0.001 | 94.2 | 0.680 | 0.265 | 0.011 | 305.2 | 0.581 | 0.248 | 0.022 | 58.1 |
| | other ethnic group (inc. Chinese) | 0.535 | 0.261 | 0.043 | 109.0 | -0.608 | 0.908 | 0.51 | 23.2 | 0.726 | 0.322 | 0.026 | 128.3 |
| Tenure (Ref: Mortgage) | Own | 0.527 | 0.146 | <0.001 | 109.1 | 0.728 | 0.287 | 0.014 | 63.1 | 0.395 | 0.306 | 0.211 | 22.7 |
| | Rent LA or HA | 0.080 | 0.126 | 0.532 | 30.0 | 0.513 | 0.218 | 0.027 | 24.0 | 0.267 | 0.188 | 0.166 | 33.1 |
| | Rent privately | -0.029 | 0.163 | 0.862 | 30.5 | 0.360 | 0.246 | 0.153 | 33.7 | 0.569 | 0.175 | 0.002 | 109.9 |
| | Other | 0.040 | 0.390 | 0.92 | 18.4 | 0.321 | 0.584 | 0.588 | 20.2 | 0.589 | 0.304 | 0.059 | 48.1 |
| Type of accommodation (Ref: A flat or mansionette, studio, other) | House or bungalow | -0.236 | 0.113 | 0.038 | 108.6 | -0.304 | 0.178 | 0.094 | 46.4 | -0.583 | 0.144 | <0.001 | 44.3 |
| Highest education attainment (Ref: o level GCSE grades A-C) | Higher degree | -0.264 | 0.221 | 0.235 | 147.6 | -1.001 | 0.584 | 0.09 | 72.7 | -0.452 | 0.397 | 0.258 | 69.6 |
| | First degree | -0.396 | 0.129 | 0.003 | 181.6 | -0.600 | 0.337 | 0.083 | 42.1 | -0.152 | 0.236 | 0.522 | 81.9 |
| | Diplomas in higher education | -0.279 | 0.150 | 0.068 | 64.8 | -0.254 | 0.275 | 0.357 | 100.1 | -0.040 | 0.218 | 0.856 | 125.9 |
| | A AS S levels | -0.053 | 0.147 | 0.721 | 36.2 | -0.375 | 0.265 | 0.163 | 53.7 | 0.023 | 0.185 | 0.9 | 138.4 |
| | GCSE grades D-G | 0.174 | 0.124 | 0.167 | 49.2 | 0.024 | 0.185 | 0.897 | 168.3 | 0.387 | 0.153 | 0.012 | 183.3 |
| | Other academic qualifications | 0.061 | 0.267 | 0.82 | 32.6 | -0.143 | 0.486 | 0.771 | 22.0 | 0.382 | 0.270 | 0.16 | 130.4 |
| | None of these qualifications | 0.276 | 0.111 | 0.017 | 40.4 | 0.222 | 0.156 | 0.157 | 140.0 | 0.515 | 0.137 | <0.001 | 175.3 |

| Predictor variable | Levels of predictor variable | Refusal vs Productive | | | | No contact vs Productive | | | | Other unproductive, ineligible & untraced vs Productive | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Coef. | Std. Err. | P>|t| | MI.df | Coef. | Std. Err. | P>|t| | MI.df | Coef. | Std. Err. | P>|t| | MI.df |
| Did main resp. Ever breast feed CM? | Breast fed CM | -0.462 | 0.075 | <0.001 | 325.3 | -0.354 | 0.112 | 0.002 | 357.7 | -0.157 | 0.094 | 0.096 | 285.9 |
| Family working status | Partner not in work or leave | 0.200 | 0.118 | 0.092 | 142.1 | 0.510 | 0.194 | 0.01 | 80.6 | 0.326 | 0.185 | 0.083 | 54.0 |
| | Single parent | 0.112 | 0.112 | 0.317 | 74.2 | 0.452 | 0.159 | 0.007 | 50.0 | 0.098 | 0.153 | 0.523 | 69.4 |
| When joined MCS | New family | 0.394 | 0.169 | 0.02 | 385.1 | 0.959 | 0.305 | 0.002 | 385.4 | -0.071 | 0.190 | 0.708 | 385.0 |
| Income item non-response | Missing income data | 0.368 | 0.116 | 0.004 | 24.2 | 0.060 | 0.176 | 0.736 | 57.9 | 0.229 | 0.156 | 0.147 | 50.5 |
| Constant | _cons | -2.222 | 0.222 | 0 | 313.5 | -4.146 | 0.396 | 0 | 132.4 | -2.698 | 0.276 | <0.001 | 98.7 |

Model stats: Multiple-imputation estimates (svy: logit) , Number of Imputations= 10, Minimum obs (N) = 18526, Minimum dof = 18.4. LA or HA =Local authority or Housing association

7.7     The estimated mean, minimum and maximum values of overall non-response adjusted weights for single country and whole UK analysis are presented in Tables 7.3 and 7.4. The sampling weights as well as longitudinal weights from sweep 1 (aovwt1|2) and sweep 2 (bovwt1|2) are also reported for comparison. Both tables show a very small change in the estimated non-response adjusted weights from design or sampling weights.

7.8     The non-response adjusted weights vary by sweep and cohort family while the sampling weights are fixed, do not change over time and only vary by stratum.

7.9      We plan to produce a more detailed description of the procedures used to produce sweep 3 non-response adjusted weights in another technical report.

Table 7.3: MCS1-3 minimum, mean and maximum non-response adjusted weight estimates for country specific analyses

| Stratum | weight1 | | aovwt1 | | | | bovwt1 | | | | covwt1 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mean | n | min | mean | max | n | min | mean | max | n | min | mean | max | n |
| England - ADV | 1.32 | 4828 | 1.28 | 1.42 | 3.21 | 4617 | 1.08 | 1.31 | 3.27 | 4209 | 1.11 | 1.35 | 3.33 | 4069 |
| England – DIS | 0.71 | 4806 | 0.691 | 0.793 | 1.85 | 4522 | 0.591 | 0.791 | 2.58 | 3880 | 0.608 | 0.807 | 2.65 | 3759 |
| England – ETHN | 0.24 | 2591 | 0.247 | 0.296 | .693 | 2394 | 0.213 | .323 | 1.25 | 1961 | 0.219 | 0.327 | 1.28 | 1889 |
| Wales – ADV | 1.77 | 832 | 1.7 | 1.84 | 3.73 | 832 | 1.45 | 1.75 | 4.88 | 692 | 1.49 | 1.79 | 5.01 | 669 |
| Wales – DIS | 0.65 | 1928 | 0.637 | 0.725 | 1.52 | 1928 | 0.546 | 0.737 | 1.65 | 1569 | 0.561 | 0.752 | 1.7 | 1512 |
| Scotland – ADV | 1.23 | 1145 | 0.681 | 1.32 | 2.37 | 1145 | 0.598 | 1.29 | 2.64 | 932 | 0.614 | 1.32 | 2.72 | 917 |
| Scotland – DIS | 0.75 | 1191 | 0.416 | 0.828 | 1.67 | 1191 | 0.354 | 0.887 | 2.22 | 882 | 0.363 | 0.896 | 2.2 | 897 |
| Northern Ireland - ADV | 1.41 | 723 | 0.83 | 1.66 | 3.39 | 723 | 0.712 | 1.65 | 3.4 | 586 | 0.731 | 1.67 | 3.5 | 594 |
| Northern Ireland  - DIS | 0.76 | 1200 | 0.438 | 0.917 | 1.68 | 1200 | 0.399 | 1.01 | 2.26 | 879 | 0.41 | 1.03 | 2.32 | 940 |
| **Total** | **0.902** | **19244** | **0.247** | **1** | **3.73** | **18552** | **0.213** | **0.989** | **4.88** | **15590** | **0.219** | **1.01** | **5.01** | **15246** |

