Centre for Longitudinal Studies

CLS Cohort Studies

Working Paper 2010/4

Non-response in the 1970 British Cohort Study (BCS70) from birth to 34 years

Sosthenes C. Ketende
John McDonald
Shirley Dex

October 2010

Non-response in the 1970 British Cohort Study (BCS70) from birth to 34 years

Sosthenes C. Ketende
John McDonald
Shirley Dex

Centre for Longitudinal Studies
Institute of Education
London

October 2010

First published in October 2010 by the
Centre for Longitudinal Studies
Institute of Education, University of London
20 Bedford Way
London WC1H 0AL
www.cls.ioe.ac.uk

ISBN 978-1-906929-19-0

The Centre for Longitudinal Studies (CLS) is an ESRC Resource Centre based at the Institution of Education. It provides support and facilities for those using the three internationally-renowned birth cohort studies: the National Child Development Study (1958), the 1970 British Cohort Study and the Millennium Cohort Study (2000). CLS conducts research using the birth cohort study data, with a special interest in family life and parenting, family economics, youth life course transitions and basic skills. The views expressed in this work are those of the author(s) and do not necessarily reflect the views of the Economic and Social Research Council. All errors and omissions remain those of the author(s).

This document is available in alternative formats.
Please contact the Centre for Longitudinal Studies.
tel: +44 (0)20 7612 6875
email: info@cls.ioe.ac.uk

Contents

1. Introduction

Longitudinal data have the potential to offer social science and policy makers many benefits. For example, such data can provide information about individuals' lives that can be used to evaluate policy interventions, and information about the durations of time people spend in states that are less than ideal (e.g. poverty). Analyses of longitudinal data are necessary for social scientists to get as close as possible to identifying cause and effect in individuals' lives. However, for the potential of longitudinal data to be fully achieved, data needs to be collected systematically over time from the same individuals. The analysis potential of longitudinal data is put at risk if individuals drop out of the successive data collection contacts, especially if they drop out in large numbers, and if those who fail to continue are a biased sample of the original sample. Attrition is one of the perennial worries in conducting prospective longitudinal surveys, either panel or cohort studies. It is in our interest, therefore, to try to learn as much as possible about those who either cannot be found in the follow up waves, or who refuse to cooperate when approached to participate again. It may be possible to devise fieldwork strategies to improve response rates in successive longitudinal data collection waves by learning more about the problem. Unlike the many cross-sectional surveys, longitudinal data collection does have information about individuals who responded and participated at previous data collection waves, and this information can be used to analyse whether they respond at successive data collection waves.

Individuals drop out of longitudinal surveys for a range of potential reasons. Broadly speaking these reasons fall into two types: either they cannot be found when the time for follow-up comes round, traditionally called non-contact (although this could be divided into being first un-locatable, and second un-contactable once the correct address has been found); alternatively individuals may refuse to cooperate. This may be because of particular circumstances at the time they are approached that make it inconvenient, or because of a bad experience when they were last interviewed (e.g. a perception of excessive burden from the previous interview, or they did not like the interviewer). This mixture of potential influences on individuals' decisions whether to cooperate, across both personal factors and factors related to the survey or the survey operations, make analyses of non-contact and refusal a complex business. Nonetheless, through systematic analyses of the growing number of large-scale longitudinal data sources, especially in the USA, knowledge has advanced over time about the characteristics of those respondents who are more or less likely to continue to participate.

The main aim of this paper is to analyse the available data from a British longitudinal data set, to see if there are lessons to learn that may help future fieldwork practice for longitudinal studies. For this we focus on a longitudinal survey that has been relatively neglected in terms of analyses of non-response, the 1970 British Cohort Study (BCS70). We first examine non-response at successive waves, and later investigate whether there is anything to learn about response from the fact that sub-studies were carried out on these data at different points over its lifetime. However, first we present a brief review of findings from analyses of non-response in other

longitudinal data sets. This is followed by a description of the data and methods used in the analyses. The size of attrition in BCS70 is then described, followed by the results from the analysis of wave response. The examination of the sub-studies is then presented, followed by the final conclusions.

2. Earlier literature

Studies of non-response in longitudinal studies have revealed that it most often has systematic elements and is not random. However, at the same time, many studies have found that the systematic components of non-response account for a very small part of the variation in response. Conclusions have often been drawn, therefore, that non-response, even though systematic, is not such a serious problem to the representativeness of the study in the case of many longitudinal panel studies. The analysis of non-response in the US Panel Study of Income Dynamics (PSID) was probably the first study to reach this conclusion (Fitzgerald, 1998a); and it has since been re-iterated by others (Macurdy et al, 1998 for NLSY; Hawkes and Plewis, 2006 for NCDS; Watson, 2003 for ECHP). [1]

Analyses of non-response have focussed on two sets of predictors in examining its systematic components: first the characteristics of the individuals or households were examined (e.g. Lillard and Panis, 1998; Fitzgerald et al, 1998). Such characteristics are obtained from earlier waves of the studies and may be time-constant or time-varying characteristics about individuals. Second, and more recent, characteristics of the fieldwork process, so-called paradata, have been collected and examined; again this is usually from earlier waves of the survey (e.g. Campanelli et al, 1997; Lepkowski and Couper, 2002; Groves and Couper, 1998; Lynn et al, 2002). Finding systematic elements in the survey process offers direct routes to intervening in future survey practice to prevent or reduce non-response.

Characteristics from earlier waves associated with non-response clearly vary according to the nature of the study and the survey units. Nonetheless, there are some common and differing findings across studies. For example, it is common to find higher non-response rates among men in comparison with women; those who recently moved house compared with those in longer tenure; those in rented accommodation compared with owner occupiers; young people and older age groups compared with the middle aged; the never married (or divorced/separated) compared with married; those on welfare compared with off welfare; disabled compared with non disabled; and lower educated compared with higher educated (SIPP, McArthur, 1988; ECHP, Behr et al, 2005; PSID, FitzGerald et al, 1998b). Many US studies have found lower response rates from black Americans (McArthur for SIPP, FitzGerald for PSID, and Olsen for NLSY; Allen et al, 1991). Some UK analyses have also found lower responses rates for minorities; but usually all minorities combined in a single

[1] However, analyses of the impact of non-response on particular topics have found attrition sometimes does produce biases in the results; but on other topics or data sets it does not. Studies on the impact of attrition on particular outcomes, of which there are many, are too numerous to cover in this paper.

category, as either non-white or non-UK born (Foster, 1998; Lynn and Clarke, 2001). However, while findings on these characteristics are more systematic across studies of non-response, other characteristics vary more between studies. For example, low income or poor households sometimes appear to have higher non-response (Behr et al for ECHP) and sometimes not (McArthur, 1998 for SIPP found non-poor had higher non-response). Similarly whether a survey participant was unemployed or employed, or has a smaller versus larger family size, can both be associated with higher and lower non-response rates. Other notable variations in response rates have been found across countries from analyses of the European Community Household Panel Study (ECHP), although some analysts have pointed to this being partly related to undocumented differences in survey practices across the countries included (Behr, et al, 2005; Watson, 2003; Vandecasteele and Debels, 2006). A further examination has been made in some studies of whether item non-response on particular questions, or an incomplete questionnaire, has been linked with unit non-response at the next wave (Vandecasteele and Debels, 2006), with some correlations being found.

Research on non-response has examined the role of fieldwork procedures and paradata; for example, the length of the fieldwork period, interviewer effects, interviewer continuity from one wave to another, number of call-backs, and the gender and ethnicity of the interviewers. Evidence has been found that many of these survey process characteristics are correlated with subsequent non-response. For example, SIPP and BHPS show a relationship between more call backs at an earlier wave and higher likelihood of non-response at the subsequent wave; ECHP found that response was higher when the interviewer was the same person wave on wave (Behr et al, 2005); but also found correlations between non-response and the duration of the interview, the mode of interview, the number of visits, and the length of fieldwork period (Vandercasteele and Debels, 2006).

Some studies suggest that these fieldwork process measures explain less of the variation in response than do survey participants' characteristics (Nicoletti and Peracchi, 2005 for ECHP). However, Olsen argues that it is survey methodology – structuring the survey process as well as possible – that holds the key to successfully achieving higher response (Olsen, 2005). Olsen claims that survey response can be expected to rise from as much as 71.5 to 87.5 per cent by changing the survey process in a number of ways, which is a sizeable increase.

Lynn et al (2002) drew attention to a weakness of much previous research. They argued that modelling of non-response typically either confounds ease of contact with reluctance to participate, or isolates one without considering simultaneously the effect of the other. Best practice is now regarded as modelling the propensity to be located (or contacted) as well as the nested propensity of individuals to cooperate and provide survey responses. However, the attempts to analyse these two main types of non-response in BHPS data did not find any evidence of correlation between these two (Lynn et al, 2002). Nicoletti and Peracchi (2005) carried out a similar analysis of the ECHP data and reached the same conclusions. In order to estimate such models, it is necessary to have data available about the survey process, and not just about the survey participants (in previous waves). Such data are not always

available for older longitudinal surveys which were initiated long before the latest thinking was publicised on modelling survey response. The data set to be examined here unfortunately did not routinely store such data until recently.

3. The 1970 British Birth Cohort Study (BCS70)

BCS70 began in 1970 by collecting data about perinatal mortality from a Census of babies born in England, Scotland, Wales and Northern Ireland in a particular week in April, approximately 17,000 in total. These babies formed the cohort members although in the early years, most information was provided by parents of the children. The main focus in the first survey was on the circumstances and outcomes of birth. Over the years, the study has broadened its focus to cover many aspects of cohort members and their families such as health, education, employment and social development. In comparison to the earlier 1958 birth cohort, the National Child Development Study (NCDS), BCS70 has encountered more data collection problems. More details about the BCS70 study can be found on the Centre for Longitudinal Studies website www.cls.ioe.ac.uk and in Plewis et al (2004).

Data collection took place first at the time of the birth, referred to as wave0, and then in six further waves, wave1 to wave6, at ages 5, 10, 16, 26, 30 and 34. However, the original BCS70 cohort children in Northern Ireland were not followed up in these successive waves. Up to and including age 16, it was primarily the parents of the cohort child who had to decide, on behalf of the child, whether to respond at each wave. However, at age 26, the cohort child had to make their own decision to be part of the study. Recorded data for these past waves of BCS70 data did not include details of whether non-response was due to non-contact or to refusal, so analysis of this distinction was not possible.

This paper presents information about cross-sectional and longitudinal responses using all of the individuals and their families who ever participated in the study respectively. However, some groups have been excluded from these analyses. As mentioned above, children born in Northern Ireland at the baseline survey were not followed up in later waves and are not included in these analyses therefore, although they are included in the descriptive data presented about BCS70 wave0. Similarly, children who joined the study after it started, as a result of migration, and children who died early, are also excluded from all of the main analyses of attrition in this paper.

4. Methods

Analysis of attrition is carried out using logistic regression response models where someone who responded at a particular wave t is given a value of one on the response variable, and zero otherwise. Response variables were created and models were estimated for the age 5 through to age 34 response consecutively. In each case, predictor variables observed at birth and available from wave0 data were used

in the models. The list of wave0 variables used in the models is presented in Table 3 below. Furthermore, a second logistic response model for the age 34 survey was also estimated using variables observed at age 30.

In addition, we give special attention to the transition between wave3 (age16) and wave4 (age 26) when main respondents changed from being parents to cohort members themselves.

5. Wave response and longitudinal participation

Table 1 shows the achieved samples in each wave of the BCS70 cohort from birth to age 34. The definition of response used in this Table is the percentage of the target sample; this was highest at birth; 96 per cent and lowest at age 26 where only 55 per cent of the estimated target sample was observed. Age 26 was the first time cohort members had to opt into the survey for themselves and this is thought to have affected the response rate, given that this was also a postal contact, and there was also a lengthy gap of 10 years since they had last been contacted. At age 34, the latest available data, 61 per cent of the estimated target sample was observed.

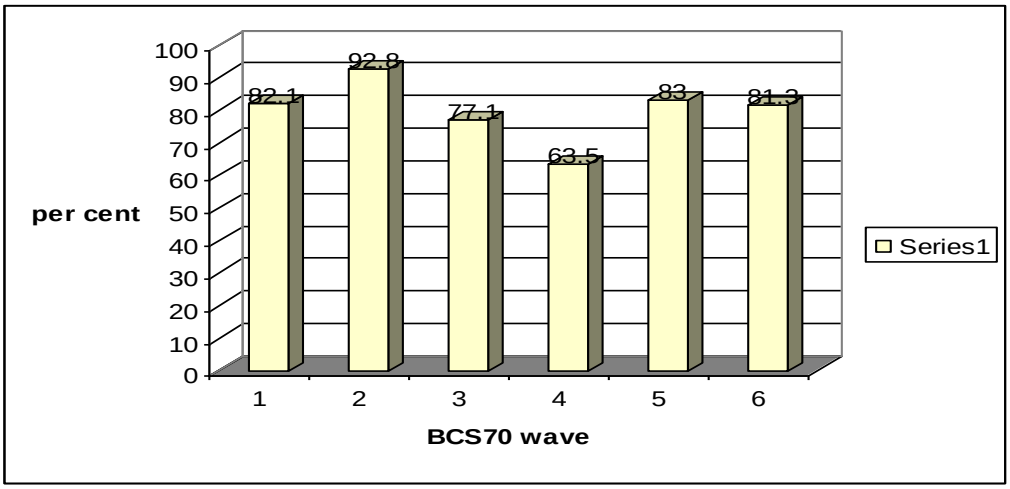
Table 1: BCS70 estimated longitudinal target and observed sample, wave0 to 6

WAVE (AGE)	**wave0 (age 0)**	**wave1 (age 5)**	**wave2 (age10)**	**wave3 (age16)**	**wave4 (age26)**	**wave5 (age30)**	**wave6 (age34)**
Achieved sample (per cent of target sample)	16571 **(95.9%)**	12939 **(79.0%)**	14350 **(88.8%)**	11206 **(70.2%)**	8654 **(55.2%)**	10833 **(70.4%)**	9316 **(60.9%)**
Non-response (per cent of target sample)	716 (4.1%)	2815 (17.2%)	1116 (6.9%)	3328 (20.8%)	4965 (31.7%)	2213 (14.4%)	2137 (14.0%)
Uncertain eligibility (per cent of target sample)	0 (0.0%)	625 (3.8%)	686 (4.2%)	1440 (9.0%)	2063 (13.2%)	2341 (15.2%)	3836 (25.1%)
Target sample **(Estimated)**	17287 (100%)	16379 (100%)	16152 (100%)	15974 (100%)	15682 (100%)	15387 (100%)	15289 (100%)

Note: All cases in the BCS70 are included

The response rates as a percentage of the eligible sample at each wave are shown in Figure 1. These also fluctuate with wave4 being the lowest at 63.5 per cent and wave2 having the highest response at 92.8 per cent. Wave5 and wave6 had higher rates than many earlier waves.

Figure 1. Response rates for successive BCS70 waves as a percentage of the eligible sample.



The proportion of the sample observed in any one wave that was also present in all earlier waves, the so-called 'balanced panel', started out at 98 per cent at age 5 and dropped to 65 per cent at age 26, the first time when the main respondent changed from parents to the cohort members themselves (Table 2). There was a further drop to 45 per cent at age 30 and this percentage stayed constant at age 34. Wave3 response was much lower than expected at age 16, and is thought to have been adversely affected by teachers who were taking industrial action at the time of the survey.

Table 2: BCS70 summary of response in earlier sweeps

WAVE	Wave1 (Age 5)	Wave2 (Age 10)	Wave3 (Age 16)	Wave4 (Age 26)	Wave5 (Age 30)	Wave6 (Age 34)
Observed Sample	12939	14350	11206	8654	10833	9316
Observed in all earlier sweeps	**12692 (98.1%)**	**11716 (81.6%)**	**8911 (79.5%)**	**5643 (65.2%)**	**4973 (45.9%)**	**4287 (46.0%)**
Observed in all but one earlier sweep		2206 (15.4%)	1770 (15.8%)	2183 (25.2%)	3743 (34.6%)	3135 (33.7%)
Observed in all but two earlier sweeps			485 (4.3%)	675 (7.8%)	1591 (14.7%)	1436 (15.4%)
Observed in all but three earlier sweeps				153 (1.8%)	451 (4.2%)	378 (4.1%)
Observed in all but four earlier sweeps					75 (0.7%)	77 (0.8%)
Observed in all but five earlier sweeps						3 (0.0%)
Not observed in any earlier sweep	247 (1.9%)	428 (3.0%)	40 (0.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

Note: All cases in the BCS70 are included

It is worth commenting that BCS70 response rates do not follow what has come to be seen as the expected pattern for longitudinal wave-on-wave responses rates. This common pattern, observed especially in annual household panel studies, sees wave2 response rates suffer the largest fall, but wave-on-wave response rates from

wave3 onwards being consistently high and at rates that are often in the 90 per cent and above range. There are a number of reasons why BCS70 did not realize this common pattern; the reasons include:

- uncertainties in the funding leading to very long gaps between contacts;
- the use of lower quality modes (e.g. postal surveys); and
- a factor that is inherent to birth cohort studies, but not to panel surveys, the change over from parent to cohort member in commitment to the study needing to take place along the way.

6. Predicting wave response

The variables shown in Table 3 are those observed at birth and are used to estimate cross-sectional response at each BCS70 wave from age 5 to 34. Variables available from wave0 were mainly about the parents and the family context into which the child was born. These include the mother's age at the birth of the cohort child, the father's social class using the older classification of the Registrar General's six categories, approximate indicators of mother's and father's level of education, number of children, marital status of parents, whether mother lived in London, whether breastfeeding was attempted and the gender and birth weight of the cohort child. These data show that the original BCS70 cohort contained 52 per cent boys (48% girls); over two thirds of mothers of the cohort children were between 20 and 30 years old at the birth; nearly 60 per cent of the cohort children's mothers and 54.7 per cent of fathers left school at age 15. There was a higher percentage, 5 per cent, of the fathers where their age at leaving school was unknown. Seventy per cent of the cohort child families had up to two children and 92 per cent of the parents were married. Furthermore, 14 per cent of families participated in the age 22 months sub-sample. Other variables from wave0 were tried in the models but dropped because they were insignificant and did not add to the model's explanatory power.[2]

[2] The variables tried but later dropped from the model included: region of birth of the mother; region of birth of the father; region of birth of the mother's mother; wider socio-economic group categories for the father; mother's social class category in 1970; employment status of father in 1970; employment status of mother in 1970.