Notes for table.7.3

*Weight 1* Sampling weight (Plewis, 2004)

*aovwt1*   The overall weight at sweep 1 which is a product of the sampling weight (weight 1) and non-response weight for the sweep 1

*bovwt1*   The longitudinal weight at sweep 2 which is a product of sweep 1 overall weight (**aovwt1)** and non-response weight at sweep 2

*covwt1*   The longitudinal weight at sweep 3 which is a product of sweep 2 overall weight (**bovwt1)** and non-response weight at sweep 3

Table 7.4: MCS1-3 minimum, mean and maximum non-response adjusted weight estimates for analyses of the UK sample

| Stratum | weight2 | | aovwt2 | | | | bovwt2 | | | | covwt2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mean | n | min | mean | max | n | min | mean | max | n | min | mean | max | n |
| England – ADV. | 2 | 4828 | 1.78 | 1.97 | 4.48 | 4617 | 1.5 | 1.83 | 4.56 | 4209 | 1.54 | 1.88 | 4.64 | 4069 |
| England – DIS. | 1.09 | 4806 | 0.975 | 1.12 | 2.62 | 4522 | 0.835 | 1.12 | 3.64 | 3880 | 0.858 | 1.14 | 3.74 | 3759 |
| England – ETHN. | 0.37 | 2591 | 0.35 | 0.419 | 0.982 | 2394 | 0.302 | 0.461 | 1.77 | 1961 | 0.31 | 0.465 | 1.82 | 1889 |
| Wales – ADV. | 0.62 | 832 | 0.549 | 0.593 | 1.2 | 832 | 0.469 | 0.564 | 1.57 | 692 | 0.481 | 0.578 | 1.61 | 669 |
| Wales – DIS. | 0.23 | 1928 | 0.207 | 0.236 | 0.496 | 1928 | 0.178 | 0.24 | 0.538 | 1569 | 0.183 | 0.244 | 0.553 | 1512 |
| Scotland – ADV. | 0.93 | 1145 | 0.473 | 0.916 | 1.65 | 1145 | 0.416 | 0.899 | 1.84 | 932 | 0.427 | 0.918 | 1.89 | 917 |
| Scotland – DIS. | 0.57 | 1191 | 0.29 | 0.579 | 1.17 | 1191 | 0.247 | 0.62 | 1.55 | 882 | 0.254 | 0.626 | 1.54 | 897 |
| Northern Ireland - ADV. | 0.47 | 723 | 0.255 | 0.509 | 1.04 | 723 | 0.218 | 0.507 | 1.04 | 586 | 0.224 | 0.513 | 1.07 | 594 |
| Northern Ireland - DIS. | 0.25 | 1200 | 0.132 | 0.277 | 0.508 | 1200 | 0.121 | 0.306 | .683 | 879 | 0.124 | 0.311 | 0.702 | 940 |
| **Total** | **0.998** | **19244** | **0.132** | **1** | **4.48** | **18552** | **0.121** | **1.01** | **4.56** | **15590** | **0.124** | **1.02** | **4.64** | **15246** |

Notes for table 7.4

*Weight 2* Sampling weight (Plewis, 2007b)

*aovwt2*  The overall weight at sweep 1 which is a product of the sampling weight (weight 2) and non-response weight for the sweep 1

*bovwt2*  The longitudinal weight at sweep 2 which is a product of sweep 1 overall weight (**aovwt2)** and non-response weight at sweep 2

*covwt2*  The longitudinal weight at sweep 3 which is a product of sweep 2 overall weight (**bovwt2)** and non-response weight at sweep 3

## 8. Number of cohort children in the productive sample at sweep 3

8.1 There were 15,459 cohort children from the 15,246 productive families at sweep 3 which is a net loss of 340 cohort children from the sweep 2 achieved number (see Table 2.2). However, the net loss of cohort families between sweep 2 and 3 is 344 families; four families more that the net loss in cohort children. This indicates that we have regained families with multiple children at sweep 3 from the 1,444 cases who were not productive at sweep 2.

8.2 Table 8.1 shows the distribution of cohort members by age at sweep 3. Nearly 85 percent of children were five (5.0-5.9) years old. About 15 percent were 4 (4.4-4.9) years old. Only 38 children were 6 years old at the time of the interview.

Table 8.1: Distribution of cohort member's age at sweep 3

| Age at interview | | Frequency | Percent |
|---|---|---|---|
| Years | Months | | |
| 4.4-4.9 | 53 | 4 | 0.03 |
| | 54 | 4 | 0.03 |
| | 55 | 24 | 0.16 |
| | 56 | 318 | 2.06 |
| | 58 | 786 | 5.08 |
| | 59 | 1,189 | 7.69 |
| 5-5.9 | 60 | 1,746 | 11.29 |
| | 61 | 2,252 | 14.57 |
| | 62 | 2,315 | 14.98 |
| | 64 | 2,224 | 14.39 |
| | 65 | 1,818 | 11.76 |
| | 66 | 1,472 | 9.52 |
| | 67 | 711 | 4.6 |
| | 68 | 326 | 2.11 |
| | 70 | 159 | 1.03 |
| | 71 | 73 | 0.47 |
| 6-6.5 | 72 | 31 | 0.2 |
| | 73 | 7 | 0.05 |
| Total | | **15,459** | **100** |

## 9. Response at sweep 4 (MCS4)

9.1    Table 9.1 shows that 72% (13,857/19,244) of the MCS sample were productive at sweep 4.   The country specific field[1] refusal rate (REF1) ranges from 18% in England to 22% in Northern Ireland. Stratum specific refusal rates for families in ethnic or disadvantaged wards are consistently higher compared to families in advantaged wards in all UK countries. However, the differences were marginal in Wales and Northern Ireland. Families in disadvantaged wards in Northern Ireland had the highest refusal rate at 23% while those in advantaged wards in England had the lowest refusal rate at 14%. These refusal rate figures show that refusal is to a large extent, the leading reason for non-response.  However, some families who have refused in previous sweep(s) have participated in later sweep(s), see Table 12.1.