Table 3: Distribution of independent variables observed at birth which are used in the models

Variable	Sample	
	N	%
Sex of the baby		
Male	8222	51.8
Female	7663	48.2
Total	**15885**	**100**
Mothers age at delivery – GROUPE D		
14-19	1560	9.8
20-24	5690	35.8
25-29	4952	31.2
30-34	2385	15.0
35+	1298	8.2
Total	**15885**	**100**
Social class of father in 1970		
Social class 1	773	4.9
Social class 2	1789	11.3
Social class 3 non-manual	1798	11.3
Social class 3 manual	7050	44.4
Social class 4	2261	14.2
Social class 5	973	6.1
Other/unknown	1241	7.8
Total	**15885**	**100**
Mothers age at leaving school		
14	1051	6.6
15	9398	59.2
16	2634	16.6
17	1148	7.2
18	1654	10.4
Total	**15885**	**100**
Fathers age at leaving school		
14	1241	7.8
15	8696	54.7
16	2161	13.6
17	902	5.7
18	2108	13.3
Unknown (plus Absent fathers)	777	4.9
Total	**15885**	**100**
Parity		
1	5959	37.5
2	5178	32.6
3	2602	16.4
4+	2146	13.5
Total	**15885**	**100**

Variable	Sample	
	N	%
Present marital status		
Single/widowed/divorced/separated	1160	7.3
Married	14725	92.7
Total	**15885**	**100**
Was lactation attempted		
Attempted	6021	37.9
Not attempted	9864	62.1
Total	**15885**	**100**
Mother resides in London		
Not in London	13862	87.3
London	2023	12.7
Total	**15885**	**100**

The results from a multivariate logistic response model using these wave0 variables are displayed Table 4. The same independent variables from wave0 were used to predict response for every BCS70 wave.

Families whose cohort baby was female were more likely to participate at wave2 and wave3 but not at wave1. This was a somewhat unexpected finding. Many studies have shown that women usually display higher response rates to surveys than men. But here it was the parents who were participating in the survey, and not the children themselves. We see that female cohort members, as might be expected, were more likely to participate in waves 4 through 6 when cohort members were the main respondents.

Table 4: Odds ratios of predictors of response measured at baseline (age 0) from a set of cross-sectional logistic regression model

Respondent		Parent			Cohort member		
Parental variables observed in 1970	**Variable levels**	**Wave 1 (age 5)**	**Wave 2 (age 10)**	**Wave 3 (age 16)**	**Wave 4 (age 26)**	**Wave 5 (age 30)**	**Wave 6 (age 34)**
Gender of the cohort baby (Ref: Boy)	Girl	1.062	1.134**	1.283***	1.837***	1.529***	1.513***
Mothers age at birth of CM (Ref: 20-24 years)	14-19	0.735***	0.854*	0.836**	0.774***	0.847**	0.773***
	25-29	1.030	1.069	1.064	1.122**	1.107*	1.132**
	30-34	1.162*	1.154	1.085	1.238***	1.171**	1.252***
	35+	1.199*	1.342**	1.179*	1.300***	1.216*	1.351***
Social class of father in 1970 (Ref: SC 3 manual)	Social Class 1	1.069	0.994	1.229*	1.244*	1.163	1.161
	Social Class 2	1.031	1.029	1.077	1.190**	1.116	1.163*
	Social Class 3 non-manual	1.148	1.252**	1.311***	1.345***	1.311***	1.318***
	Social Class 4	0.951	0.903	0.938	0.879**	0.909	0.876**
	Social Class 5	0.801**	0.848	0.879	0.701***	0.776***	0.760***
	Other or unknown	0.524***	0.656***	0.888	0.925	0.855	0.947
Birth-weight of cohort member in Kilograms		1.697***	1.768***	1.347***	1.382***	1.386***	1.297***

Respondent		Parent			Cohort member		
Parental variables observed in 1970	**Variable levels**	**Wave 1 (age 5)**	**Wave 2 (age 10)**	**Wave 3 (age 16)**	**Wave 4 (age 26)**	**Wave 5 (age 30)**	**Wave 6 (age 34)**
Mother's age at completion of school (Ref:15 yrs)	14	0.716***	0.625***	0.819*	0.773**	0.891	0.811**
	16	1.006	0.890	1.123*	1.046	1.091	1.114*
	17	0.925	0.741***	0.997	1.099	0.997	1.060
	18	0.784**	0.696***	0.981	1.018	0.907	0.972
Father's age at completion of school (Ref:15 yrs)	14	0.834*	0.776**	0.889	0.961	0.831*	0.952
	16	0.905	0.796**	1.030	1.040	0.950	0.966
	17	0.745**	0.819	1.105	1.176*	1.070	1.064
	18	0.697***	0.705***	0.926	0.983	0.942	1.015
	Unknown	0.619***	0.627***	0.683***	0.794*	0.675***	0.706***
Parity (Ref: 1-CM only)	2	0.922	0.960	0.845***	0.889**	0.911*	0.871**
	3	0.773***	0.810**	0.789***	0.746***	0.805***	0.710***
	4+	0.660***	0.815*	0.668***	0.548***	0.630***	0.559***
Mother is married		1.276**	1.701***	1.418***	1.650***	1.587***	1.660***
Breast feeding was not attempted		0.840***	0.856**	0.867***	0.876***	0.932	0.883***
Participated in 22 months sub-sample		2.203***	1.714***	1.369***	1.185***	1.248***	1.193***
Mother resides in London		0.551***	0.547***	0.474***	0.710***	0.616***	0.614***
The estimation sample		15885	15885	15885	15885	15885	15885
Wave observed sample		12939	14350	11206	8654	10833	9316
Pseudo R2		0.060	0.060	0.036	0.047	0.037	0.039
Degrees of freedom		28	28	28	28	28	28
Chi-squared		1016.2	857.5	737.8	1044.6	771.1	844.0

Notes: Exponentiented coefficients (Odd ratios), * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Cases in Northern Ireland at age 0 who migrate to GB before age 5 and participated at age 5 are also excluded

The age of the mother at the birth of the cohort child was a significant predictor of response at all waves with families of younger mothers at birth being less likely to respond and older mothers more likely to respond at each wave.

Social class of the father of the cohort child measured in 1970 was also a significant predictor of response in all waves but more so in waves when cohort members were the main respondent, age 26 and beyond. Families with a father whose social class was non manual were about 30 per cent more likely to participate in waves2 through to wave6 compared to families with a father in the skilled manual social class. Having a father from the lowest social class groups (4 and 5) was generally associated with lower response in all waves in comparison to a father in the skilled manual social class.

The birth weight of the cohort child was a statistically significant predictor of response in all waves with large associations in earlier waves than in later waves. The heavier the cohort child at birth the more likely the family was to participate in the survey. This relationship continued in the same direction at age 26, 30 and 34 surveys when cohort members were the main respondent.

Mothers' and fathers' ages at leaving school were significant predictors of responding at wave1 and wave2 (ages 5 and 10) but not at age 16 (wave3) and above. Compared with mothers or fathers who left school age 15, cohort members leaving school at other ages tended to have lower responses.

Families with more than two children (including the cohort child) were significantly less likely to participate in wave1, wave2 and wave3 and the cohort members who had more than one sibling were also less likely to participate in wave4, wave5 and wave6 than cohort members who were the only child at birth.

Families in which parents were married were significantly more likely to participate in all waves and cohort members whose mothers were married were also more likely to participate in all waves.

Families where no attempt was made to breastfeed the cohort child were significantly less likely to participate in all waves. Cohort members who were not breastfed were also significantly less likely to participate in surveys after age 26, when they had to decide for themselves to respond.

Mothers residing in London at the birth of the cohort child were significantly less likely to participate in all waves (up to and including age 16) and the cohort children were also less likely to participate in all waves from age 26 onwards, than mothers residing outside of London.

7. Wave response before and after a non-response at wave 3

The parents of the cohort children were the main respondents in all waves to the age 16 survey (wave3) whereupon the cohort child had to decide for themselves whether to respond (wave4 onwards). Table 5 shows the distribution of families who participated at wave 3 according to their response at previous waves (wave1 and wave2), and according to their subsequent responses in wave4 to wave6. This table only includes families that participated in at least one wave after wave3. As expected, families that did not respond or participate in wave3 were also less likely to have participated in the previous waves (wave1 and wave2). Similarly, cohort members themselves, whose families had not participated at wave3, were less likely to respond at wave4, wave5 and wave6. However, it is interesting to note that some cohort members participated at wave4, wave5 and wave6 despite their parents being non-responders at wave3, although with slightly lower response rates than cohort members whose families responded in wave3. Unfortunately, we could not separate refusals at wave3 from other types of non-respondents so no further examination was possible with the data.

Table 5. Wave response pre and post age 16 non-response

	Response at age 16				
	Respondents	Non-respondents		Total	
Response	%	%	No.	%	No.
At Age 5 (wave1)					
No	12.8	23.6	761	15.3	2123
Yes	87.2	76.4	2463	84.7	11783
Total (N)	**100(10682)**	**100**	**3224**	**100**	**13906**
Pearson chi2(1) = 225, Pr < 0.001					
At Age 10 (wave2)					
No	5.3	15.5	500	7.7	1071
Yes	94.7	84.5	2724	92.3	12835
Total (N)	**100(10682)**	**100**	**3224**	**100**	**13906**
Pearson chi2(1) = 359, Pr < 0.001					
At Age 16 (wave3)					
No	0	100	3224	23.2	3228
Yes	100	0	0	76.8	10678
Total (N)	**100(10682)**	**100**	**3224**	**100**	**13906**
At Age 26 (wave4)					
No	38.0	48.1	1550	40.3	5608
Yes	62.0	51.9	1674	59.7	8298
Total (N)	**100(10682)**	**100**	**3224**	**100**	**13906**
Pearson chi2(1) = 104, Pr < 0.001					
At Age 30 (wave5)					
No	23.5	31.2	1007	25.3	3517
Yes	76.5	68.8	2217	74.7	10389
Total (N)	**100(10682)**	**100**	**3224**	**100**	**13906**
Pearson chi2(1) = 78, Pr < 0.001					
At Age 34 (wave6)					
No	32.7	46.1	1487	35.8	4982
Yes	67.3	53.9	1737	64.2	8924
Total (N)	**100(10682)**	**100**	**3224**	**100**	**13906**
Pearson chi2(1) = 193, Pr < 0.001					

Predicting response at age 34 using characteristics from age 30

We also decided to focus, as an additional example, on the most recent wave responses of cohort members at age 34. In this case we have derived a set of independent variables from the previous wave at age 30 to use as predictors. However, it is also possible to use several pieces of information about the survey process, namely, whether the address given to the interviewer at wave 5 was correct as printed, and how may calls were made to the house at the wave 5. The other variables chosen to use as predictors included the cohort member's gender, marital status at age 30, smoking status at age 30, their interest in politics at age 30, whether they said at age 30 they intended to move house in the near future, and how

well they were managing financially at age 30. Another set of variables were tried but dropped since they did not offer any significant results, nor did they add to the model's explanatory power.[3] The results are displayed in Table 6.

Columns 3 and 4 show the productive sample number and proportion at age 34 where the age 30 variable in question was not missing. So, for example, of cohort members who were married at age 30, 82.9 per cent, gave productive responses at age 34, higher than those who had been cohabiting (80.3%) or single (76.7%) at age 30. Of the other variables examined using such descriptive statistics, the highest productive responses at age 34 came from those who had never smoked at age 30 (82.9%), were women (82.5%), were not very interested in politics at age 30 (82.2%) said they did not intend to move in the near future (83.6%), and said they were managing their finances comfortably at age 30 (82.9%). Productive responses were also higher where interviewers were given the correct address (81.9%) and where only one call was made to the address (84%) both at the age 30 survey. The lowest productive responses came from cohort members who smoked every day at age 30 (74.9%), were not at all interested in politics at age 30 (76.6%), were finding it difficult to manage at age 30 (70.7%) and had 5 or more calls from the fieldwork agency (75.1%).

The final column in Table 6 shows the odds ratio estimates from a multivariate logistic regression model of cohort members' responses at wave6 (age 34), along with their 95% confidence intervals. We see that cohort members who were men or at age 30 were single, smoked every day, were not at all interested in politics, were not managing well financially, or who intended to move in the near future, they were all significantly less likely to participate at the age 34 survey. Also, if their address was incorrectly printed on the form for the age 30 contact, or they had more than 2 calls made to their address by interviewers at the age 30 survey, they were similarly less likely to respond and participate in the age 34 (wave 6) survey.

[3] Variables tried but dropped from the age 30 interview included: total number of people in the household; whether cohort member has a partner in the household; was the cohort member willing to complete the Self Completion questionnaire; whether the cohort member's name was correct on the ARF details; how often did the cohort member see their mother; how often did the cohort member see their father; had they done any YTS or youth training courses; had a longstanding illness; did illness limit the paid work they could do; religion.

Table 6: Predictors of response at 34 using age 30 characteristics

Variables	Categories	Productive sample at age 34		Total	Multivariate model estimates	
		n	%	N	Odds Ratios	(95% CI)
CM's age 30 marital status	Married	3728	82.9	4497	1	
	Cohabiting	1914	80.3	2383	0.998	(0.873, 1.140)
	Single	2624	76.7	3420	0.829**	(0.737, 0.932)
CM's age 30 smoking status	Never	3778	82.9	4558	1	
	Used to	1627	83.2	1956	1.054	(0.912, 1.218)
	Occasionally	628	78.9	796	0.885	(0.728, 1.076)
	Every day	2268	74.9	3027	0.760***	(0.673, 0.857)
CM's gender	Male	3904	77.6	5029	1	
	Female	4420	82.5	5360	1.264***	(1.141, 1.401)
How interested are you in politics – age 30	Very	343	79.6	431	1	
	Fairly	2472	81.7	3025	0.989	(0.760, 1.287)
	Not very	3118	82.2	3794	0.942	(0.725, 1.224)
	Not at all	2357	76.6	3076	0.747*	(0.575, 0.971)
Was CM's address correct as printed? At wave 5	Yes	6524	81.9	7961	1	
	No	1750	74.2	2357	0.709***	(0.632, 0.796)
Total number of calls made to address wave 5	1	1803	84.0	2146	1	
	2	1980	82.8	2391	0.926	(0.787, 1.089)
	3	1485	80.1	1855	0.799**	(0.676, 0.945)
	4	1067	79.3	1346	0.779**	(0.650, 0.933)
	5+	1987	75.1	2647	0.649***	(0.558, 0.755)
Does CM intend to move in near future age 30	Yes	3264	75.8	4305	1	
	No	4947	83.6	5919	1.551***	(1.401, 1.716)
How well cm managing financially these days age 30	Comfortably	2869	82.9	3460	1	
	Alright	3223	81.4	3960	0.937	(0.829, 1.060)
	Just about	1681	77.1	2180	0.807**	(0.701, 0.929)
	Difficult	529	70.7	748	0.641***	(0.528,0.777)
Model statistics	Estimation sample				10078 (cases excluded with missing item values)	
	Pseudo R2				0.0346	
	Degrees of freedom				18	
	Chi-squared				345.7	

Note: 95% confidence intervals in brackets, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
CM – cohort member

8. Sub studies in BCS70

Since 1970, there have been five sub-studies using the BCS70 cohort members. These were conducted when cohort members were aged 22 months, 2-3 years, 42 months, 7 years and 21 years of age (CLS website). Additionally, at the age 34 main survey, there was an additional element to the survey, called the Parent and Child interview (Simmonds et al, 2007). This element aimed to collect additional information about the cohort members and their children and a questionnaire was administered to one in two cohort members who had already become parents by age 34. This element was an extension of the core interview and was only applied if the cohort member had a natural or adopted children aged under 17 years in the household.

9. Sub-study effects

Sub-studies as part of a longitudinal study could be expected to have one of two main effects. If they are lengthy, and especially if they are administered at the same time as the main survey contact, they risk adding to respondent burden and generating more refusals in future survey waves. On the other hand, if sub-studies are undertaken between survey waves, and especially if the usual gap is fairly lengthy, then they hold the potential to increase involvement in the study, keep the address database more up-to-date for the sub-study group, and thereby raise response rates among sub-study participants in subsequent waves. There is some limited literature relevant to these issues, as summarised below.

Respondent burden is typically concerned with the length of questionnaires. Bogen (1996) reviewed the research on the effect of questionnaire length on response rates for the first and subsequent waves. Bogen suggests there is some evidence that interview length did not have a significant effect on subsequent response rates in the context of panel surveys (Frankel and Sharp, 1981; Sharp and Frankel, 1983; Taylor and Lynn, 1996). Branden, *et al* (1995) found that interview length (time) or questionnaire length (number of questions) had either no effect or a positive effect on sample retention rates. However, Zabel (1994) reported that attrition rates were reduced after a decrease in interview length. Kantorowitz (1998) found no increase in refusal rate, either at current or subsequent waves when the Israeli Labour Force Survey was lengthened. She also found no relationship between interview length and response rate when the supplementary questionnaire was drastically shortened. Galvin *et al* (2000), using the SIPP, found that people whose interviews were short were much more likely to drop out of the panel than people whose interviews were longer. However, when looking at respondent burden in a longitudinal study there is an additional component. This is the perceived cost to cohort or panel members of future survey participation. The effect of the perceived longitudinal burden on the survey participation resulted in a 5 per cent decrease in the response rate in one study (Apodaca, *et al*, 1998). Lynn, *et al* (1997) found that explicitly telling sample members that the survey was longitudinal, but only at a later wave, resulted in slight reduction in response at that wave, but an overall improved net response at subsequent waves.