Table 9.1: From "Issued" sample to final sample at MCS4 by Stratum

| | England | | | | Wales | | | Scotland | | | NI | | | UK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Adv | Dis | Ethn | Total | Adv | Dis | Total | Adv | Dis | Total | Adv | Dis | Total | All |
| Issued sample[2] | 4828 | 4806 | 2591 | 12225 | 832 | 1928 | 2760 | 1145 | 1191 | 2336 | 723 | 1200 | 1923 | 19244 |
| Ineligible | 180 | 95 | 52 | 327 | 14 | 26 | 40 | 57 | 34 | 91 | 12 | 18 | 30 | 488 |
| ELIG (%) | 96.2 | 97.9 | 97.8 | 97.2 | 98.3 | 98.6 | 98.5 | 94.9 | 96.9 | 95.9 | 98.3 | 98.4 | 98.4 | 97.3 |
| Uncertain eligibility | 119 | 267 | 193 | 579 | 17 | 70 | 87 | 36 | 82 | 118 | 13 | 51 | 64 | 848 |
| Unproductive | 741 | 1067 | 672 | 2480 | 179 | 436 | 615 | 223 | 276 | 499 | 164 | 293 | 457 | 4051 |
| REF1 (%) | 14.3 | 19.4 | 21.9 | 17.9 | 18.7 | 19.6 | 19.3 | 17.4 | 21.2 | 19.4 | 21.9 | 22.8 | 22.5 | 18.8 |
| Productive | 3788 | 3377 | 1674 | 8839 | 622 | 1396 | 2018 | 829 | 799 | 1628 | 534 | 838 | 1372 | 13857 |

Notes on Table 9.1

*Adv.:*    Advantaged ward, *Dis:* Disadvantaged ward, *Ethn:* Ethnic minority ward.

*Issued sample:*      These are the productive families from sweep 1 plus productive 'new families' at sweep 2.

*Ineligible :*        Child deaths (n=23), sensitive (n=25) and temporary and  permanent emigrants

including those who have returned to UK after sweep 4 (n=440).,

*Uncertain eligibility*: Untraced movers: 848 cases (who might have died *or* emigrated) .

*Unproductive:*    Three sub-groups of cases with their UK sizes in brackets: (i) Non-contact (149),(ii) Refusal  (3516), (iii) Other non-response (386).

*REF1*           Refusal rate in the field. This is the ratio of refusals to all issued cases known or estimated to be eligible as defined by Lynn et al. (2001). Estimated eligibility rates are used for cases with uncertain eligibility as explained above.

---

[1] Not all cases in MCS sample were issued to the field
[2] This is not the true issued sample but the MCS sample.

*Productive:* All families with some data (from at least one instrument – Main-interview or self completion, partner, proxy, child measurements and assessments) other than data carried forward from MCS3.

9.2 The 19,244 'issued sample' in Table 9.1 was not the actual sample issued to the field at sweep 4. The MCS4 issued sample was 17,031 cases or 2,213 cases fewer. The reasons for not being issued and the numbers of cases in brackets were: emigration (340), death of a cohort child (22), refusal (1,705), being permanently untraced (136) and ten families who were considered "sensitive" as a result of the cohort child being taken into care etc. The number of not issued cases is cumulative over all four sweeps of the MCS i.e. the number of cohort child deaths is a cumulative number since sweep 1 (or sweep 2 for the new families).

9.3 The decline from the MCS initial[3] sample of 28,927 families to 13,857 productive families at sweep 4 is shown in Table 9.2. The overall refusal rate [REF1] is highest in Northern Ireland at 35% and lowest in England at 33%. Stratum specific overall refusal rates are highest in the Ethnic wards in England and the advantaged wards in Northern Ireland at 37%. The lowest stratum specific overall refusal rate is in English advantaged wards at 30%.

Table 9.2: From initial samples to final sample at MCS4 by stratum.

| | England | | | | Wales | | | Scotland | | | NI | | | UK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Adv | Dis | Ethn | Total | Adv | Dis | Total | Adv | Dis | Total | Adv | Dis | Total | All |
| **Initial Sample** | 6859 | 7506 | 4400 | 18665 | 1076 | 2787 | 3863 | 1581 | 1750 | 3331 | 1109 | 1959 | 3068 | 28927 |
| **Ineligible** | 367 | 288 | **155** | 810 | 33 | 103 | 136 | 115 | 98 | 213 | 31 | 58 | 89 | 1248 |
| **uncertain Eligibility** | 509 | 1020 | 710 | 2239 | 70 | 297 | 367 | 123 | 227 | 350 | 136 | 335 | 471 | 3427 |
| **Unproductive** | 2195 | 2721 | 1861 | 6777 | 351 | 991 | 1342 | 514 | 626 | 1140 | 408 | 728 | 1136 | 10395 |
| **REF2 (%)** | 30.2 | 33.5 | 36.0 | 32.9 | 31.4 | 34.1 | 33.5 | 32.7 | 34.6 | 33.7 | 36.7 | 34.5 | 35.3 | 33.3 |
| **Productive** | 3788 | 3377 | 1674 | 8839 | 622 | 1396 | 2018 | 829 | 799 | 1628 | 534 | 838 | 1372 | 13857 |

Notes on Table 9.2

*Adv:* Advantaged ward. *Dis:* Disadvantaged ward. *Ethn:* Ethnic minority ward.

*Initial sample:* The sum of the initial sample from Table 7.3 in Plewis (2007b) plus, for England, the 'New families' from Table 2.1.

*Ineligible* From Table 7.3 and Table 4.1 in (Plewis, 2007b).

*Uncertain eligibility* From Table 7.3 and Table 4.1 in (Plewis, 2007b).

*REF2* is the overall refusal rate, based on the initial sample and defined in the same way as REF1 using the two eligibility rates given previously.

---

[3] This includes New Families who were not in the true initial sample

*Unproductive* From Table 7.3 and Table 4.1 in (Plewis, 2007b).

*Productive* As in Table 8.1.

Table 9.3: Sweep 4 response rates (%) by stratum

| | England | | | | Wales | | | Scotland | | | NI | | | UK |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Adv | Dis | Ethn | Total | Adv | Dis | Total | Adv | Dis | Total | Adv | Dis | Total | All |
| **RR1** | 60 | 49 | 41 | 51 | 61 | 54 | 56 | 58 | 50 | 54 | 51 | 46 | 48 | 52 |
| **RR1 (W)** | 64 | 55 | 46 | 54 | 66 | 58 | 59 | 62 | 53 | 55 | 54 | 46 | 49 | 55 |
| **RR2** | 82 | 72 | 66 | 74 | 76 | 73 | 74 | 76 | 69 | 73 | 75 | 71 | 73 | 74 |
| **RR2 (W)** | 82 | 72 | 66 | 77 | 76 | 73 | 75 | 76 | 69 | 74 | 75 | 71 | 74 | 76 |
| **CON** | 97 | 93 | 91 | 94 | 98 | 95 | 96 | 96 | 92 | 94 | 97 | 95 | 96 | 95 |
| **COOP** | 84 | 77 | 72 | 79 | 78 | 77 | 77 | 79 | 75 | 77 | 77 | 75 | 76 | 78 |

Notes on Table 9.3

*RR1*  This is the *overall* response rate to the study at sweep 4 defined as the ratio of productive cases to all cases in the initial sample for the study known or estimated to be eligible (Lynn et al., 2001). An overall eligibility rate of 74% was used.

*RR2*  This is the response rate for the study for sweep 4, based on Tables 9.1 and defined as the ratio of productive cases to all cases known or estimated to be eligible in the MCS longitudinal sample.

*RR1W*  This is the weighted version of RR1.

*RR2W*  This is the weighted version of RR2, allowing for varying selection probabilities across strata (see Plewis, 2007b, Table 5.5). It differs from RR2 only for the four countries and the UK as a whole.

*CON*  This is the contact rate – the proportion of all cases in which a household member was reached by the interviewer with whom there was contact in person.

*COOP*  This is the cooperation rate – the number of productive cases as a proportion of cases who were contacted during the fieldwork period.

9.4  Contact rates (CON) at sweep 4 were almost identical to rates in sweep 3 but cooperation rates (COOP) were much lower at sweep 4, see Table 9.3. The UK cooperation rate at sweep 4 is 7% lower than the estimate at sweep 3.

9.5  The unweighted overall response rate (RR1) is 52% for the whole of UK but varies considerably between UK countries as well as between strata within country, see Table 9.3. At 48%, the overall response rate in Northern Ireland is the lowest, followed by England at 51%, Scotland at 54% and Wales at 56% is the highest. The rate in England is largely reduced by families in Ethnic wards where the 41% response rate is lowest across all MCS wards.

# 10. Predicting response at sweep 4 for weight adjustment

10.1    The same procedure used for predicting non-response at sweep 3 was again used at sweep 4. Missing data for predictor variables due to non-monotone non-response were also dealt with in the same way as sweep 3, see Section 7. These procedures allowed for estimation of non-response weights for all issued cases at sweep 4. Sweep 4 non-response predictor variables were mostly the same as at sweep 3. Gender of the cohort child and main respondent's consent for data linkage to birth records were both included in the final model for the first time.

10.2    The estimation process was much easier at sweep 4 than sweep 3 because of the use of Stata 11 (StataCorp, 2009) which includes multiple imputation packages by default. The new features simplified the process of creating variables as well as fitting models.

10.3    As a result of using multiple imputation, 18,736 cases were used in the logistic regression modeling of response. The 508 cases excluded were those known to be emigrants at the time of the survey, or families ineligible for the survey due to the death of the cohort child or families who were considered to be sensitive cases. Missing values for variables measured at sweep 3 were imputed using values of the same variables measured at previous sweeps. For families productive at sweep 4 but not at sweep 3, their sweep 4 values were also used to impute missing values of the same variable at sweep 3. Missing values were imputed 10 times and a logistic model of responding at sweep 4 was estimated 10 times, once for each imputed dataset. This yields 10 estimated non-response weights at sweep 3 and the weights issued for sweep 3 are the average of the 10 weights.

10.4    Table 10.1 shows coefficient estimates of predictors of unit non-response from a logistic regression model. The non-response weights at sweep 4 are the inverse of predicted probabilities from this model. The coefficient estimates in Table 9.1 are similar to estimates at sweep 3 with few changes in the coefficient size for some variables.

10.5    Table 10.2 shows estimates of predictors of different types of response from a multinomial logistic regression model. This model shows for example that, while families renting from local authority or housing associations are more likely to not be contacted, those renting from private landlords are more likely to be untraced, or result into other types of unproductive outcome.

## Table 10.1: Predictors of response: Estimates from a logit response model.

| Predictor variable | Levels of the predictor variable | Coef. | Std. Err. | P>\|t\| | MI.df |
|---|---|---|---|---|---|
| Age in years | Main respondent's age at the birth of the cohort member | 0.029 | 0.004 | <0.001 | 374 |
| Education levels | Main respondents highest education qualification | -0.078 | 0.012 | <0.001 | 80.6 |
| Stratum within country [Ref: England - Advantaged] | England- Disadvantaged | -0.150 | 0.067 | 0.027 | 383.6 |
| | England- Ethnic | -0.181 | 0.095 | 0.057 | 385 |
| | Wales- Advantaged | -0.252 | 0.124 | 0.044 | 385.8 |
| | Wales- Disadvantaged | -0.050 | 0.082 | 0.541 | 382 |
| | Scotland- Advantaged | -0.238 | 0.083 | 0.004 | 381.5 |
| | Scotland- Disadvantaged | -0.277 | 0.092 | 0.003 | 381.6 |
| | Northern Ireland – Advantaged | -0.244 | 0.121 | 0.045 | 386.3 |
| | Northern Ireland – Disadvantaged | -0.085 | 0.100 | 0.397 | 383.2 |
| Sweep 1 consent to Birth records or NHS linkage [Ref: Consent given] | Consent not given | -0.919 | 0.079 | <0.001 | 374.7 |
| Sweep in which family entered study [Ref: Sweep 1] | New family (Sweep 2) | -0.130 | 0.096 | 0.175 | 384.2 |
| Residential mobility between sweep 3 and 4 [Ref: Moved ] | Did not move | 0.425 | 0.057 | <0.001 | 182 |
| Did not breastfeed cohort child? [Ref: Yes] | Did not breastfeed | -0.382 | 0.054 | <0.001 | 382.6 |
| Cohort Child's ethnic group - 8 category classification (UK) [Ref: White] | Mixed | -0.296 | 0.100 | 0.003 | 380.5 |
| | Indian | -0.289 | 0.144 | 0.046 | 364.7 |
| | Pakistani | 0.005 | 0.113 | 0.963 | 358 |
| | Bangladeshi | -0.044 | 0.142 | 0.754 | 380.8 |
| | Black Caribbean | -0.325 | 0.185 | 0.079 | 360.3 |
| | Black African | -0.286 | 0.169 | 0.091 | 381.3 |
| | Other ethnic group (inc. Chinese, other Asian, Other Black) | -0.287 | 0.146 | 0.051 | 355.9 |
| Gender of cohort member [Ref: Male] | Female | 0.129 | 0.044 | 0.003 | 386.2 |
| Whether main respondent is in work or not [Ref: Main respondent in work or leave] | Not in work or leave | -0.158 | 0.058 | 0.008 | 63.5 |
| Main Housing tenure [Ref: Own] | Mortgage | 0.172 | 0.133 | 0.207 | 28.2 |
| | Renting from Local Authority or Housing Association | 0.130 | 0.135 | 0.342 | 45.5 |
| | Renting privately | 0.016 | 0.158 | 0.918 | 32.8 |
| | Other | 0.078 | 0.187 | 0.678 | 52.3 |
| Type of accommodation [Ref: House or bungalow] | Flat or  mansionette | -0.238 | 0.076 | 0.002 | 237.4 |
| | Studio, room, bedsit, other. | -0.168 | 0.310 | 0.59 | 60.2 |
| Income item non-response [Ref: Banded income data given] | Missing banded income data | -0.142 | 0.094 | 0.137 | 41.9 |
| Parents response summary [Ref: Single parent] | Both parents | 0.077 | 0.071 | 0.282 | 64 |
| | One of two | -0.433 | 0.081 | <0.001 | 247.6 |
| Constant | | 0.646 | 0.218 | 0.004 | 80.1 |