Features of the survey design can also affect the possibility of locating respondents. Such features include the length of the panel, the length of the gap between interviews and the extent and nature of contact between the survey organisation and sample member between waves (Laurie *et al*, 1999; Lepkowski and Couper, 2002; Lynn, 2003; Taylor and Lynn, 1996). Shortening the gap can have beneficial effects on the non-contact component of attrition therefore.

The initial aim of this exercise was to assess the impact of sub-studies on the main survey waves in British Cohort Study (BCS70). In order to carry out a rigorous evaluation of the effects of sub-studies on attrition in BCS70, data are needed about which people were approached to be in the study, and how they were selected. Given that selection of cases into a sub-study is likely to be systematic rather than random, the nature of this selection process needs to be known in order to allow for

this in attempting to evaluate the effect of being in a sub-study. It is possible that the selection into a sub-study has captured, either knowingly or unknowingly, the most motivated of the whole cohort. This would lead the sub-study participants' group response at subsequent waves to be higher than that for non-participants and compared with response rates for the whole sample, but not because of any effects of being in the sub-study. The average response for non-participants would also appear to decline compared with the average for the whole sample prior to the sub-study, from having removed the best responders from this group.

Unfortunately BCS70 records did not offer full information about who was selected and approached to be in the sub studies, and the criteria used in selection. All that is known is which people actually took part in the sub-study. Because of the un-availability of details about sub-study sampling and response data, we were only able to look at future and past response rates for those who participated in the sub-studies and those who did not. Furthermore, data for one of the sub-studies, at age 7 years was not deposited in the UK Data Archive and so was unavailable for analysis. Data for the age 2-3 years sub-study was also not used because it only covered the South-West region and nearly all cohort member families participated.

Records of participants were available for the following BCS70 sub-studies:

- Participants in the age 22 months sub-study;
- Participants in the age 42 months sub-study;
- Participants in the 21 years sub-study; and
- Participants in the parent and child sub-sample which took place along side the main age 34 interview.

However, on investigation, the response patterns of those who were in the age 42 month sub-study were very similar to those who participated in the age 22 months, so only one of these, the age 22 months sub-study, is reported in this paper.

The age 22 month and age 21 year sub-studies, took place between other main BCS70 waves, between age 0 and age 5 in the case of the 22 month sub-study, and between age 16 and age 26 in the case of the 21 year sub-study. It is likely that the main effect of such sub-studies is likely to be on being able to contact cohort members and have their up-to-date addresses in the address database. The Parent and Child sub-study took place along side the age 34 main interview. It runs the risk, therefore, of increasing respondent burden for sub-study participants. However, no data were available at the time of writing about response in subsequent waves of BCS70 (the age 38 data will be available in the near future to examine the effects).

10. Age 22 month sub-study

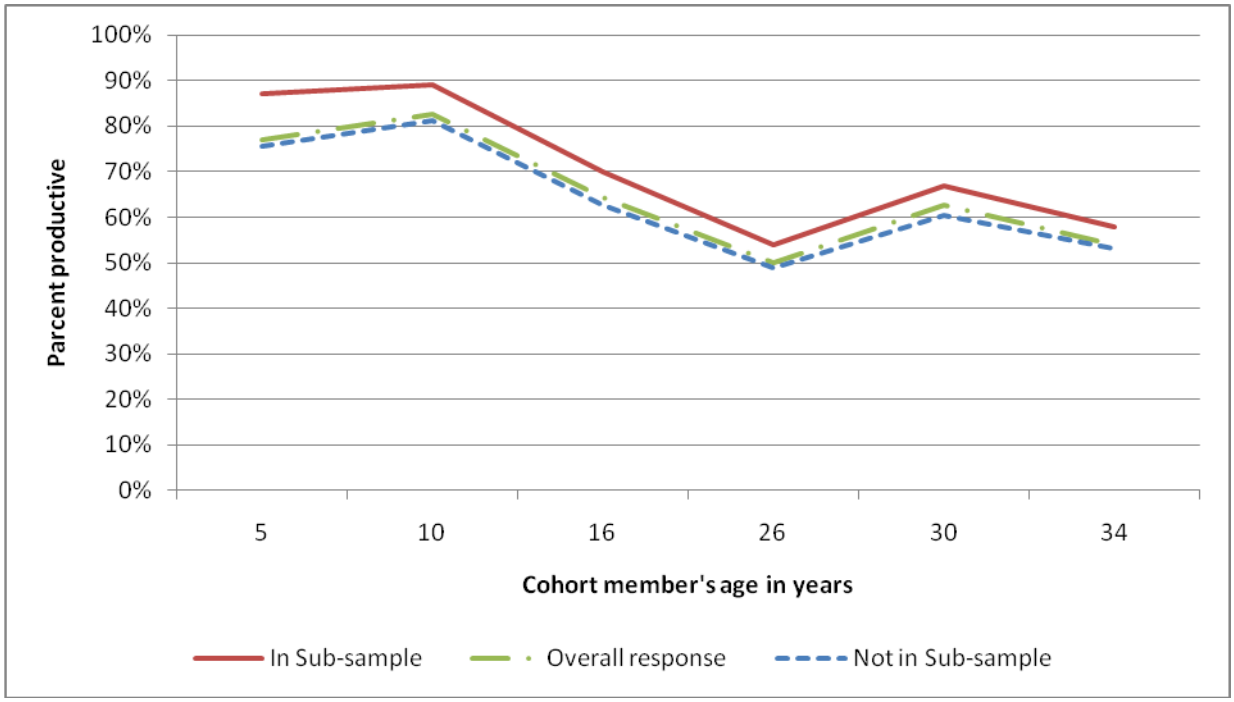
The number of cohort family participants in the age 22 month sub-study was 2361, approximately 15 per cent of the wave0 sample. Table 7 shows a comparison between participants in the age 22 months sub-study and those who were not in the study. In all waves after the sub-study, those who participated in the sub-study were more likely to participate than those who either did not participate or were not

sampled into the sub-study. The same productive response rate data are also displayed in Figure 2. These results are in line with our expectations about the effects of this type of sub-study.

Table 7. Response rates of participants in the age 22 months sub-study in all future waves

Wave	Participated in 22 months sub-study				
	No	Yes		Total	
Productive at Age 5	**%**	**%**	**No.**	**%**	**No.**
No	24.9	12.7	300	23.1	3838
Yes	75.1	87.3	2061	76.9	12744
Total (N)	**100 (14221)**	**100**	**2361**	**100**	**16582**
Pearson chi2(1) = 168.6522 Pr = 0.000					
Productive at Age 10					
No	18.5	11.0	259	17.4	2890
Yes	81.5	89.0	2102	82.6	13692
Total (N)	**100(14221)**	**100**	**2361**	**100**	**16582**
Pearson chi2(1) = 79.7976 Pr = 0.000					
Productive at Age 16					
No	36.6	29.4	693	35.6	5904
Yes	63.4	70.6	1668	64.4	10678
Total (N)	**100(14221)**	**100**	**2361**	**100**	**16582**
Pearson chi2(1) = 46.9466 Pr = 0.000					
Productive at Age 26					
No	50.6	46.0	1085	50.0	8284
Yes	49.4	54.0	1276	50.0	8298
Total (N)	**100(14221)**	**100**	**2361**	**100**	**16582**
Pearson chi2(1) = 17.6427 Pr = 0.000					
Productive at Age 30					
No	38.1	32.5	768	37.3	6193
Yes	61.9	67.5	1593	62.7	10389
Total (N)	**100(14221)**	**100**		**100**	**16582**
Pearson chi2(1) = 27.3237 Pr = 0.000					
Productive at Age 34					
No	46.9	41.8	987	46.2	7658
Yes	53.1	58.2	1374	53.8	8924
Total (N)	**100(14221)**	**100**		**100**	**16582**
Pearson chi2(1) = 21.2329 Pr = 0.000					

Figure 2. Response rates of BCS70 cohort members according to whether they were in the age 22 month sub-study



Having been in the 22 month sub-study was also entered as a dummy variable in the multivariate regression analysis carried out on each BCS70 wave of data (Table 4). This enabled us to see whether, after controlling for other factors, participating in the study was still associated with higher productive response at each BCS70 wave. The results in Table 4 show that the significantly higher response from being in the 22-month sub-study was confirmed, after controlling for a range of other characteristics. The largest effect was seen at wave1 (age 5), with smaller but still significant effects seen up to age 34.

Closing the 5-year gap between main survey contacts (ages 0 and 5) at 22 months, not quite half way through the period, may have been responsible for the beneficial effects on response rates. For example, the contact details of sub-study participants may have been updated, and this may have helped to raise their productive response rates. However, we cannot be sure that the sub-study is responsible for these effects without more rigorous analysis. We can note that this effect did hold up after controlling for other characteristics. There was a substantial 12 per cent higher productive response at age 5 from those who participated in the sub-study, compared with those who did not. The fact that this benefit to response rates continued in the later waves, albeit not with a narrower gap, is also something to note and investigate further in other studies.