**Model stats: Multiple-imputation estimates (svy: logit) , Number of Imputations= 10,  Estimation sample (N) =  18,736**

Table 10.2: Estimates of predictors of different types of response from a multinomial logistic regression model.

| Predictor variable | Levels of the predictor variable | Refusal Vs Productive | | | | No Contact Vs productive | | | | Other unproductive, ineligible & untraced Vs Productive | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Coef. | Std. Err | P.t | DF | Coef. | Std. Err | P.t | DF | Coef. | Std. Err | P.t | DF |
| Age in years | Main respondent's age at the birth of the cohort member | -0.025 | 0.004 | <0.001 | 362.4 | -0.045 | 0.017 | 0.009 | 252.9 | -0.037 | 0.008 | <0.001 | 355.2 |
| Education levels | Main respondents highest education qualification | 0.076 | 0.013 | <0.001 | 103.8 | 0.047 | 0.053 | 0.375 | 96.9 | 0.090 | 0.023 | <0.001 | 101.6 |
| Stratum within country [Ref: England - Advantaged] | England- Disadvantaged | 0.129 | 0.070 | 0.065 | 382.1 | 0.767 | 0.293 | 0.009 | 383.7 | 0.140 | 0.119 | 0.24 | 384.2 |
| | England- Ethnic | 0.126 | 0.100 | 0.209 | 385.2 | 0.780 | 0.336 | 0.021 | 381.9 | 0.259 | 0.167 | 0.123 | 378.9 |
| | Wales- Advantaged | 0.257 | 0.114 | 0.025 | 386.2 | -0.997 | 1.018 | 0.328 | 386.9 | 0.272 | 0.353 | 0.44 | 384.2 |
| | Wales- Disadvantaged | 0.071 | 0.090 | 0.43 | 380.1 | 0.830 | 0.413 | 0.045 | 385.1 | -0.122 | 0.130 | 0.349 | 384.4 |
| | Scotland- Advantaged | 0.158 | 0.103 | 0.128 | 385.5 | 0.161 | 0.532 | 0.763 | 386.2 | 0.498 | 0.148 | 0.001 | 347.8 |
| | Scotland- Disadvantaged | 0.193 | 0.110 | 0.082 | 383 | 0.491 | 0.509 | 0.335 | 381.1 | 0.510 | 0.160 | 0.002 | 374.8 |
| | Northern Ireland – Advantaged | 0.360 | 0.122 | 0.003 | 385.3 | 0.388 | 0.548 | 0.48 | 385.7 | -0.726 | 0.364 | 0.047 | 383.5 |
| | Northern Ireland – Disadvantaged | 0.117 | 0.108 | 0.278 | 379.2 | 0.636 | 0.355 | 0.074 | 381.2 | -0.131 | 0.154 | 0.397 | 382.5 |
| Sweep 1 consent to Birth records or NHS linkage [Ref: Consent given] | Consent not given | 0.979 | 0.084 | <0.001 | 378.7 | 0.267 | 0.339 | 0.432 | 385.3 | 0.787 | 0.162 | <0.001 | 363.5 |
| Sweep in which family entered study [Ref: Sweep 1] | New family (Sweep 2) | -0.084 | 0.125 | 0.503 | 383.9 | 0.647 | 0.374 | 0.084 | 382 | 0.479 | 0.157 | 0.002 | 376.6 |
| Residential mobility between sweep 3 and 4 [Ref: Moved ] | Did not move | 0.109 | 0.075 | 0.149 | 79.9 | 0.226 | 0.297 | 0.447 | 366.4 | -1.593 | 0.090 | <0.001 | 88.3 |
| Did not breastfeed cohort child? [Ref: Yes] | Did not breastfeed | 0.405 | 0.061 | <0.001 | 379.1 | 0.077 | 0.215 | 0.72 | 377.7 | 0.357 | 0.083 | <0.001 | 382.2 |
| Cohort Child's ethnic group - 8 category classification (UK) [Ref: White] | Mixed | 0.159 | 0.121 | 0.189 | 379.9 | -0.374 | 0.735 | 0.611 | 385.5 | 0.684 | 0.159 | <0.001 | 381.5 |
| | Indian | 0.292 | 0.175 | 0.096 | 384.3 | -0.380 | 0.591 | 0.521 | 379.7 | 0.362 | 0.231 | 0.119 | 238 |
| | Pakistani | -0.083 | 0.133 | 0.531 | 365.2 | -0.313 | 0.489 | 0.522 | 375 | 0.282 | 0.196 | 0.151 | 356.5 |
| | Bangladeshi | -0.060 | 0.152 | 0.695 | 371.6 | 0.259 | 0.343 | 0.45 | 352.8 | 0.336 | 0.301 | 0.265 | 377.8 |
| | Black Caribbean | 0.272 | 0.192 | 0.157 | 356.1 | 0.343 | 0.472 | 0.468 | 358.5 | 0.498 | 0.280 | 0.076 | 349.9 |
| | Black African | 0.189 | 0.164 | 0.252 | 378.2 | 0.541 | 0.345 | 0.118 | 311.2 | 0.506 | 0.241 | 0.036 | 377.9 |
| | Other ethnic group (inc. Chinese, other Asian, Other Black) | 0.032 | 0.158 | 0.84 | 353 | 0.358 | 0.442 | 0.418 | 369.3 | 0.849 | 0.244 | 0.001 | 318 |
| Gender of cohort member [Ref: Male] | Female | -0.132 | 0.050 | 0.008 | 386 | -0.077 | 0.189 | 0.683 | 386.6 | -0.132 | 0.074 | 0.075 | 382.8 |

| Variable | Category | Coef | SE | p | Stat | Coef | SE | p | Stat | Coef | SE | p | Stat |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Whether main respondent is in work or not [Ref: Main respondent in work or leave] | Not in work or leave | 0.144 | 0.064 | 0.028 | 77.1 | -0.024 | 0.214 | 0.911 | 233.5 | 0.232 | 0.113 | 0.047 | 33.1 |
| Main Housing tenure [Ref: Own] | Mortgage | -0.245 | 0.134 | 0.076 | 36.1 | -0.107 | 0.684 | 0.875 | 202 | 0.120 | 0.266 | 0.654 | 40.4 |
| | Renting from Local Authority or Housing Association | -0.224 | 0.146 | 0.132 | 45.2 | 0.182 | 0.655 | 0.781 | 255 | 0.248 | 0.260 | 0.343 | 75.7 |
| | Renting privately | -0.208 | 0.181 | 0.258 | 33.8 | 0.543 | 0.666 | 0.417 | 176.6 | 0.466 | 0.282 | 0.105 | 48.3 |
| | Other | -0.215 | 0.200 | 0.282 | 122.3 | 0.373 | 0.874 | 0.671 | 73.3 | 0.347 | 0.347 | 0.323 | 41.1 |
| Type of accommodation [Ref: House or bungalow] | Flat or mansionette | 0.184 | 0.091 | 0.044 | 339.7 | 0.804 | 0.279 | 0.005 | 88.9 | 0.253 | 0.109 | 0.022 | 147.8 |
| | Studio, room, bedsit, other. | -0.134 | 0.362 | 0.713 | 83.3 | 0.542 | 1.096 | 0.623 | 51.7 | 0.639 | 0.426 | 0.137 | 72.4 |
| Income item non-response [Ref: Banded income data given] | Missing banded income data | 0.119 | 0.105 | 0.264 | 38 | 0.105 | 0.375 | 0.78 | 44.8 | 0.201 | 0.147 | 0.179 | 44 |
| Parents response summary [Ref: Single parent] | Both parents | 0.039 | 0.095 | 0.686 | 34.9 | -0.395 | 0.274 | 0.151 | 190.1 | -0.300 | 0.100 | 0.003 | 173.4 |
| | One of two | 0.574 | 0.095 | <0.001 | 186.2 | 0.121 | 0.349 | 0.728 | 180.5 | 0.126 | 0.160 | 0.433 | 66.9 |
| Constant | | -1.429 | 0.241 | <0.001 | 71.4 | -4.306 | 0.950 | <0.001 | 250.9 | -1.495 | 0.422 | 0.001 | 97.9 |

**Model stats: Multiple-imputation estimates (svy: mlogit) , Number of Imputations= 10, Estimation sample (N) = 18,736.**

10.6  The estimated mean, minimum and maximum values of overall non-response adjusted weights for single country, whole UK and Great Britain only analyses are presented in Tables 10.3, 10.4 and 10.5 respectively. The sampling weights as well as longitudinal (overall) weights from sweep 1 (aovwt1|2), sweep 2 (bovwt1|2) and sweep 3 (bovwt1|2) are also reported for comparison. Both tables show a small change in the estimated non-response adjusted weights from design (sampling weights). Weights shown for GB analyses are design (sampling) weights, and sweep 2 to 4 which are adjusted for non-response.

10.7  The non-response adjusted weights vary by sweep and cohort family while the sampling weights are fixed, do not change over time and only vary by stratum.

**Table 10.3: MCS1 to 4 minimum, mean and maximum non-response adjusted weight estimates for country specific analyses***

| Stratum | weight1 | | aovwt1 | | | | bovwt1 | | | | covwt1 | | | | dovwt1 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mean | N | min | mean | max | n | min | mean | max | n | min | mean | max | n | min | mean | max | n |
| England – ADV | 1.32 | 4828 | 1.28 | 1.42 | 3.21 | 4617 | 1.08 | 1.31 | 3.27 | 4209 | 1.11 | 1.35 | 3.33 | 4069 | 0.977 | 1.34 | 4.29 | 3788 |
| England – DIS | 0.71 | 4806 | 0.691 | 0.793 | 1.85 | 4522 | 0.591 | 0.791 | 2.58 | 3880 | 0.608 | 0.807 | 2.65 | 3759 | 0.556 | 0.911 | 6.12 | 3377 |
| England – ETHN | 0.24 | 2591 | 0.247 | 0.296 | 0.693 | 2394 | 0.213 | 0.323 | 1.25 | 1961 | 0.219 | 0.327 | 1.28 | 1889 | 0.198 | 0.399 | 2.09 | 1674 |
| Wales – ADV | 1.77 | 832 | 1.7 | 1.84 | 3.73 | 832 | 1.45 | 1.75 | 4.88 | 692 | 1.49 | 1.79 | 5.01 | 669 | 1.14 | 1.65 | 5.23 | 622 |
| Wales – DIS | 0.65 | 1928 | 0.637 | 0.725 | 1.52 | 1928 | 0.546 | 0.737 | 1.65 | 1569 | 0.561 | 0.752 | 1.7 | 1512 | 0.433 | 0.712 | 2.1 | 1396 |
| Scotland – ADV | 1.23 | 1145 | 0.681 | 1.32 | 2.37 | 1145 | 0.598 | 1.29 | 2.64 | 932 | 0.614 | 1.32 | 2.72 | 917 | 0.456 | 1.14 | 5.11 | 829 |
| Scotland – DIS | 0.75 | 1191 | 0.416 | 0.828 | 1.67 | 1191 | 0.354 | 0.887 | 2.22 | 882 | 0.363 | 0.896 | 2.2 | 897 | 0.282 | 0.853 | 2.49 | 799 |
| NI- ADV | 1.41 | 723 | 0.83 | 1.66 | 3.39 | 723 | 0.712 | 1.65 | 3.4 | 586 | 0.731 | 1.67 | 3.5 | 594 | 0.465 | 1.28 | 6.19 | 534 |
| NI- DIS | 0.76 | 1200 | 0.438 | 0.917 | 1.68 | 1200 | 0.399 | 1.01 | 2.26 | 879 | 0.41 | 1.03 | 2.32 | 940 | 0.26 | 0.818 | 2.85 | 838 |
| Total | 0.902 | 19244 | 0.247 | 1.O | 3.73 | 18552 | 0.213 | 0.989 | 4.88 | 15590 | 0.219 | 1.01 | 5.01 | 15246 | 0.198 | 1.00 | 6.19 | 13857 |

Notes for Table.9.3

*Weight 1* Sampling weight (Plewis, 2004)

*aovwt1*   The overall weight at sweep 1 which is a product of the sampling weight  (weight 1) and non-response weight for the sweep 1

*bovwt1*   The longitudinal weight at sweep 2 which is a product of sweep 1 overall weight (**aovwt1)** and non-response weight at sweep 2

*covwt1*   The longitudinal weight at sweep 3 which is a product of sweep 2 overall weight (**bovwt1)** and non-response weight at sweep 3

*dovwt1*   The longitudinal weight at sweep 4 which is a product of sweep 3 overall weight (**covwt1)** and non-response weight at sweep 4