11. Age 21 years sub study

The age 21 year sub-study involved 1522 participants, and again took place half way between, in this case, a ten year gap between main survey waves at 16 and 26. Table 8 shows a comparison between participants in the age 21 years sub-study and

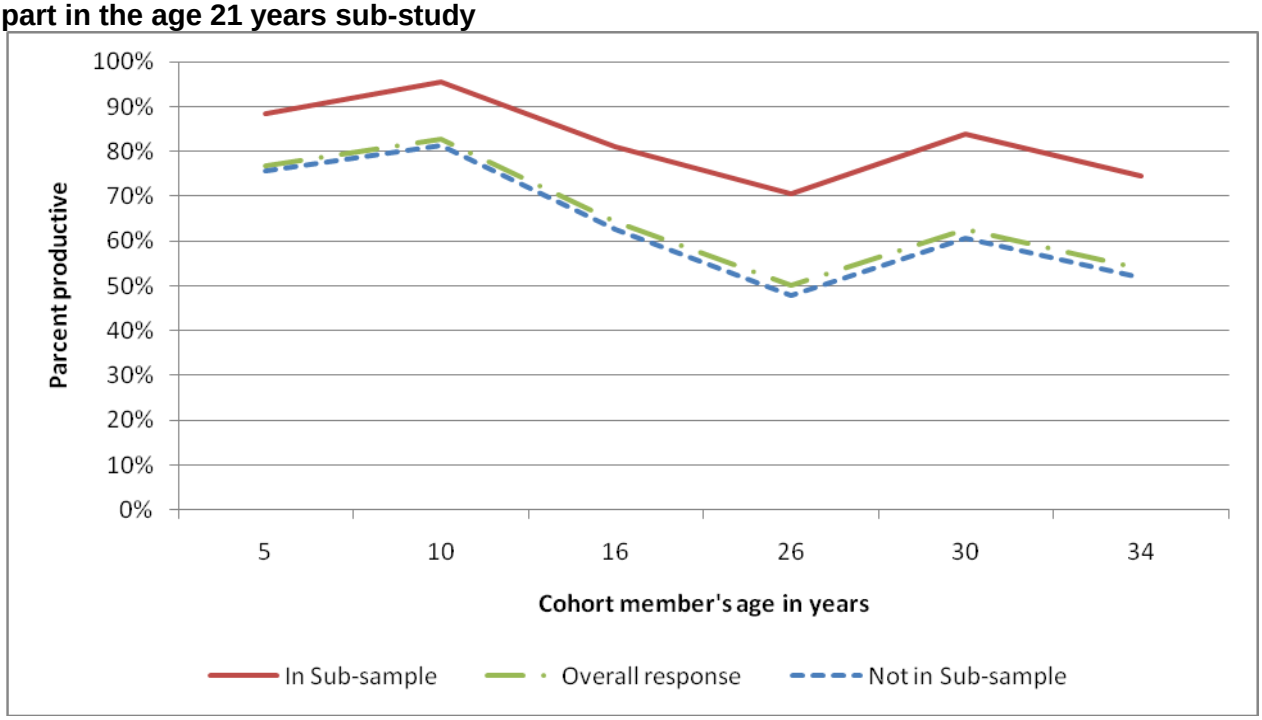
those who were not participants. In all waves before and after the sub-study, those who participated in the sub-study were more likely to participate in the main surveys than those who either did not participate or were not sampled into the sub-study. The same productive response rates are also displayed in Figure 3 for sub-study participants and non-participants. As was the case for the 22 month sub-study, these results are in line with our expectations about the effects of this type of sub-study. Closing this 10-year gap between main survey contacts (ages 16 and 26) half way after 5 years with another contact, may have been responsible for these beneficial effects on response rates. This additional contact may have helped to update the contact details of sub-study participants and in this way helped to raise the productive response rates of sub-study participants. At the next interview following the sub-study, its participants had a productive response rate of 70.6 per cent, compared to the equivalent rate for non participants of 48 per cent. This is a huge gap of 22.4 per cent. However, without more rigorous analysis we cannot be sure this is an effect from the sub-study. We should also note that at the waves prior to the age 21 sub-study, the sub-study participants also had higher productive response rates than non-participants at these earlier interviews. This points to sub-study participants being a more highly motivated group from the beginning. There were other issues to consider around the time of the age 26 interview which may also have played some role: that the age 26 interview was one in which decisions on participation switched from parents to cohort members, but that in addition, this was a postal questionnaire. The fact that benefits to response rates of similar magnitudes continued for sub-study participants in the later waves is also something to note and investigate further in other studies.

Table 8. Wave response rates of BCS70 cohort members according to whether they participated in the age 21 years sub-study.

Respondents	**Wave**	**Participant in 21 Year Sub-study**				
		No	**Yes**		**Total**	
	Productive Age 5	**%**	**%**	**No.**	**%**	**No.**
Parent	No	24.3	11.7	178	23.1	3838
	Yes	75.7	88.3	1344	76.9	12744
	Total (N)	**100(15060)**	**100**	**1522**	**100**	**16582**
	Pearson chi2(1) = 123.5191, Pr = 0.000					
	Productive at Age 10					
Parent	No	18.7	4.6	70	17.4	2890
	Yes	81.3	95.4	1452	82.6	13692
	Total(N)	**100(15060)**	**100**	**1522**	**100**	**16582**
	Pearson chi2(1) = 191.6652, Pr = 0.000					
	Productive at Age 16					
Parent	No	37.3	19.0	289	35.6	5904
	Yes	62.7	81.0	1233	64.4	10678
	Total(N)	**100(15060)**	**100**	**1522**	**100**	**16582**
	Pearson chi2(1) = 201.8151, Pr = 0.000					

Respondents	Wave	Participant in 21 Year Sub-study				
		No	Yes		Total	
Cohort member	**Productive at Age 26**					
	No	52.0	29.4	448	50.0	8284
	Yes	48.0	70.6	1074	50.0	8298
	Total(N)	**100(15060)**	**100**	**1522**	**100**	**16582**
	Pearson chi2(1) = 282.3329, Pr = 0.000					
Cohort member	**Productive at Age 30**					
	No	39.5	16.2	246	37.3	6193
	Yes	60.5	83.8	1276	62.7	10389
	Total(N)	**100(15060)**	**100**	**1522**	**100**	**16582**
	Pearson chi2(1) = 321.4205, Pr = 0.000					
Cohort member	**Productive at Age 34**					
	No	48.3	25.6	390	46.2	7658
	Yes	51.7	74.4	1132	53.8	8924
	Total(N)	**100(15060)**	**100**	**1522**	**100**	**16582**
	Pearson chi2(1) = 284.9741, Pr = 0.000					

Figure 3. Wave response rates of BCS70 cohort members by whether they took part in the age 21 years sub-study



12. Age 34 parent and child sub-sample

The age 34 Parent and Child sub-sample consisted of 2846 parents who were given an additional questionnaire as part of the age 34 main contact. Table 9 displays a comparison of some characteristics of participants in the Parent and Child sub-study, and those in the core sample only. This comparison was done for the whole sample,

and for men and women cohort members separately. There were more female cohort members in the Parent and Child sub-sample in Wales than in other GB countries. There were more men in the Parent and Child sub-sample who were interviewed in the previous survey at age 30 than those who did not participate at age 30. There were more married cohort members in the Parent and Child sub-sample than non-married cohort members. The proportion of cohort members in the Parent and Child sub-sample decreased with increasing levels of education across all comparison groups.

We might worry that this sub-study had the potential to increase respondent burden and risk reducing productive responses at the next wave. It is unfortunately, as yet, too early to tell.

Table 9. Characteristics of the BCS70 age 34 Parent and Child sub-sample

Variable	Whole sample			Men			Women		
	Core sample	Parent and Parent+child	Total	Core sample	Parent and Parent+child	Total	Core sample	Parent and Parent+child	Total
	%	%	%	%	%	%	%	%	%
Country at interview									
England	85.8	84.7	85.5	85.5	84.7	85.3	86.1	84.7	85.6
Wales	4.9	6.1	5.3	5.2	5.2	5.2	4.6	6.7	5.4
Scotland	9.3	9.2	9.2	9.3	10.2	9.5	9.3	8.5	9.0
Total %	100	100	100	100	100	100	100	100	100
N	6815	2846	9661	3501	1122	4623	3314	1724	5038
Statistics	chi2(2) = 5.7756 , Pr = 0.056			chi2(2) = 0.7653 , Pr = 0.682			chi2(2) = 10.0919 , Pr = 0.006		
Whether interviewed in previous wave									
Yes	92.5	94.6	93.1	91.0	93.4	91.6	94.1	95.4	94.6
No	7.5	5.4	6.9	9.0	6.6	8.4	5.9	4.6	5.4
Total %	100	100	100	100	100	100	100	100	100
N	6819	2846	9665	3504	1122	4626	3315	1724	5039
Statistics	chi2(1) = 13.9008 , Pr = 0.000			chi2(1) = 6.4627 , Pr = 0.011			chi2(1) = 3.5689 , Pr = 0.059		
Marital status - de facto (derived)									
Married	46.6	72.2	54.1	42.4	79.7	51.4	51.0	67.3	56.6
Cohabiting (living as a couple)	21.9	16.8	20.4	23.2	18.4	22.0	20.5	15.8	18.9
Single (and never married)	25.4	5.3	19.5	29.1	0.4	22.1	21.5	8.6	17.1
Separated, divorced and widowed	6.2	5.6	6.0	5.4	1.6	4.4	7.0	8.2	7.4
Total %	100	100	100	100	100	100	100	100	100
N	6795	2845	9640	3489	1122	4611	3306	1723	5029
Statistics	chi2(3) = 682.8535 , Pr = 0.000			chi2(3) = 582.7273 , Pr = 0.000			chi2(3) = 179.0438 , Pr = 0.000		
Cohort member's age at interview									
33	13.3	11.9	12.8	12.8	11.9	12.5	13.8	11.9	13.1