**Table 10.4: MCS1 to 4 minimum, mean and maximum non-response adjusted weight estimates for analyses of the whole of UK sample***

| Stratum | weight2 | | aovwt2 | | | | bovwt2 | | | | covwt2 | | | | dovwt2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mean | n | min | mean | max | n | min | mean | max | n | min | mean | max | n | min | mean | max | n |
| England – ADV | 2.0 | 4828 | 1.78 | 1.97 | 4.48 | 4617 | 1.50 | 1.83 | 4.56 | 4209 | 1.54 | 1.88 | 4.64 | 4069 | 1.24 | 1.71 | 5.45 | 3788 |
| England – DIS | 1.09 | 4806 | 0.975 | 1.12 | 2.62 | 4522 | 0.835 | 1.12 | 3.64 | 3880 | 0.858 | 1.14 | 3.74 | 3759 | 0.716 | 1.18 | 7.88 | 3377 |
| England – ETH | 0.37 | 2591 | 0.35 | 0.419 | 0.982 | 2394 | 0.302 | 0.461 | 1.77 | 1961 | 0.310 | 0.465 | 1.82 | 1889 | 0.256 | 0.518 | 2.69 | 1674 |
| Wales – ADV | 0.62 | 832 | 0.549 | 0.593 | 1.20 | 832 | 0.469 | 0.564 | 1.57 | 692 | 0.481 | 0.578 | 1.61 | 669 | 0.387 | 0.558 | 1.78 | 622 |
| Wales – DIS | 0.23 | 1928 | 0.207 | 0.236 | 0.496 | 1928 | 0.178 | 0.24 | 0.538 | 1569 | 0.183 | 0.244 | 0.553 | 1512 | 0.148 | 0.244 | 0.721 | 1396 |
| Scotland – ADV | 0.93 | 1145 | 0.473 | 0.916 | 1.65 | 1145 | 0.416 | 0.899 | 1.84 | 932 | 0.427 | 0.918 | 1.89 | 917 | 0.358 | 0.898 | 4.02 | 829 |
| Scotland – DIS | 0.57 | 1191 | 0.290 | 0.579 | 1.17 | 1191 | 0.247 | 0.62 | 1.55 | 882 | 0.254 | 0.626 | 1.54 | 897 | 0.223 | 0.674 | 1.97 | 799 |
| NI- ADV | 0.47 | 723 | 0.255 | 0.509 | 1.04 | 723 | 0.218 | 0.507 | 1.04 | 586 | 0.224 | 0.513 | 1.07 | 594 | 0.186 | 0.513 | 2.48 | 534 |
| NI- DIS | 0.25 | 1200 | 0.132 | 0.277 | 0.508 | 1200 | 0.121 | 0.306 | 0.683 | 879 | 0.124 | 0.311 | 0.702 | 940 | 0.102 | 0.322 | 1.13 | 838 |
| Total | 0.998 | 19244 | 0.132 | 1.00 | 4.48 | 18552 | 0.121 | 1.01 | 4.56 | 15590 | 0.124 | 1.02 | 4.64 | 15246 | 0.102 | 1.00 | 7.88 | 13857 |

Notes for Table.9.4

*Weight 2* Sampling weight (Plewis, 2007b)

*aovwt2*  The overall weight at sweep 1 which is a product of the sampling weight  (weight 2) and non-response weight for the sweep 1

*bovwt2*  The longitudinal weight at sweep 2 which is a product of sweep 1 overall weight (**aovwt2)** and non-response weight at sweep 2

*covwt2*  The longitudinal weight at sweep 3 which is a product of sweep 2 overall weight (**bovwt2)** and non-response weight at sweep 3

*dovwt2*  The longitudinal weight at sweep 4 which is a product of sweep 3 overall weight (**covwt2)** and non-response weight at sweep 4

**Table 10.5: MCS1 to 4 minimum, mean and maximum non-response adjusted weight estimates for analyses of the Great Britain (GB) sample or for GB versus Northern Ireland comparisons.**

| Stratum | weightgb | | | | bovwtgb | | | | covwtgb | | | | dovwtgb | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | min | mean | max | n | min | mean | max | n | min | mean | max | n | min | mean | max | n |
| England – ADV | 1.78 | 1.78 | 1.78 | 4617 | 1.58 | 1.75 | 3.96 | 4209 | 1.40 | 1.71 | 4.29 | 4069 | 1.13 | 1.59 | 5.83 | 3788 |
| England – DIS | 0.97 | 0.97 | 0.97 | 4522 | 0.86 | 1.03 | 2.29 | 3880 | 0.765 | 1.06 | 3.77 | 3759 | 0.63 | 1.12 | 6.98 | 3377 |
| England – ETHN | 0.33 | 0.33 | 0.33 | 2394 | 0.294 | 0.387 | 1.08 | 1961 | 0.265 | 0.426 | 1.71 | 1889 | 0.222 | 0.491 | 3.21 | 1674 |
| Wales – ADV | 0.55 | 0.55 | 0.55 | 832 | 0.491 | 0.549 | 1.04 | 692 | 0.444 | 0.536 | 1.2 | 669 | 0.364 | 0.534 | 1.4 | 622 |
| Wales – DIS | 0.20 | 0.20 | 0.20 | 1928 | 0.179 | 0.214 | 0.416 | 1569 | 0.161 | 0.219 | 0.75 | 1512 | 0.131 | 0.225 | 0.909 | 1396 |
| Scotland – ADV | 0.82 | 0.82 | 0.82 | 1145 | 0.735 | 0.848 | 1.55 | 932 | 0.665 | 0.828 | 1.80 | 917 | 0.560 | 0.826 | 3.23 | 829 |
| Scotland – DIS | 0.51 | 0.51 | 0.51 | 1191 | 0.457 | 0.582 | 1.19 | 882 | 0.414 | 0.593 | 1.88 | 897 | 0.355 | 0.658 | 2.82 | 799 |
| NI- ADV | 1.41 | 1.41 | 1.41 | 723 | 0.712 | 1.65 | 3.40 | 586 | 0.732 | 1.68 | 3.5 | 594 | 0.465 | 1.28 | 6.19 | 534 |
| NI- DIS | 0.76 | 0.76 | 0.76 | 1200 | 0.399 | 1.01 | 2.26 | 879 | 0.41 | 1.03 | 2.32 | 940 | 0.26 | 0.818 | 2.85 | 838 |
| Total | 0.2 | 0.954914 | 1.78 | 18552 | 0.179 | 1.03 | 3.96 | 15590 | 0.161 | 1.03 | 4.29 | 15246 | 0.131 | 1.00 | 6.98 | 13857 |

Notes for Table.9.5

*weightgb* weight for MCS sample in Great Britain analyses(Plewis, 2007b)

*bovwtgb* The longitudinal weight at sweep 2 which is a product of sweep 1 overall weight (*weightgb*) and non-response weight at sweep 2

*covwtgb* The longitudinal weight at sweep 3 which is a product of sweep 2 overall weight (*bovwtgb*) and non-response weight at sweep 3

*dovwtgb* The longitudinal weight at sweep 4 which is a product of sweep 3 overall weight (*covwtgb*) and non-response weight at sweep 4

## 11. Number of cohort children in the productive sample at sweep 4

11.1   There were 14,042 cohort children from the 13,857 productive families at sweep 4. Table 11.1 shows the distribution of cohort members by age in months at sweep 4. Nine in ten children were interviewed when they were between 6.5 and 7.5 years old.  Nearly 75% percent of children were 3 months short of or 3 months after their 7th birthday, see Table 11.1.