Variable	Whole sample			Men			Women		
	Core sample	Parent and Parent+child	Total	Core sample	Parent and Parent+child	Total	Core sample	Parent and Parent+child	Total
Continued									
34	85.5	87.0	85.9	85.9	86.7	86.1	85.1	87.1	85.8
35	1.3	1.2	1.2	1.4	1.4	1.4	1.1	1.0	1.1
Total %	100	100	100	100	100	100	100	100	100
N	6784	2846	9630	3479	1122	4601	3305	1724	5029
Statistics	chi2(2) = 3.5991 , Pr = 0.165			chi2(2) = 0.6447 , Pr = 0.724			chi2(2) = 3.7596 , Pr = 0.153		
NS-SEC 8 analytic version (derived)									
Higher managerial and professional occupations	18.1	13.6	16.8	21.8	19.5	21.2	13.5	8.1	11.8
Lower managerial and professional occupations	31.5	29.7	31.0	27.0	27.5	27.2	37.1	31.6	35.3
Intermediate occupations	12.5	13.0	12.6	6.8	5.5	6.5	19.6	19.9	19.7
Small employers and own account workers	9.0	10.2	9.3	11.9	13.6	12.3	5.3	7.0	5.9
Lower supervisory and technical occupations	11.2	10.6	11.0	16.0	15.7	15.9	5.2	6.0	5.4
Semi-routine occupations	10.7	14.6	11.8	8.1	8.1	8.1	14.0	20.5	16.1
Routine occupations	7.1	8.3	7.4	8.4	10.0	8.8	5.4	6.8	5.8
Total %	100	100	100	100	100	100	100	100	100
N	5740	2249	7989	3188	1071	4259	2552	1178	3730
Statistics	chi2(6) = 49.1212 , Pr = 0.000			chi2(6) = 8.5764 , Pr = 0.199			chi2(6) = 55.6535 , Pr = 0.000		
Employment status									
Self-employed : small establishment (1-24 employees)	2.9	3.3	3.0	4.0	4.9	4.2	1.6	1.8	1.6
Self-employed : no employees	9.1	8.5	8.9	11.4	10.4	11.2	6.1	6.9	6.4
Manager : large establishment (25+ employees)	10.6	8.8	10.1	12.0	13.2	12.3	8.9	4.7	7.6
Manager : small establishment (1-24 employees)	5.3	5.0	5.2	5.3	6.8	5.6	5.4	3.3	4.8
Foreman or supervisor	27.6	25.2	26.9	28.4	27.0	28.0	26.6	23.6	25.6
Employee (not elsewhere classified)	44.5	49.2	45.8	38.9	37.7	38.6	51.4	59.7	54.0
Total %	100	100	100	100	100	100	100	100	100
N	5752	2252	8004	3196	1072	4268	2556	1180	3736
Statistics	chi2(5) = 18.4170 , Pr = 0.002			chi2(5) = 7.7327 , Pr = 0.172			chi2(5) = 39.6884 , Pr = 0.000		

Variable	Whole sample			Men			Women		
	Core sample	Parent and Parent+child	Total	Core sample	Parent and Parent+child	Total	Core sample	Parent and Parent+child	Total
Continued									
Highest academic qualification – (derived)									
None	9.6	8.7	9.3	11.3	8.4	10.6	7.9	8.9	8.2
CSE	14.4	17.0	15.1	14.7	17.1	15.3	14.0	16.9	15
GCE O level/ GCSE	31.5	36.4	33.0	32.0	37.8	33.4	31.0	35.5	32.5
A level/ SSCE/ A-S level	9.5	8.6	9.2	9.4	8.0	9.1	9.5	8.9	9.3
Degree/ dip. h.ed/ oth teaching qual./ nursing qual	28.2	24.3	27.1	26.1	23.1	25.4	30.5	25.1	28.6
Higher degree / PGCE	6.8	5.0	6.3	6.5	5.6	6.3	7.1	4.6	6.3
Total %	100	100	100	100	100	100	100	100	100
N	6783	2841	9624	3484	1119	4603	3299	1722	5021
Statistics	chi2(5) = 48.0435 , Pr = 0.000			chi2(5) = 23.9573 , Pr = 0.000			chi2(5) = 37.5106 , Pr = 0.000		
Computer at home									
Yes	79.9	86.4	81.8	80.1	87.8	81.9	79.7	85.5	81.7
No	20.1	13.6	18.2	19.9	12.2	18.1	20.3	14.5	18.3
Total %	100	100	100	100	100	100	100	100	100
N	6791	2843	9634	3488	1121	4609	3303	1722	5025
Statistics	chi2(1) = 56.4617 , Pr = 0.000			chi2(1) = 34.0390 , Pr = 0.000					
Smoking habits (derived)									
Never smoked	45.2	45.0	45.1	43.4	44.4	43.7	47.1	45.3	46.5
Ex smoker	23.1	24.6	23.5	22.2	25.3	22.9	24.1	24.2	24.1
Occasional smoker	7.2	5.6	6.7	7.7	5.9	7.3	6.6	5.4	6.2
Up to 10 a day	9.2	9.7	9.4	8.4	8.1	8.4	10.1	10.8	10.3
11 to 20 a day	13.1	13.2	13.1	15.1	13.7	14.8	11.0	12.8	11.6
More than 20 a day	2.2	1.9	2.1	3.1	2.4	2.9	1.2	1.5	1.3
Daily but frequency not stated	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.1	0.0
Total %	100	100	100	100	100	100	100	100	100

Variable	Whole sample			Men			Women		
	Core sample	Parent and Parent+child	Total	Core sample	Parent and Parent+child	Total	Core sample	Parent and Parent+child	Total
Continued									
N	6791	2843	9634	3488	1121	4609	3303	1722	5025
Statistics	chi2(6) = 10.8185 , Pr = 0.094			chi2(6) = 10.3666 , Pr = 0.110			chi2(6) = 8.2078 , Pr = 0.223		
BMI weight status category (derived)									
Underweight (<18.5)	1.4	1.7	1.5	0.8	0.4	0.7	2.1	2.6	2.3
Normal (18.5-24.9)	48.6	47.0	48.1	39.6	36.2	38.8	58.3	54.1	56.9
Overweight (25-29.9)	33.9	33.6	33.8	42.8	43.1	42.9	24.2	27.2	25.3
Obese (30 and above)	16.1	17.8	16.6	16.8	20.4	17.7	15.4	16.1	15.6
Total %	100	100	100	100	100	100	100	100	100
N	6582	2773	9355	3434	1110	4544	3148	1663	4811
Statistics	chi2(3) = 5.5070 , Pr = 0.138			chi2(3) = 10.5017 , Pr = 0.015			chi2(3) = 8.8195 , Pr = 0.032		
Whether any kids in household (derived)									
Yes	34.8	70.2	45.5	28.1	61.0	36.3	41.6	76.0	53.5
No	65.2	29.8	54.5	71.9	39.0	63.7	58.4	24.0	46.5
Total %	100	100	100	100	100	100	100	100	100
N	6265	2692	8957	3153	1047	4200	3112	1645	4757
Statistics	chi2(1) = 948.6321 , Pr = 0.000			chi2(1) = 368.5904 , Pr = 0.000			chi2(1) = 510.2204 , Pr = 0.000		
Any kids aged 0-13yrs in household (derived)?									
Yes	34.5	70.1	45.2	27.7	60.7	35.9	41.4	76.0	53.4
No	65.5	29.9	54.8	72.3	39.3	64.1	58.6	24.0	46.6
Total %	100	100	100	100	100	100	100	100	100
N	6265	2692	8957	3153	1047	4200	3112	1645	4757
Statistics	chi2(1) = 960.7506 , Pr = 0.000			chi2(1) = 373.1176 , Pr = 0.000			chi2(1) = 516.7151 , Pr = 0.000		

21. Conclusions

This paper has examined response in longitudinal surveys by analysing the available data from the 1970 British Birth Cohort study. The pattern of responses across waves varied to some extent from patterns noted in other longitudinal studies (mainly annual household panel surveys) but given the inconsistent and uncertain support for this study in its early stages, it is perhaps surprising that it has held up as well as it has. The predictors of response and non-response from around the time of the cohort child's birth were found to have systematic elements across most if not all waves of data. In many cases, these were consistent with the findings from analyses of response in other longitudinal studies. Response was lower for cohort members who were men, having a mother who was younger at the birth, a mother who did not attempt to breastfeed, a lower birth weight baby, in a family with 2 or more children, born of non-married parents, a manual father and living in London. Many of these findings are indicators of comparative disadvantage and in addition, they were visible as indicators right from the start of the birth cohort study. This general finding about disadvantage as a marker of low response was reinforced by the separate analysis of age 34 responses; there, not managing financially was also associated with lower response. The practical implications of these findings point to the need to offer respondents from disadvantaged circumstances some sort of monetary incentive to participate in longitudinal surveys. The fact that interest or lack of interest in politics was not a consistent predictor associated with response suggests we cannot rely on civic engagement for continuing participation in longitudinal studies.

Birth cohort studies need to be particularly careful around the time they plan to switch over from parents consenting to being in a study to the child making this decision for themselves. Sending postal questionnaires at this point is not advisable, nor is a very long gap after the child reaches maturity, before another cohort contact is made. Sub-studies may have beneficial effects on response when they close the gap between contacts although more rigorous analysis is needed to be sure. Clearly it would always be better to have a wave contact that captured data for all of the longitudinal survey members. None the less, where sufficient resources are not available for a full wave contact, these results, if confirmed in other rigorous studies, suggest it is worth carrying out a smaller sub-study of some of the members, for the boost it may give to their response. If sub-studies can be rotated around different sub-samples of cohorts, this may prove to be an effective way to generate a sustained increase in response over the life of the survey among a large group of cohort members.

Lastly, it is important for full records to be archived of the details of sampling and selection involved in carrying out sub-studies, as well as the data collected, so that rigorous evaluations of their effects can be carried out later.

References

Allen, K., Simpson, G., Winn, D., and Giovino, G. (1991) 'Response rates and characteristics of adolescents in the 1989 teenage attitudes and practices survey' in *Proceedings of Survey Research Methods Section of the American Statistical Association*, 575-580.

Apodaca, R., Lea, S. and Edwards, B. (1998) The effect of longitudinal burden on survey participation, *1998 Proceedings of Survey Research Methods Section of the American Statistical Association,* 906-910.http://www.amstat.org/sections/srms/proceedings
BCS70 website
http://www.cls.ioe.ac.uk/studies.asp?section=0001000200020011 [accessed in March 2009].

BCS70 22-months survey guide to dataset.
http://www.cls.ioe.ac.uk/studies.asp?section=00010002000200110001 [accessed on 10 October2007].

Behr, A., Bellgardt, E., and Rendtel, U. (2005) 'Extent and determinants of panel attrition in the European Community Household Panel, *European Sociological Review*, Vol. 21 (5), pp. 489-512.

Bogen, K. (1996) The Effect of Questionnaire Length on Response Rates - A Review of the Literature, *1996 Proceedings of the Survey Research Methods Section of the* 67 *American Statistical Association*, 1020-1025.http://www.amstat.org/sections/srms/proceedings

Branden, L., Gritz, R.M., and Pergamit, M.R. (1995) The effect of interview length on nonresponse in the National Longitudinal Survey of Youth, *Proceedings of the 1995 Census Bureau Annual Research Conference*, 129-154.

Campenelli, P., Sturgis, P. and Purdon, S. (1997) *Can you hear me knocking: An investigation into the impact of interviewers on survey response rates*, London: National Centre for Social Research.

Fitzgerald,J., Gottschalk,P. and Moffitt,R. (1998a) 'Analysis of sample attrition in panel data: the Michigan Panel Study of Income Dynamics' *Journal of Human Resources*, Vol. 33 (2), 300-344.

Fitzgerald, J., Gottschalk, P. and Moffitt, R. (1998b) 'An analysis of the impact of sample attrition on the second generation of respondents in the Michigan Panel Study of Income Dynamics', *Journal of Human Resources*, Vol. 33 (2), 251-299.

Foster, K. (1998) 'Evaluating non-response on household surveys: Report of a study linked to the 1991 Census' Government Statistical Service Methodology Series No.8.

Frankel, J., and Sharp, L.M. (1981) Measurement of Respondent Burden, *Statistical Reporter* 81(4), 105-111.

Galvin, L.N., Sae-Ung, S. and King, K. (2000) Effect of Interview Length and Proxy Interviews on Attrition to the Survey of Income and Program Participation, *2000Proceedings of Survey Research Methods Section of the American Statistical Association,* 636-640. http://www.amstat.org/sections/srms/proceedings

Groves, R. M. and Couper, M.P. (1998) '*Non response in Household Interview Survey*s, New York: Wiley.

Hawkes, D. and Plewis, I. (2006) 'Modelling non-response in the National Child Development Study,*Journal of the Royal Statistical Society Series A (Statistics in Society)* 169 (3): 479-491.

Kantorowitz, M. (1998) Is it true that non response rates in a panel survey increase when supplement surveys are annexed? 121-137 in Koch, A. and Porst, R. (ed.s) *Nonresponse in Survey Research*, ZUMA Nachrichten Spezial No. 4, Mannheim: ZUMA.

Lepkowski, J.M. and Couper,M.P. (2002) 'Non response in the Second Wave of Longitudinal Household Surveys' in Groves,R.M. Dillman,D.A. Eltinge,J.L. and Little,R.J.A. (2002) *Survey Nonresponse*, New York: John Wiley and Sons, pp. 259-72.

Lillard, L.A. and Panis, C.W.A. (1998) 'Panel attrition from the Panel Study of Income Dynamics: Household income, marital status, and mortality' *Journal of Human Resources*, 33 (2), pp.437- 457.

Lynn, P. (2006) 'Editorial: Attrition and non-response' *Journal of the Royal Statistical Society Series A (Statistics in Society)* 169 (3): 393-94.

Lynn, P. and Clarke, P. (2001) 'Separating refusal bias from non-contact bias: Evidence from UK National Surveys', ISER Working Paper No. 2001-24. Colchester: University of Essex.

Lynn, P. Clarke, P. Martin, J. and Sturgis, P. (2002) 'The effects of extended interview efforts on nonresponse bias 'in Groves, R.M. Dillman, D.A. Eltinge, J.L. and Little, R.J.A. (2002) *Survey Nonresponse*, New York: John Wiley and Sons, pp.135-148.

Lynn, P., Taylor, B. and Brook, L. (1997) Incentives, Information and Number of Contacts: Testing the effects of these factors on response to a panel survey, *Survey Methods Centre Newsletter* **17(3)**, 7-12.

Macurdy,T., Mroz,T. and Gritz, R.M. (1998) 'An evaluation of the National Longitudinal Survey of Youth', *Journal of Human Resources*, 33, 345-36.

Nicoletti, C. and Peracchi, F. (2005) 'Survey response and survey characteristics: microlevel evidence from the European Community Household Panel', *Journal of the Royal Statistical Society Series A (Statistics in Society)* 168: 763-781.

Olsen, R. (2005) 'The problem of survey attrition: Survey Methodology is key', *Monthly Labor Review*, 128, pp.63-70.

Plewis, I. (2004) *National Child Development Study and 1970 British Cohort Study Technical Report: Changes in the NCDS and BCS70 Populations and Samples over Time (1st ed.)* London: Institute of Education, University of London.

Plewis, I. (2007) 'Non-response in a birth cohort study: The case of the Millennium Cohort Study', *International Journal of Social Research Methodology*, Vol.10 (5) pp.325-34.
Plewis, I., Calderwood,L., Hawkes, D., Nathan, G. (2004). Changes in the NCDS and BCS70 populations and samples over time. London, Centre for Longitudinal Studies.

Plewis, I., Calderwood, L., Hawkes, D., Nathan, G. (2004). Changes in the NCDS and BCS70 populations and samples over time. London, Centre for Longitudinal Studies.

Sharpe, L.M., and Frankel, J. (1983) Respondent Burden: A Test of Some Common Assumptions, *Public Opinion Quarterly* **47(1)**, 36-53.

Simmonds, N., Fuller, E., Lessof, C. and Foudouli, V. 2007: *"1970 British Cohort Study: 2004-2005 Survey Technical Report"*, http://www.cls.ioe.ac.uk/studies.asp?section=00010002000200140003 [accessed in March 2009].

Taylor, S. and Lynn, P. (1996) England and Wales Youth Cohort Study (YCS): The effect of time between contacts, questionnaire length, personalisation and other factors on response to the YCS, *Department for Education and Employment Research Paper no.8*, Sheffield: DfEE.

Vandercasteele, L. and Debels, A. (2006) 'Attrition in panel data: The effectiveness of weighting' in *European Sociological Review*, Vol. 23 (1) , pp. 81-97.

Watson, D. (2003) 'Sample attrition in waves 1 and 5 in the European Community Household Panel', *European Sociological Review*, Vol. 19 (4) , pp.361-78

Zabel, J.E. (1994) A Comparison of Attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation, unpublished manuscript, US Census Bureau.

Centre for Longitudinal Studies
Institute of Education
20 Bedford Way
London WC1H 0AL
Tel: 020 7612 6860
Fax: 020 7612 6880
Email cls@ioe.ac.uk
Web http://www.cls.ioe.ac.uk