Table 11.1: Distribution of cohort member's age at sweep 4

| Age in Months | Frequency | Unweighted Percent |
|---|---|---|
| ≤80 | 281 | 2.02 |
| 81 | 433 | 3.12 |
| 82 | 758 | 5.47 |
| 83 | 1,100 | 7.94 |
| 84 | 1,479 | 10.67 |
| 85 | 1,715 | 12.38 |
| 86 | 1,712 | 12.35 |
| 87 | 1,686 | 12.17 |
| 88 | 1,461 | 10.54 |
| 89 | 1,306 | 9.42 |
| 90 | 853 | 6.16 |
| 91≥ | 1073 | 7.74 |
| **Total families** | **13,857** | **100** |

## 12. MCS longitudinal response up to sweep 4

12.1   Table 12.1 shows MCS sweep 1 to sweep 4 response history by UK country. Rows starting with S1 show response history for families present and productive  at sweep 1 and rows starting with S2 show the new families response history. Overall, 63% MCS families have participated in all sweeps they were eligible for, see rows S1,2,3,4 and S2,3,4. The fourth row shows families were not productive at sweep 2 but came back at sweep 3 where 1,029 out of 1,444 families participated at wave 4 as well. The fifth row shows 445 productive families at sweep 4 who were not productive at sweep 3. The seventh row shows 168 families who, before sweep 4, participated only at sweep 1.

12.2 A total of 494 new families participated at sweep 4, the majority of which have participated in all sweeps since they joined MCS at sweep 2. There were 26 productive families at sweep 4 who did not participate at sweep 3.

Table 12.1: MCS Response History by MCS1 UK country.

| Productive these Sweeps | England | Wales | Scotland | Northern Ireland | Total |
|---|---|---|---|---|---|
| **S1,2,3,4** | **60.4** | **64.5** | **59.2** | **60.9** | **60.9** |
| | **(7387)** | **(1779)** | **(1384)** | **(1171)** | **(11721)** |
| S1,2,3 | 7.6 | 8.1 | 9.1 | 7.9 | 7.9 |
| | (927) | (223) | (212) | (151) | (1513) |
| S1,2 | 6.3 | 6.6 | 6.7 | 5.7 | 6.3 |
| | (772) | (182) | (156) | (109) | (1219) |
| **S1,3,4** | **4.7** | **4.9** | **7.1** | **8.0** | **5.3** |
| | **(573)** | **(136)** | **(166)** | **(154)** | **(1029)** |
| **S1,2,4** | **2.2** | **2.8** | **2.7** | **1.8** | **2.3** |
| | **(272)** | **(77)** | **(62)** | **(34)** | **(445)** |
| S1,3 | 2.1 | 1.6 | 2.2 | 3.0 | 2.2 |
| | (262) | (43) | (52) | (58) | (415) |
| **S1,4** | **0.9** | **0.9** | **0.7** | **0.7** | **0.9** |
| | **(113)** | **(26)** | **(16)** | **(13)** | **(168)** |
| S1 | 10.0 | 10.7 | 12.3 | 12.1 | 10.6 |
| | (1227) | (294) | (288) | (233) | (2042) |
| **S2,3,4** | **3.8** | **0.0** | **0.0** | **0.0** | **2.4** |
| | **(468)** | **(0)** | **(0)** | **(0)** | **(468)** |
| S2,3 | 0.8 | 0.0 | 0.0 | 0.0 | 0.5 |
| | (100) | (0) | (0) | (0) | (100) |
| **S2,4** | **0.2** | **0.0** | **0.0** | **0.0** | **0.1** |
| | **(26)** | **(0)** | **(0)** | **(0)** | **(26)** |
| S2 | 0.8 | 0.0 | 0.0 | 0.0 | 0.5 |
| | (98) | (0) | (0) | (0) | (98) |
| Total | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| | (12225) | (2760) | (2336) | (1923) | (19244) |

Notes: Unweighted percent. Unweighted sample in parentheses. S1=Families who joined MCS from sweep 1, S2=Families who joined MCS at sweep 2 (New families).

# References

Calderwood, L., Ketende, S. and McDonald, J. (2008). Patterns of longitudinal participation in the Millennium Cohort Study. Paper Prepared for Panel Surveys Workshop in Essex 14th-15th July 2008.

Hansen, K. (ed) (2008) *Millennium Cohort Study First, Second and Third Surveys: A Guide to the Datasets Third Edition*. Centre for Longitudinal Studies, Institute of Education, University of London, UK.

Hansen, K and Joshi, H. (eds) (2008) *Millennium Cohort Study Second Survey: A User's Guide to Initial Findings.* Centre for Longitudinal Studies, Institute of Education, University of London, UK.

Hawkes, D. and Plewis, I. (2006). *Modelling non-response in the National Child Development Study*. Journal of the Royal Statistical Society, Series A, 169, 479-491.

Galati, J.C, Royston, P., Carlin,J.B.( 2007). *MIM: Stata module to analyse and manipulate multiply imputed datasets.* Statistical Software Components S456825, Boston College Department of Economics, revised 26 Aug 2008.

Lynn, P., Beerten, R., Laiho, J. and Martin, J. (2001), *Recommended Standard Final Outcome Categories and Standard Definitions of Response Rate for Social Surveys*. Working papers of the Institute for Social and Economic Research, paper 2001-23. Colchester: Univer*nium Cohort Study First Survey: Technical Report on Sampling (3rd. Edition)*. London: Centre for Longitudinal Studies, Institute of Education, University of London.

Plewis, I. (2007a) *Non-response in a birth cohort study: the case of the Millennium Cohort Study*. International Journal of Social Research Methodology, 10, 325-334.

Plewis, I. (2007b), *Millennium Cohort Study First Survey: Technical Report on Sampling (4th. Edition)*. London: Centre for Longitudinal Studies, Institute of Education, University of London

Plewis,I., Ketende,S.C., Joshi, H., Hughes, G. ( 2008). *The Contribution of Residential Mobility to Sample Loss in a Birth Cohort Study: Evidence from the First Two Waves of the UK Millennium Cohort Study*. Journal of Official Statistics, Vol. 24, No. 3, 2008, pp. 365–385

Royston, P. 2005. *Multiple imputation of missing values: Update of ice*. Stata Journal 5(4): 527-536.

Rubin, D.B. (1987, 2004). *Multiple imputation for nonresponse in surveys*. New York: John Wiley. Reprinted as a Wiley Classic, 2004.

StataCorp. 2005. *Stata Statistical Software: Release 10*. College Station, TX: StataCorp LP.

StataCorp. 2009. *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP.