The influence of life course socioeconomic position on cognitive function and cognitive decline in older age: the impact of missing data

Rebecca Landy

Thesis submitted for the degree of Doctor of Philosophy
University College London

Department of Epidemiology and Public Health
University College London

2012

I, Rebecca Landy, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Cognitive function has been associated with many physical and mental health conditions, as well as mortality. Cognitive decline is one aspect of ageing that causes anxiety among the general population. Understanding the risk factors which affect cognitive function over the life course is therefore important. One potential risk factor is socioeconomic position (SEP).

This thesis investigates the impact of SEP across the life course on crystallized cognitive function and memory decline. The 1946 British birth cohort study and Whitehall II study of British civil servants were used for these analyses. Missing data is a potential source of bias in longitudinal studies, with both SEP and cognitive function predictive of dropout. This thesis therefore considers the impact of methods for dealing with missing data on the findings. A complete case analysis is compared with multiple imputation and Heckman selection models.

To compare the suitability of these methods a simulation study was carried out. The Heckman selection method did not perform well in the simulation study. Multiple imputation was the best method of the three considered for data missing not at random.

The impact of SEP on cognitive function varied by cohort, as well as SEP and cognitive measures, with father's occupational SEP, but not childhood household amenities, associated with crystallized cognitive function in the NSHD after adjustment for later life SEP. Accumulation models were usually supported when considering the life course hypotheses. In some analyses the conclusions varied depending on the missing data methodology utilized.

Overall, there was no consistent conclusion as to whether childhood SEP remained a significant predictor of cognitive function in adulthood, but it was not a significant predictor of cognitive decline in Whitehall II after adjustment for later life SEP. Multiple imputation was found to be an appropriate method of dealing with missing data in most situations.

Acknowledgements

Thank you to my primary supervisor, Dr Rebecca Hardy, who has provided guidance, encouragement and valuable feedback over the course of my PhD. Thank you also to Jenny Head, for her advice on the statistical aspects of this thesis, and to Dr Marcus Richards for his advice in developing the epidemiological aspects.

For her assistance with developing and writing the code for the simulation study I thank Amy Kirkwood, and for generously sharing her Stata expertise over the last few years (long may it continue!). The simulations would not have been possible without the numerous kind people who let me use their computers; I am very grateful to Steven Bell, Rebecca Lacey, Ellen Schwab, Jitka Pikhartova, Gabriel Masset, Anja Scheiwe, Lizzy Leigh, Noriko Cable, Jenny Head, Kim Bouillon, Caroline Coope and everyone in the Whitehall II study room.

Thanks to Sean Clouston for his assistance with understanding Heckman selection models. I am lucky to have been welcomed by three research groups throughout my PhD, and I am grateful to the members of the MRC Unit for Lifelong Health and Ageing, Whitehall II and the International Centre for Lifecourse Studies for all the support, encouragement and opportunities that they provided throughout my PhD.

For keeping me sane and providing encouragement throughout my PhD, I would like to thank my PhD colleagues, family and friends, especially Lizzy Leigh, Benjy Gilbert and Verd Dost for proofreading large sections of this thesis, and Jon Roiser for putting up with all my questions and his willingness to help. Finally, I would like to acknowledge the ESRC for funding my PhD studentship.

Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 The importance of studying cognitive function

Cognitive function has been associated with many physical (1) and mental (2) health conditions, as well as mortality (3;4). A systematic review (5) found that higher cognitive test scores in childhood and youth were related to lower rates of all-cause mortality in middle to late adulthood. Decline in cognitive function has also been associated with higher mortality risk (6). In 2009 a special issue of the journal Intelligence focussed on cognitive epidemiology, with a discussion article commenting that "social scientists and practitioners cannot afford to neglect cognitive ability when modelling epidemiological and health care phenomena" (7). Similarly, Singh-Manoux recently commented that "impaired cognitive status is one of the biggest challenges of the future due to its impact on both the individual and society" (8). The cognitive trajectory of individuals with preclinical Alzheimer's disease has been studied. It is important to understand the differences between these trajectories and those of 'normal' cognitive ageing (9), in order to identify individuals on a preclinical Alzheimer's disease cognitive trajectory, thus raising possibilities for earlier interventions to take place.

Understanding the risk factors which affect cognitive ability is therefore of importance, as such factors could represent the underlying causes of any association between cognitive ability and health, as well as the impact of cognitive decline itself on individuals and their families. Earles and Salthouse found that self-rated health accounted for 15-20% of age-related variance in cognitive scores, implying that health factors, although explaining some, do not explain the majority of the cognitive decline observed in ageing (10). One factor which has been consistently strongly related to cognitive ability is socioeconomic position (SEP). It has been suggested that an individual's SEP background, including childhood SEP, may impact structural and functional brain development (11), and that childhood mental test scores may "act as a record of insults to the brain" that have occurred up to that point (12).

There is a complex relationship between SEP, lifetime intelligence and later health, which is not yet fully understood; even the relationship of later health to intelligence, education and SEP has not been agreed upon (12). It may be that social circumstances confound the relationship between early life cognition and adult mortality (13). The systematic review by Calvin et al (5) found that childhood SEP had no impact on the relationship between childhood cognitive function and mortality, whereas education and adult SEP attenuated this effect. In the Scottish Mental Survey childhood cognition was found to be related to all-cause mortality in adulthood (14), with those who had higher test scores as children at lower risk, independent of childhood SEP.

Socioeconomic inequalities in health have been observed throughout a range of time periods and locations; Gottfredson (15) proposed that intelligence is the "fundamental cause" of these socioeconomic inequalities in health. Singh-Manoux et al (16) examined the extent to which cognitive ability explained such inequalities using Whitehall II data. They concluded that intelligence was unlikely to be the fundamental cause of social inequalities in health as variation in cognitive ability does not fully explain the relationship between SEP and health. Similarly, Batty et al (17) concluded that scores from an IQ test did not entirely explain the socioeconomic gradients in health, but that controlling for IQ did reduce the magnitude of the gradients. In the Maastricht Aging Study (18), the greatest cognitive decline was observed in the lowest SEP group, and more than a third of the association between adult occupational level and longitudinal change in cognitive function was explained by the lower baseline intellectual abilities in lower occupational groups.

It is clear that the relationship between SEP and cognition through the life course has implications for later health. This thesis will therefore examine how cognition varies by SEP across the life course in two cohort studies. To carry out this examination it is necessary to have data which has been collected throughout subjects' lives to investigate the inter-relationships between SEP and cognition from childhood through to adulthood.

One of the major limitations of such longitudinal studies is missing data and study dropout. Many studies find that individuals with lower cognitive function and lower SEP are more likely to drop out of the study. Carrying out analyses which ignore the impact of missing data could result in biased estimates of the associations between SEP and cognitive function and may lead to misleading or erroneous conclusions. Numerous simulation studies have shown that results can be biased unless methods are applied to allow for the effects of missing data. This thesis thus also addresses the issue of missing data when investigating the relationship between SEP and cognitive function, in order to obtain results that are accurate and unbiased. It will do this by comparing methods used to account for missing data, assessing their impact on findings and thereby establishing the most appropriate method to apply to such examples to limit this impact.

1.2 Literature review

The first topic to be reviewed is the trajectory that cognitive function takes throughout the life course. After this, the literature on the relationship between SEP and cognitive function, both at one point in time and measured throughout the life course, is reviewed. This section will also refer to the methods used for addressing the issue of missing data in the studies reviewed. Measures of cognitive function are distinguished from measures of cognitive impairment, such as the Mini Mental State Examination (MMSE), which is used to diagnose dementia. Unless specifically mentioned, studies of cognitive impairment have not been included in this literature review.

1.2.1 Cognitive function throughout the life course

In 1996 Neisser et al (19) published an authoritative review of ‘intelligence’ on behalf of the American Psychological Association. Though the word intelligence can cover a wide domain, the concept of intelligence was explained in this review as the “ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought’ (19).

Cognitive function covers a range of domains, such as processing speed, working memory and verbal comprehension (20). Two aspects of cognitive function are fluid intelligence, which is considered a measure of the ability to process information, and crystallized intelligence, which is thought of as knowledge accumulated over time (20). It is well established that cognitive function changes throughout the life course. During early childhood, cognitive function develops rapidly, for example when learning to talk. Fluid intelligence reaches a peak and then declines, whereas crystallized intelligence, such as vocabulary, remains at a similar level into old age, or can even improve during ageing (Figure 1.1, adapted from (21)). The absolute level of fluid intelligence at any point in later life therefore depends both on the peak level achieved and the rate of decline. A meta-analysis of cross sectional studies found that cognitive speed decreases by around 20% at age 40 and 40-60% at age 80 (22). However the mean National Adult Reading Test (NART) scores, a measure of the ability to read irregular words, were no different in people with and without dementia, after controlling for childhood IQ scores, in one study (23).

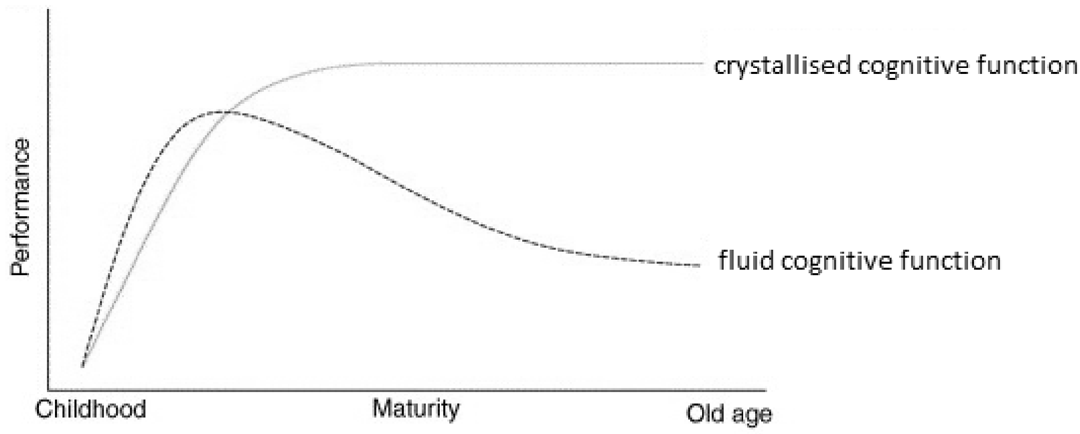


Figure 1.1: Cognitive function throughout the life course

There are many advantages to using longitudinal studies to investigate cognitive decline. Longitudinal studies consist of repeated measures of the same individuals, which enable change over time to be studied within individuals. Such studies are almost always prospective. Longitudinal studies allow ageing effects (within-subject changes) to be distinguished from cohort effects (between-subject changes) when the participants cover a range of ages. For birth cohorts with a narrow age range, it is not possible to separate cohort effects from ageing effects unless such cohorts are compared against each other. Longitudinal studies can provide estimates of rates of decline for individuals, as well as risk factors for cognitive decline; however they may underestimate change due to practice effects and selective dropout (22).

Wilson et al (24) measured seven cognitive abilities annually in older Catholic clergy members. They found that decline occurred for every ability, with the lowest decline observed in word knowledge, a measure of crystallized function. They also found that there were wide differences between individuals at all ages. The observed rate of change in a given domain was not related to the baseline level in that domain, but the rates of change in different domains were moderately associated, leading to the conclusion that the cognitive trajectory in old age is mainly the result of person-specific factors, rather than an “inevitable developmental process”. However a study of elite academics and elderly blue-collar workers suggested that cognitive decline was universal in non-verbal intelligence tests (25). In longitudinal studies there is both within- and between-individual variability; as some people experience very little cognitive change with ageing while others experience much greater levels of decline, between-individual variability increases with age (22). Rabbitt (26) concluded that cognitive skills do not "all go together when they go", with the within-individual variance increasing with age as well as the between-individual variance.

One study of cognitive decline which has used appropriate missing data methodology is that of Muniz Terrera (27), which investigated the shape of cognitive decline by comparing the model fit of a linear model, a quadratic model, and two piecewise models, using longitudinal data on the MMSE over 9 years. The MMSE is a test which is often used as a screen for possible dementia. When the raw MMSE scores were examined, the scores increased between the second, third and fourth waves of data collection. This could be due to the selective dropout observed; at baseline the mean age was 81. The analyses therefore used a selection model framework (see section 2.4.2). The model providing the best fit to the data was a positive quadratic, with the rate of decline decreasing with increasing age. The main consequence of the attrition bias was less observed decline when a complete case analysis was carried out.

A difference in results has been observed between cross-sectional and longitudinal data, as can be seen in the plots below (Figure 1.2), taken from Schaie's work using the Seattle Longitudinal Study. The Seattle Longitudinal Study has a cross-sequential design, where new participants were recruited at each study wave, as well as the follow up of participants already in the study. Schaie concluded that this confirmed the presence of cohort effects, with scores generally increasing the more recently the cohorts were born (28). This may be an example of the Flynn effect, which refers to a secular increase in intelligence test scores, observed since testing began. Cross-sectional results have suggested that the rate of cognitive decline, as measured by the MMSE, is faster in those with lower NART scores (29).

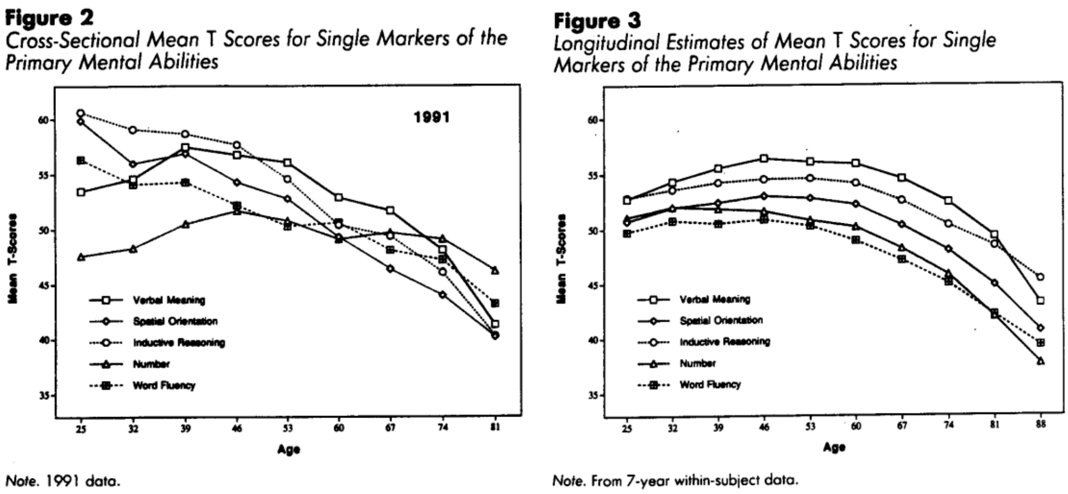


Figure 1.2: Cross sectional (left) and longitudinal (right) trajectories of mental abilities (Schaie)

The longest time-frame over which cognitive data are currently available comes from the Scottish Mental Survey of 1932 (30). This is a historical study; the Scottish Mental Surveys took place in 1932 and 1947, when all pupils aged 11 in Scotland took the Moray House Test, a verbal ability test, with subsamples given other cognitive tests. Beginning in 1997 survivors from the samples who took these tests were recruited to a follow-up study. The participants took a series of cognitive tests, including a repeat of the Moray House Test that they had taken at age 11 and the Raven's Progressive Matrices. The correlation of the Moray House Test score at age 11 and Raven's Progressive Matrices score in the follow-up at age 65 was 0.58 (using the 1947 Scottish Mental Surveys), higher than the correlation between ages 11 and 80 (0.45, using the 1932 Scottish Mental Surveys). This may suggest that the correlation between childhood cognition and adult cognition weakens with increasing age, or it may represent a cohort effect. Despite being the longest follow-up study of the stability of mental ability, data were available at only two periods of life, age 11 and old age. This study cannot therefore study the whole life course cognitive trajectory, or capture the peak.

Further, a study of this type is not representative of the population born in 1921 or 1936. This is because it is limited to those who not only survive to old age, but who are also still healthy and living independently (a condition of being in the follow-up part of the study), and choose to participate in the study. It has been shown in the Scottish Mental Surveys that lower IQ in childhood was related to higher mortality rates, thus indicating that the follow-up sample was not cognitively representative of the original cohort (14), yet no adjustment was made for missing data in the Scottish Mental Surveys analyses (14).

The question of whether men and women have the same patterns of cognitive change has been debated. Studies have generally found that women suffer greater cognitive decline and cognitive impairment than men (30;31). However this may be due, in part, to women's increased life expectancy.

1.2.2 Earlier cognitive function and cognitive decline

Before investigating the influence of SEP on cognitive function later in life, it is important to understand the shape of the cognitive trajectory throughout the life course, and to understand how cognitive function earlier in life is associated with both cognitive function and cognitive decline later in life.

Several studies have investigated whether early cognitive function is related to rate of cognitive decline in older age, and the results have been somewhat conflicting. Some have shown that the rate of decline is slower in those with higher early life cognition (32;33), while others have shown no association (34;35), and one study has shown a very weak positive association (36) (see Appendix 1).

One study (36) which included individuals with a wide range of ages, from 49 to 92, and in which youthful cognitive scores were predicted from adult Mill Hill test scores, showed a very weak positive association, with slower decline among those participants with lower 'childhood' cognitive scores. However, the 'childhood' cognitive scores were represented by adult vocabulary test scores, which may not be precise enough for this situation (32).

Gow et al (34) used two different methods to analyse the same data in a longitudinal study and observed different results. When linear regression was used (the method of analysis in three other studies (32;33;36)) a negative association was found between childhood cognitive function and cognitive decline. In contrast, no association was found when latent growth curve modelling was used (34;35), although both these studies used the same dataset. One possible contributing factor to the different results is the different ways in which missing data were treated; the conditional change linear regression analyses were all complete case analyses, which assume MCAR missingness, whereas full-information maximum likelihood (FIML) was used with the latent growth curve models, which allows for MAR missingness. Gow et al concluded that the results from the latent growth curve model were more accurate since linear regression cannot completely account for test-specific variance.

1.2.3 Childhood SEP and childhood cognitive function

Many studies have examined SEP at one point in time and its relationship to cognitive function either at the same time or later. This thesis will first consider studies which have investigated whether cognitive function during childhood varies by childhood SEP. Childhood SEP is usually measured using characteristics of the child's parents.

Overall studies show that higher IQ is related to higher SEP during childhood from ages as young as 9 months.

Bradley & Corwyn (37) reviewed the literature on the relationship between SEP and intellectual ability/academic competence in early life. They concluded that poverty and low parental education were associated with lower school achievement and IQ later in childhood. They also considered various SEP measures in relation to cognitive development, and found that each SEP measure used in the Health Examination Survey (family income, paternal and maternal education) was strongly associated with intellectual attainment, with maternal education being the best predictor (38).

White (39) conducted a literature review and meta-analysis; by comparing multiple cross-sectional studies he found some evidence that the association between SEP and cognitive function decreased with age during childhood. White also concluded that the correlation between SEP and cognitive function had declined over time, with earlier studies finding higher correlations. Two explanations were offered; one, that the availability of community organizations and preschool had increased for people from all SEP levels, and two, that the reduced strength of the correlation was a result of the successful compensatory education. A more recent meta-analysis (40), using White's procedures, found a slight decrease in the association since White's review, although they did not investigate whether the relationship had changed over time in the studies included in the meta-analysis.

The association between SEP and cognitive function has been found in a variety of countries and age-groups (41). For example, a large study of children born in 1963 living in Warsaw, Poland, where dwellings, schools and health facilities were allocated without taking account of social class, found that mental performance was still related to parental occupation and education (42).

Although most existing studies are cross-sectional, some longitudinal studies have investigated how the effect of childhood SEP on cognitive function changes with age. In contrast to the White review, some studies observed an increasing gradient with age (40;43;44), and some a similar effect at all ages (45). One study found that the relationship of SEP with cognitive function increased between ages 2 to 7, but then decreased again at age 12 (46).

Jefferis (44) made use of repeated cognitive data in the 1958 British birth cohort to examine how SEP at birth affected cognitive development and cognitive function at age 33. Maths scores at ages 7, 11 and 16 were converted to z scores (mean of 0 and standard deviation of 1), and an increasing z score with age was viewed as an improvement in relative achievement. A repeated measures multilevel model was used for the longitudinal analysis, to take account of the clustering of maths scores at different ages within individuals. Social background was found to have a large effect on maths scores, with the percentage of variance in these scores explained by social class at birth increasing from around 3% at age 7 to 12% at age 16. The difference between the mean z scores for social classes I and II compared to IV and V increased with age from 7 to 16, when there was a full standard deviation advantage for children born into social classes I or II. Li (47) extended this study by categorising reading scores, and found that SEP factors were strongly associated with both reading and mathematics scores.

Feinstein (48) used the 1970 birth cohort to examine the effect of SEP at birth on cognitive trajectory up to age 10, finding that SEP is much more influential than earlier cognitive ranking on cognitive rank at age 10, with those from a low SEP with a high initial cognitive ranking dropping below those from a high SEP with a low initial cognitive ranking by age 10.

Overall, evidence suggests that childhood SEP does affect the absolute level of childhood cognitive function, however it is less clear whether the strength of the association changes with age. Lawlor et al. (45) used multiple imputation to account for missing data in risk factors and confounders, and Li (47) used a repeated measures multilevel model, which uses all the available cognitive outcome data, but requires complete risk factor information. However the majority of the papers carried out complete case analyses.

1.2.4 Adult SEP and adult cognitive function

Many studies have found a relationship between higher adult SEP, using a variety of measures, and better cognitive function (see Appendix 2). Although there are researchers who do not consider education to be a measure of SEP, especially when investigating the relationship between SEP and cognitive function, the relationship between education and cognitive function later in life has been widely investigated.

Education was found to be significantly associated with a range of cognitive measures, including memory (49-56) and crystallized cognitive function (53;55;57;58), as well as other cognitive tests (31;49;51-56;58;59). Occupation was associated with memory (49;50;53;57;60), crystallized cognitive function (36;53;57;58) and other cognitive tests (36;49;53;57;58;60). Only one study found that occupation was not associated with cognitive function; in a prospective cohort study Bosma et al (18) found no association between occupation and cognitive functioning at baseline. However, the measure of cognitive function used was "bother due to forgetfulness in daily life", and the age range of participants was 24-81. Fritsch (54) found that occupational demands of the longest held job, derived from measures of mental, physical and social traits associated with various occupations, was predictive of cognitive function.

Cagney and Lauderdale (51) investigated the effects of wealth and household income, as well as the effect of education, on cognitive function (including memory, knowledge, language and orientation), and found that income and net worth had a much smaller effect than education.

1.2.5 Adult SEP and adult cognitive decline

The majority of papers examining the relationship between adult SEP and cognitive decline (see Appendix 3) used education as the measure of SEP. The results are less consistent for cognitive decline than for cognitive function at one point in time. Anstey and Christensen (61) carried out a review of 14 papers investigating education as a predictor of cognitive decline. The results were contradictory; five of the studies concluded that the rate of decline was slower for the more educated, four studies found that this effect was restricted to a subgroup; five found that the association was restricted to types of outcome measures, and two found no association. Anstey and Christensen concluded that the results seemed to depend on the cognitive test used; all seven studies which used mental status studies, often the MMSE, found that education had a positive association, whereas neither study that used fluid measures found this.

More recently Valenzuela and Sachdev (62) reviewed the literature on brain reserve and cognitive decline, part of which involved examining the effects of education and occupation on longitudinal cognitive decline. Of the thirteen studies reviewed which examined the effect of education, ten found a significant effect, with a large positive effect overall, whereas of the only four papers which considered the effect of occupation, three showed small positive effects, but the overall effect was non-significant.

The debate is still continuing; some studies have concluded that SEP is inversely associated with rate of cognitive decline (63-65), whereas others concluded that there was no association. Some research has found mixed results, for example Leibovici et al (66) concluded that the result depended on the nature of the cognitive test: for tests with a high learned component the highly educated declined less, but there was little difference in those tests with a higher ‘nature’ than ‘nurture’ component, such as attention and visuo-spatial analysis. Ardila (67) also found that the results depended on the cognitive measures used, with the more educated group declining slower for word recall, but faster in the verbal fluency test. Schmand (68) found that lower levels of SEP were associated with a higher rate of cognitive decline among the older participants only. Singh-Manoux (69) found that the results depended on the SEP variable used; education did not affect the rate of cognitive decline, but occupation did, with participants in the high occupation group showing greater cognitive decline.

Jorm (57) found no significant differences between groups in any of a variety of cognitive tests. However, Jorm grouped occupations according to the Australian Standard Classification of Occupations, which divides these into artistic, conventional, enterprising, investigative, realistic and social, which is quite different to most groupings of occupations. Lee (52) defined cognitive decline as a binary outcome (decline vs. no decline), whereas the other studies discussed in this section investigated the rate of decline.

Dugravot (70) investigated the effect of SEP on change in cognitive function using two different methods; the first used ANOVA to estimate the effect of SEP on the change score, and the second used ANCOVA to estimate the effect of SEP on the change score, adjusting for baseline cognitive score, a method known as a conditional model of change. Different results were found in the two situations; no effect of SEP was found when ANOVA was used, but the low SEP group was found to have a higher rate of cognitive decline with ANCOVA. This discrepancy is known as Lord’s paradox (71); however Plewis (72) explains that rather than being a paradox, the two approaches address different questions. The ANCOVA method would be appropriate if the baseline cognitive scores were the same in the different SEP groups. As this is not the case, the ANOVA results are more reliable in this situation.

1.2.6 Childhood SEP and adult cognitive function

It is of interest to know whether childhood SEP has longer lasting effects on cognition, as social inequalities in cognition emerge during childhood, potentially suggesting underlying mechanisms, and identifying the age at which interventions to reduce inequalities might be most effective. This section discusses studies which investigate the influence of childhood SEP on adult cognitive function. All of the studies have SEP measured in childhood and in adulthood, with most studies having two adulthood measures, usually including education. Studies which specifically investigated the life course hypotheses are considered in a separate section (see section 1.2.8). For more details on the papers discussed in this section see Appendix 4.

The aim of reviewing the work which does not formally test any of the life course hypotheses is to investigate whether childhood SEP remains a significant predictor of adult cognitive function after adjusting for adult SEP. Eleven studies were identified which examined the relationship between childhood SEP and adult cognitive function. In analyses which were not adjusted for later life SEP, almost all of the studies found a positive association between childhood SEP and cognitive function later in life (52;73-80). Kaplan (73) reached different conclusions for the different measures of childhood SEP used; for both father's education and mother's occupation, no association with adult cognitive function was found in the unadjusted analyses, although a significant effect was found for mother's education and father's occupation.

When childhood SEP was adjusted for later life SEP some studies found that childhood SEP remained a significant predictor of adult cognitive function (75-79;81), whereas other studies found the effect of childhood SEP was fully attenuated (52;82;83). Kaplan again reached different conclusions for different SEP measures, with the effect of father's occupation fully attenuated, but an effect remaining for mother's education (73).

Singh-Manoux (82) used structural equation modelling to compare two models; a 'direct effects' model which estimated the effect of each measure of SEP independently of the other measures of SEP, and an 'indirect effects' model, where the effects of early life SEP were mediated through measures from later life. They found that the indirect effects model was a better fit, implying that the effect of childhood SEP was mediated through education and adult SEP, with no direct effect found between childhood SEP and cognition.

Richards and Sacker (81) concluded that although the direct effect of childhood SEP on crystallized cognitive function was negligible, it had a substantial effect through its impact on cognitive development; this was especially true for the fluid cognitive measures used where there was no direct effect. However Richards and Sacker adjusted for childhood cognitive function in their analyses, a variable not available in most datasets. The other study which adjusted for childhood cognitive function (83) found that the effect of childhood SEP was fully attenuated. The other two studies which found the effect of childhood SEP was fully attenuated both used very specific samples, the nurses' health study (52) and the Whitehall II study of British civil servants (82), both of which are restricted to participants with higher levels of education than a general population sample. Some of the studies used regression methods (52;73;75;78;80), whereas others used structural equation modelling (81-83); however results were inconsistent within each methodology.

Most of these studies (52;73;75;76;78-80;82;83) relied on retrospective childhood SEP data, which is vulnerable to recall bias. The reliability of the recalled data may also be influenced by the level of an individual's cognitive function when the information was recalled. Complete case analyses were carried out in three of the studies (52;78;80), while some of the studies using multilevel models required participation in a certain number of follow-ups (75;76;79), which does not make full use of all the available data. The two studies using SEM used FIML, which assumes the missingness is MAR (81;82).

The studies can be split into two groups; one group adjusted for very few variables beyond SEP (73;75;76;79;81-83), whereas the other group adjusted for a wide range of potential confounders (52;78;80). It is feasible that some of the covariates considered, such as alcohol consumption, could be on the causal pathway between SEP and cognitive function (84).

1.2.7 Childhood SEP and adult cognitive decline

Four studies have been identified that examine childhood SEP and cognitive decline in older age (see Appendix 5). Three of the four studies found no association (75;79;85); however Lee et al (52) found a relationship between father's occupational SEP and cognitive decline, with higher odds of experiencing cognitive decline for children of farmers than upper white-collar workers. The difference may be due to the method of measuring cognitive decline, as Lee et al compared the odds ratios of experiencing the worst 10% of change in cognitive score, whereas the other studies used the change in score (79;85) or whether the trajectory differed among different SEP groups (75).

As in section 1.2.6, most of these studies (52;75;79) relied on retrospective childhood SEP data. Complete case analyses were carried out in two of the studies (52;85) and the studies using multilevel models required participation in a certain number of follow-ups (75;79).

1.2.8 Life course SEP and cognitive function

Few studies have considered how life course SEP influences adult cognitive function and decline. In this section, first the life course methods and models will be introduced, and then the literature on life course SEP and cognitive function reviewed. The studies reviewed investigated how individual SEP at various points in the life course was related to cognitive function, and some also investigate the impact of socioeconomic mobility on cognitive function.

1.2.8.1 Life course methods

There are two general life course hypotheses as defined by Kuh and Ben-Shlomo. The first hypothesis covers critical and sensitive periods; a critical period refers to "a limited time window in which an exposure can have adverse or protective effects on development and subsequent disease outcome" (86), with a sensitive period allowing for smaller effects outside the window. In relation to SEP, a sensitive period is more likely to be relevant than a critical period, since SEP at all stages in the life course are likely to have some effect. The second hypothesis is the accumulation hypothesis, where exposures accumulate to increase the risk of an outcome. The accumulation hypothesis is considered by some to be the 'main explanation to observed socio-economic differences in risk of disease' (87).

Studies with information covering all stages of the life course provide data to test the many different proposed life course models, such as the four accumulation of risk models pictured below (Figure 1.3, adapted from (86)). Model (a) shows an accumulation of independent risk factors, and Model (b) shows an accumulation of clustered exposures. Models (c) and (d) are 'chains of risk' models; in Model (c) each exposure both increases the risk of the subsequent exposure, and also has an independent effect on risk of the outcome, separately from the later exposure. In Model (d) the earlier exposures only have an effect if the final exposure occurs, known as a 'trigger effect' (86).

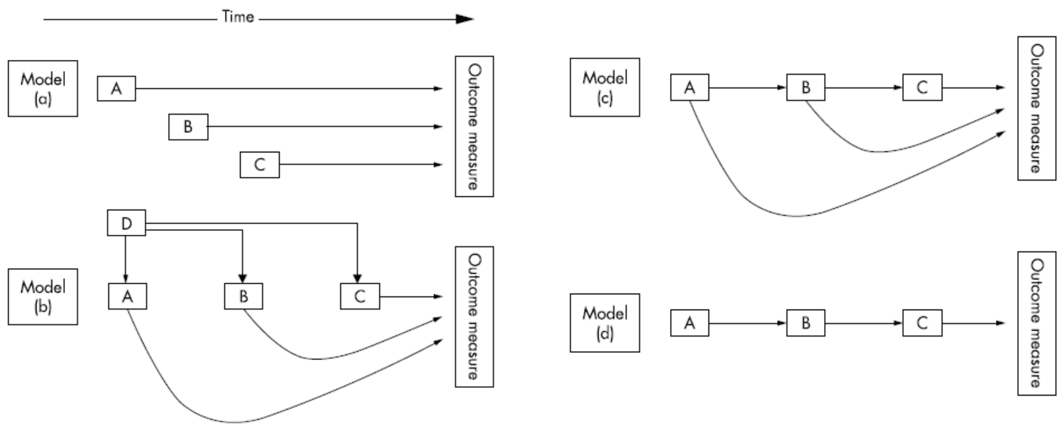


Figure 1.3: Life course models

There is also interest in the impact of social mobility, especially in the social sciences, but there is debate as to whether this represents a different life course model. When investigating social mobility between two time points, the analyses have usually compared those participants in the higher SEP category at both time points to those who were in the lower SEP category at both time points and those who were socially mobile, either combined into one group or in two groups, upwardly mobile and downwardly mobile. However this is the same analysis as would be carried out to test the accumulation hypothesis. Hence, using this definition social mobility cannot be distinguished from accumulation.

1.2.8.2 Life course SEP

As Lynch (88) pointed out, the 'most important' limitation of SEP measures is that they are normally measured at one point in time, whereas variations between groups 'are generated over the entire life course'. Hallqvist (87) proposed that the three life course hypotheses - accumulation, critical period and social mobility - could not be 'disentangled' to determine which model was most appropriate, as the data could support more than one of the hypotheses. This is due to the eight possible trajectories which can be formed from binary SEP variables at three stages in the life course. As shown in Table 1.1 from Hallqvist (87), most of the trajectories are associated with more than one exposure category defined by the life course hypotheses, and it was therefore concluded that there was ‘no way of creating an unconfounded exposure contrast of any of them’.

How each of the eight trajectories measures different exposure categories of the three competing hypotheses: accumulated experience of manual socio-economic position (0, 1, 2, and 3 periods), critical period (either 1st, 2nd, or 3rd), and social mobility (intergenerational and intergenerational upward and downward)

Trajectory[a]	Accumulated dose				Critical period			Social mobility[b]			
								Inter-		Intra-	
	0	1	2	3	1	2	3	↗	↘	↗	↘
	X										
		X					X		X		X
		X				X			X[c]	X	
			X			X	X		X		
		X			X			X			
			X		X		X	X[c]			X
			X		X	X		X		X	
				X	X	X	X				

[a]The graphic design of the trajectories refer to Fig. 1. The trajectory on the first line depicts those who have been non-manuals in all three periods, on the second line those who where non-manuals in childhood and at 25 but later became manual workers, and so on.
[b]The stable (referent) categories are defined in Table 7.
[c]Depends on the choice of age for the final position.

Table 1.1: Life course SEP trajectories (Hallqvist et al)

Mishra et al (89) proposed a method of modelling repeated binary exposure variables to compare life course models, using SEP (manual, non-manual) as an example. The method distinguished accumulation, critical period and certain specifications of a social mobility model. The accumulation model was tested by summing indicators of the binary SEP variable over the life course, and tested whether the total time spent in disadvantaged conditions was the likely model. The critical period model investigated whether SEP at a specific point in time was associated with the outcome, irrespective of SEP at other points in the life course. The social mobility model has not been well defined in the literature and debate remains over its specification. Mishra et al (89) defined two types of social mobility based on different trajectories over the life course, using a different definition to Hallqvist. The methodology describes a way of deciding which life course hypothesis best fits the data by comparing a model representing each of the life course hypotheses to a saturated model. More than one of the life course models could be accepted as being not significantly different to the saturated model, in which case the results of the life course models could be compared. This is one advantage of the methodology; when only one life course model is tested it may be supported by the data, however had other life course models also been tested they too may have been supported by the data. Mishra et al's methodology (89) is described in more detail in section 6.3.

1.2.8.3 Review of the literature on life course SEP and cognitive function

The results of studies that have examined life course SEP and cognitive function are now summarized. All studies have SEP measured in childhood and in adulthood, with most studies having two adulthood measures, usually with education as the measure for early adulthood. All of these studies have tested a specific life course model (accumulation, sensitive period, social mobility). For more details on these studies see Appendix 6. As in the earlier sections, studies of cognitive impairment were excluded unless specifically mentioned.

Two studies assessed the effect of cumulative SEP on cognitive function (77;90), where cumulative SEP was defined by summing a series of binary SEP variables. Both of these studies observed a dose-response relationship, with more time spent in a socially disadvantaged environment associated with lower cognitive function. Turrell et al (77) carried out the analyses adjusting only for age, and additionally adjusting for a range of morbidity indicators, and found no difference in the effect of SEP on cognitive function.

Socioeconomic mobility was investigated in three studies (55;77;90). The overall conclusion was that participants who were upwardly mobile had higher cognitive scores than those who had a steady low SEP, and that participants who were downwardly mobile had lower cognitive scores than those who had a steady high SEP. Luo & Waite (90) looked in more detail at those who had a mixed mobility pattern, and found this group to score between those who experienced upward or downward mobility. Hatch et al (55) investigated inter-generational social mobility, whereas the other studies used three time points; in childhood, early adulthood and later adulthood. Very few variables were included in any of the analyses beyond the SEP variables; Luo & Waite adjusted for age, sex and race/ethnicity, and Hatch et al (55) adjusted for childhood cognitive function. None of the studies specified that they were investigating a critical period hypothesis.

1.2.8.4 Review of the literature on life course SEP and cognitive decline

The only study which investigated any of the life course hypotheses with regards to cognitive decline is that of Long et al (91). Their study examined the accumulation hypothesis, with the outcome of MMSE, a measure of mental status, rather than a measure of cognitive function. The outcome was defined as decline of 3 or more MMSE points, with the cumulative SEP variable composed of three binary variables; having eighth grade or less education, having an annual income of $7,000 or less, and having either never worked or worked in a job that required only minimal skills. There was a dose-response relationship between cumulative disadvantage and risk of mental status decline. After adjusting for baseline MMSE score, participants with a cumulative disadvantage score of 3 had an 85% (95% CI: 17% - 165%) increased risk of experiencing MMSE decline of 3 or more points compared to those who had a cumulative disadvantage score of 0.

1.2.8.5 Methodological issues with life course SEP and cognitive function studies

Almost all of the studies began in midlife and are thus not representative of the socioeconomic structure that existed at the birth of participants, as those who survive to older ages are not representative of those who are born; lower SEP is consistently shown to be associated with higher premature mortality rates (92). These studies also rely on retrospective measures of SEP. Using data from the Aberdeen children of the 1950s study Batty (93) concluded that the agreement between social class of father recalled in adulthood and that recorded at birth and in childhood was only moderate. Where reported, attrition rates were generally higher in the most disadvantaged groups. Most studies carried out complete case analyses which are valid when data are missing completely at random (MCAR). Two of the studies (81;82) used FIML, which has only been shown to be appropriate under the missing data mechanisms of MCAR and missing at random (MAR) (see section 2.2) (94). These assumptions may not hold when considering SEP and cognitive function, as people with lower SEP and lower cognitive function are more likely to drop out in longitudinal studies (32;95), and it is likely that those with faster cognitive decline would be more likely to drop out. This implies that there is selection into the data analysed, related to both SEP and cognitive function, meaning that the missing data mechanism may be missing not at random (MNAR). The implications of missing data are discussed in Chapter 2.

1.2.9 Measures of SEP

SEP affects a wide variety of health outcomes, from ill health throughout life (96) to mortality (97), although there is no consensus on exactly what SEP represents (98).

Chapin defined SEP as '*the position that an individual or family occupies with reference to the prevailing average of standards of cultural possessions, effective income, material possessions, and participation in group activity in the community*' (99); whereas Krieger defined SEP to be "*an aggregate concept that includes both resource-based and prestige-based measures, as linked to both childhood and adult social class position*" (100).

The three most common indicators of SEP used in industrialized countries are income, education and occupation (101). The Registrar-General's class schema, based on occupation, is usually regarded as a hierarchy. Education can be measured in a number of ways; for example educational qualifications or years of completed schooling (100). As mentioned above, education could be perceived as representing cognition rather than SEP; however it is strongly determined by parental characteristics, and has been considered an appropriate measure of SEP within a life course framework (102). Wealth and income may play distinct roles with respect to material conditions; income is specific to occupation, whereas wealth is a broader measure of financial resources and safety nets (103). There is a general consensus that income, education and occupation together represent SEP better than any of these alone (104). However Bradley and Corwyn (37) conclude that the choice of how to measure SEP remains open, and depends on the question being asked. Macintyre et al (105) showed that the relationship between SEP and health varied according to both the SEP measure and health measure used.

A range of measures have been used to represent childhood SEP. Parental occupation and education were the most common in the papers reviewed above, but a variety of household variables have also been used, such as overcrowding and other material conditions. Neighbourhood and community-level variables have also been used to measure SEP, such as area-based levels of unemployment. However when considering the effect of area-level SEP, it is important to note that the effect will differ depending on the individual's own SEP (59).

It is also important to consider how to measure SEP in older people, as occupation or income may no longer be appropriate after an individual has retired. Also when considering a range of ages, the same level of education may represent something different to the older and younger participants, based on differing childhood education opportunities and norms. O'Reilly (106) investigated the relationship between three measures of deprivation and levels of income support, and found that the relationship with mortality varied between the different SEP measures.

Different measures are appropriate in different countries, and the measures which are relevant have changed over time; for example access to an indoor toilet is no longer a good way to distinguish between different SEP levels in the UK, although it may still be in some countries.

There has been some discussion as to SEP measures when considering a cognitive outcome, and specifically how each SEP variable may contribute to cognitive function or cognitive decline. One such paper is by Glymour and Manly (107). The four main pathways suggested are material conditions, psychological stressors, cognitive engagement and test taking skills. Material conditions and psychological stressors can then lead to differences in medical access, physical health and health behaviours, which may in turn impact on cognitive function and cognitive decline. There may also be a genetic contribution, both to cognitive ability and age-related cognitive decline (108). Glymour and Manly point out that individuals with higher levels of education or income are likely to be treated with more 'esteem and deference' in their daily life, which may allow them to avoid other stressors, which 'likely has substantial consequences for cognitive aging'. It is also important to be aware that various SEP factors 'are not interchangeable with respect to cognitive function' (51), and Gallacher et al (58) commented that although social class and education were closely related, each still contributed substantially to cognitive function.

1.3 Aims and objectives

There are two main purposes of this thesis, which aim to fill some of the gaps in the current knowledge of the relationship between SEP and cognitive function and cognitive decline.

Aim 1:

The first aim is to investigate the effect of SEP on crystallized cognitive function in late middle age, suitably accounting for missing data. This work is split into two sections; first to investigate whether any effect of childhood SEP on crystallized cognitive function remains after adjusting for later life SEP, and second to investigate the life course hypotheses, accumulation, critical period and social mobility, on crystallized cognitive function.

Objective 1:

i. To investigate the relationship between childhood SEP and adult crystallized cognitive function.

ii. To investigate whether there remains an effect of childhood SEP on crystallized cognitive function after adjusting for later life SEP, and childhood cognitive function.

iii. To investigate the impact of applying missing data techniques to the analyses for objectives i to ii.

Objective 2:

To carry out a simulation study, allowing an in-depth examination of the performance of the missing data techniques under each of the three missing data mechanisms.

Objective 3:

i. To investigate which life course hypotheses were supported using standard techniques.

ii. To investigate which life course hypotheses were supported, using and extending the life course methodology developed by Mishra et al (89).

iii. To investigate the impact of applying missing data techniques to the analyses for objectives i and ii.

Aim 2:

The second aim is to investigate the effect of SEP on the trajectory of fluid cognitive function, suitably accounting for missing data. This work starts by investigating the impact of childhood SEP on fluid cognitive trajectory, and whether an effect remains after adjusting for later life SEP. The life course hypotheses are also investigated, to examine the impact of the different life course SEP trajectories on the cognitive trajectory.

Objectives:

i. To investigate the effect of SEP at each stage of the life course on the intercept of memory trajectories.
ii. To investigate whether SEP at each stage of the life course influences the rate of memory decline.
iii. To investigate whether there remains an effect of childhood SEP on the intercept and slope of a memory trajectory after adjusting for later life SEP.
iv. To investigate which of the life course hypotheses explain the intercept and slope of memory trajectories.
v. To investigate the impact of applying missing data techniques to the analyses for objectives i to iv.

1.4 Hypotheses

Aim 1:

Objective 1:

It is hypothesised that there will be a positive association between childhood SEP and adult cognitive function. It is less clear whether an effect will remain after adjusting for later life SEP. Previous work using the NSHD, but a different method of analysis, found a small but significant effect of childhood SEP on adult cognitive function after adjusting for later life SEP and childhood cognitive function (83). However the previous work involving datasets with a highly educated sample has found no direct effect of childhood SEP after adjusting for later life SEP, including an analysis carried out on the Whitehall II study (82), again using different methodology.

The missing data mechanism is very unlikely to be missing completely at random; therefore the missing data analyses will be an important aspect of the analyses.

Objective 2:

The simulation study aims to investigate how well each of the complete case, multiple imputation and Heckman selection analyses cope with each of the missing data mechanisms. Multiple imputation relies on the assumption of MAR, and Heckman selection allows for MNAR; however both of these methods are dependent on the appropriate identification of the imputation/selection model. The multiple imputation analyses should deal well with the MAR missingness, although it is less straight forward to identify an appropriate selection model.

Objective 3:

A limited amount of research has been carried out investigating the life course hypotheses with relation to cognitive function, and the life course hypotheses have not been systematically considered previously. Both the accumulation and social mobility models have been supported, but the models have not been compared in order to identify which life course hypothesis is best supported.

Aim 2:

Most of the studies examined found no association between childhood SEP and cognitive decline, with the only study to find a relationship investigating the odds of experiencing cognitive decline, rather than the rate of cognitive decline. It is therefore likely that there will not be an effect of childhood SEP on cognitive function, although the results may differ by SEP and cognitive measure. It is hypothesized that SEP later in life will be associated with cognitive decline, at least in unadjusted analyses.

1.5 Thesis structure

Chapter 2 gives an introduction to the methodological issue of missing data, a central aspect of this thesis, explaining why it is important and the common types of method to account for missing data. The two methods implemented in this thesis, multiple imputation and Heckman selection, are then described in detail. Chapter 3 introduces the two datasets used in this thesis: the National Survey of Health and Development (NSHD) and Whitehall II study of British civil servants. The datasets and variables used in the analyses are described.

Objective 1 of Aim 1 is addressed in Chapter 4, using both the NSHD and Whitehall II datasets. A simulation study is one way of comparing the results of missing data techniques; such a study, Objective 2, was carried out and is discussed in Chapter 5. Chapter 6 considers the life course hypotheses (Objective 3), and which of these hypotheses may be appropriate when investigating the relationship between life course SEP and cognitive function. Chapter 7 addresses methodological issues that arise in Chapter 6, including missing data.

Aim 2 is addressed in Chapter 8 which considers the effect of childhood SEP on memory decline using the Whitehall II study, before and after adjusting for later life SEP. Life course SEP is also addressed in Chapter 8, and missing data are accounted for. Chapter 9 draws together the work in this thesis, with a summary of the main results, as well a discussion of the relevance and implications of the thesis.

Chapter 2: Introduction to Missing Data

This chapter provides an introduction to missing data, and discusses the different types of missing data. It then describes the various missing data mechanisms, before discussing two methods of analysis for partially observed data, which will be applied in the analyses later in this thesis. The two methods are multiple imputation and Heckman selection models, a type of joint modelling.

2.1 What are missing data?

Missing data are observations which were intended to be made at the beginning of the study, but which were not made in practice. There are many reasons why missing data may occur in epidemiological studies; for example, when a sample member refuses to take any further part in the study, or when completing a questionnaire, mistakenly believes the question does not apply to them.

2.1.1 Types of missing data

In all types of study design, both unit non-response (Figure 2.1a) and item non-response (Figure 2.1b) are possible. Unit non-response is when an individual ('unit') does not answer any of the questions, e.g. individual 3 in Figure 2.1a, who has missing data for all three variables x1, x2 and x3. Item non-response is when the individual has taken part in the study but has not answered every question ('item'), e.g. individual 2 in Figure 2.1b, who has missing data for variable x2.

ID	x1	x2	x3
1	5.2	7.1	8.2
2	34.5	33.9	35.4
3	?	?	?
4	1.1	1.3	1.4

Figure 2.1a: Unit non-response

ID	x1	x2	x3
1	5.2	7.1	8.2
2	34.5	?	35.4
3	?	14.6	13.5
4	1.1	1.3	?

Figure 2.1b: Item non-response

In longitudinal studies with multiple data collections at various time points, wave non-response and attrition also occur. Wave non-response occurs when an individual does not answer any of the questions in a particular wave of data collection. Attrition is when an individual permanently drops out of the study, and therefore provides no data after they drop out.

2.1.2 Missing by design

Data may also be missing as part of the study design. The purpose of sampling is to collect data from a proportion of the population of interest, and therefore it could be said that all those from the target population who were not sampled are missing by design. More importantly, within a study sample, data may be missing by design. For example, if a questionnaire is too long to expect all participants to answer every question, it may be split into sections with different sub-groups completing different sections.

2.1.3 Methods for reducing the amount of missing data

It is, of course, preferable to minimise the extent of missing data not due to study design at the data collection stage. High levels of missing data are generally associated with poor study design (109). Careful design of questionnaires can reduce the amount of missing data, for example, the length of the questionnaire and the wording of each question are important. Retrospective questions may have particularly high rates of non-response, as participants are less likely to remember the relevant information in sufficient detail to be able to respond. Incentives, such as cash or participation in a raffle, have been found to be effective in improving response in single postal, telephone and face-to-face surveys (110). In longitudinal studies, the frequency of follow-up is a factor that may influence the participation rate. Follow-up needs to be often enough to satisfy the scientific purpose of the study, but if it is too often, participants may resent the amount of time taken for continued participation and drop out. Even with the most careful planning, however, some missing data and dropout are unavoidable in long running cohort studies.

2.1.4 Missing Data Patterns

Let the complete data be defined as a matrix Y, where $Y = (y_{ij})$, for individual *i* and variable *j*; the shape of the matrix Y then illustrates the missing data pattern. Three of the possible patterns are shown in Figure 2.2. The first pattern (Figure 2.2a) is univariate missingness, where the missing values are limited to a single variable (Y3 in the figure). This pattern may occur in a questionnaire where full data are received for all questions except one, which might be a particularly sensitive question. Figure 2.2b illustrates monotone non-response, where the variables can be arranged so that all $Y_{j+1} \ldots Y_k$ are missing for cases where Y_j is missing, for all *j*. An example of monotone non-response is attrition, where once a participant has dropped out of the study they provide no further data, but they have provided complete data until that point. Figure

2.2c shows general non-response, which is often observed when dealing with questionnaires, where most questions have some missing data. The general non-response pattern is often found in cohort studies with multiple waves of data collection.

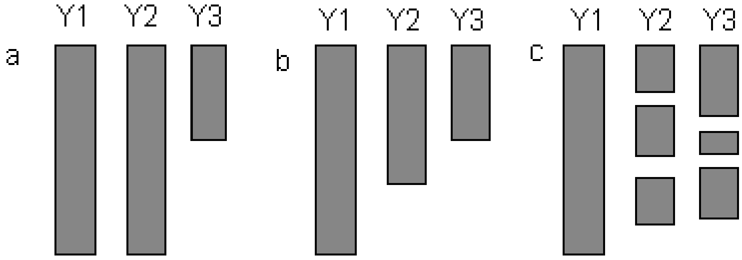


Figure 2.2: Non-response patterns: a - Univariate b - Monotone c - General

2.2 Missing Data Mechanisms

A critical issue to be considered when selecting the method of analysis is that of the mechanism behind the missing data. Using the notation of Little & Rubin (111), let the complete data matrix be defined as $Y = (y_{ij})$ and the missing data indicator matrix $M=(m_{ij})$, with $m_{ij} = 1$ if y_{ij} is missing and $m_{ij} = 0$ if y_{ij} is present. Let Y be comprised of Y_{obs}, those elements of Y that were observed, and Y_{mis}, the missing elements of Y. The missing data mechanism depends on $f(M|Y)$, the conditional distribution of the missing data indicator matrix, M, given the complete data matrix.

2.2.1 Missing Completely at Random (MCAR)

When $f(M|Y) = f(M)$ for all Y, the missing data mechanism is called Missing Completely at Random (MCAR). From this equation it is clear that the conditional distribution of M given Y does not depend on Y, either the observed or unobserved values. An example of data that is MCAR would be a single lab sample being accidentally dropped, leading to the results from that sample being missing. MCAR is unrealistic in most situations, and although the mechanisms may initially appear to be MCAR, on closer consideration it may not be.

2.2.2 Missing at Random (MAR)

When $f(M|Y) = f(M|Y_{obs})$ for all Y_{mis}, the missing data mechanism is called Missing at Random (MAR). The conditional distribution of M given Y does not depend on Y_{mis}, the unobserved values of Y, but depends only on the observed values of Y, Y_{obs}. An example of a situation in which the missing data are MAR is if individuals in a particular age group, for example young adults aged 18-25, are less likely to respond to a postal questionnaire than older individuals, perhaps because they are a more geographically mobile group. The missingness therefore depends on the participant’s age, which is known and is therefore part of Y_{obs}. An important point here is the implication that the distribution of $M|Y_{obs} = M|Y_{mis}$; that is that the distribution is identical for observed Y and missing Y. This means that the data with observed Y can be used to estimate the distribution of $M|Y$.

2.2.3 Missing Not at Random (MNAR)

When the conditional distribution of M given Y depends on Y_{mis} as well as Y_{obs}, the missing data mechanism is called Missing Not at Random (MNAR). In this case the distribution of M depends on what was not observed. An example of MNAR data is when a patient misses an appointment at which they would have taken cognitive tests because they have declined cognitively to the extent that they are no longer able to participate. The missingness is related to the cognitive measurements that would have been recorded had the patient been well enough to participate. In ageing studies it is not unfeasible that the missingness mechanism is MNAR.

2.2.4 Deciding between MCAR, MAR and MNAR

It is not possible to test for MNAR since it is only the missing data that distinguishes between MAR and MNAR assumptions (112). It is only possible to distinguish MCAR and MAR by assessing whether missingness of a variable is dependent on observed variables. A common method of dealing with the possibility of MNAR, which is likely to occur in epidemiological studies, is to carry out sensitivity analyses, testing different assumptions about the unknown underlying missing data mechanism.

2.2.5 Ignorable and non-ignorable non-response

MCAR and MAR data are sometimes said to be ‘ignorable’ missingness, as it is not necessary to model the non-response mechanism, whereas MNAR is ‘non-ignorable’ missingness (113), as it is necessary to model the non-response mechanism (109).

2.3 Types of methods for dealing with missing data

Little & Rubin (111) divided the available methods that have been proposed for dealing with the problem of missing data into four main groups: procedures based on completely recorded units, weighting procedures, imputation based procedures and model-based procedures. To be considered an appropriate method of dealing with the missing data, the estimators calculated from the data (such as the mean and standard error) must be unbiased, and give correct p-values and confidence intervals (114). These terms are defined in Chapter 5.

2.3.1 Procedures Based on Completely Recorded Units

Analyses based only on completely recorded units are known as complete case analyses. They ignore all cases that have any missing data, and are, therefore, easy to implement. Unless data are MCAR this approach can result in bias. Further, if there is considerable missing data then the sample size may be greatly reduced, resulting in loss of statistical power.

2.3.2 Weighting Procedures

The idea behind non-response weighting is that a participant's likelihood of responding can be predicted based on information that is known. Each sample member has a predicted probability of responding, p_i. The lower the probability, p_i, the greater the proportion of participants with any given set of characteristics that do not respond. Individuals with these characteristics who do respond are thus allocated a weight, $\frac{1}{p_i}$.

It is important that p_i is well predicted and is based on all observed variables which predict non-response. For weighting procedures to give unbiased results the missing data mechanism must be MAR.

2.3.3 Imputation Based Procedures

Imputation refers to a procedure by which each missing data value is 'filled in'. The way the value is imputed depends on the imputation method used. The resulting 'complete' dataset is then analysed using standard complete case methods. There are a variety of imputation based procedures, which are divided into single-imputation methods and multiple imputation methods. Single imputation methods, such as hot-deck, last observation carried forward (LOCF), regression and other ad-hoc methods, impute only a single value for each missing item. They do not, therefore, take into account the uncertainty in that value, but treat it as the true value. This results in the total uncertainty being underestimated, limiting the value of single imputation methods. Multiple imputation methods impute multiple values for each item to produce multiple complete datasets. Each value is imputed via a regression model with added variation. Each complete dataset is analysed and the results are combined according to formulae given by Rubin. These methods are unbiased when data are MAR, but can be biased when data are MNAR.

2.3.4 Model-Based Procedures

Model-based procedures are generated by defining a model for the observed data and basing inferences on the likelihood or posterior distribution under that model, with parameters estimated by procedures such as maximum likelihood (111). These require the assumption of MAR; however this assumption can be relaxed when the model of interest is associated with a model of missingness and they are estimated jointly (115).

2.4 Missing data methods investigated in this thesis

Multiple imputation (assuming MAR) and Heckman selection models (assuming MNAR) were chosen as the two methods to be investigated in this thesis. Multiple imputation has become available in common statistical packages, and has therefore been widely implemented. However, details of how the imputation model was chosen are rarely given in published papers, despite this being an important stage of the analysis. Sterne et al (116) present an example of a published paper where an initial multiple imputation analysis resulted in a surprising result; the authors later clarified that in a complete case analysis the result changed to one which supported the expected result, with a similar result found after improving the imputation procedure. This illustrates the importance of carrying out multiple imputation appropriately to produce results which can be trusted.

Heckman selection models have mainly been used in the economics literature to address sample selection bias. As the main issue with attrition in longitudinal studies is that those who drop out from a study are likely to have different characteristics from those who remain in the study, a method designed to account for selection bias is appropriate.

2.4.1 Multiple Imputation Methods

2.4.1.1 Introduction

Single imputation methods are not appropriate as the imputed value is treated as the true value, thereby reducing the estimated uncertainty, and thus will not be considered here. In 1987 Rubin (117) proposed multiple imputation as a method of generalising single imputation, allowing for this extra uncertainty by replacing each missing item with a vector of D imputed values, so D complete data sets are created from the imputations. It is then possible to account for the extra uncertainty by looking at the variation between the imputed values.

2.4.1.2 The imputation model

The default method of imputing missing values using the Stata command *ice*, created by Patrick Royston (118;119), is by sampling from the posterior predictive distribution $p(Y_{mis}|Y_{obs}, X)$. For *n* imputations, *n* independent selections are made from the posterior predictive distribution. All the variables which are thought to predict or be associated with the missing values should be included in the model used to form the posterior predictive distribution, the imputation model.

When choosing which variables to include in the imputation model, it is important to also include any variables that may be important in subsequent analyses, including the variable(s) with missing data and the outcome of interest. The converse of this rule is not necessary; if variable X_1 has been imputed from a model that includes variable X_2, X_2 does not need to be included in all analyses involving X_1 (although it can be). It has been shown that including as many explanatory variables in the imputation model as possible makes the MAR assumption more plausible (120). Although including redundant predictors may be expected to reduce the precision of the final estimates, Kenward and Carpenter (112) note that this effect is typically not large. They conclude it is therefore better to err on the side of including too many variables in the imputation model rather than too few, as excluding an important predictor of missingness could cause bias.

van Buuren (120) and Carpenter and Plewis (113) suggest similar methods for selecting the variables to include in the imputation model from a large database:

1. Include all variables that appear in the complete-data model of interest.
2. Include the variables which were predictive of missing data.
3. Check whether the variables included in step 2 were associated with the variables in the model of interest.
4. Auxiliary variables can also be included to make the assumption of MAR more likely. Auxiliary variables are predictive of any of the variables in the imputation model.

Collins et al (121) highlight the serious consequences of omitting important causes of missingness from the imputation model. They concluded that adding auxiliary variables which were not causes of missingness is ‘at worst neutral, and at best extremely beneficial’.

Multiple imputation (MI) can be adapted to be suitable for stratified samples by including strata indicators as covariates in imputation models (112).

2.4.1.3 Imputing interactions, squares and transformed variables

This section summarises an article by von Hippel (122) on how to carry out multiple imputation when there are squared or interaction terms in the model of interest. It is generally agreed that squared variables, interaction terms and transformed variables should be included in imputation models if appropriate, as otherwise it is assumed that no relationship exists between these squared variables (for example) and the outcome of interest. This would bias their coefficients towards zero in a regression analysis. There are two possible methods of treating squared and interaction terms in studies using multiple imputation. The first is to impute the missing values and then transform the variables, for example squaring or computing interactions, which ensures that the imputation model and the analysis model are compatible; for example, ensuring that all squared terms are equal to the value of the variable squared. The other method is to transform and then impute, which does not ensure that the squared term equals the value of the variable squared.

Under the 'impute, then transform' method, the regression estimates are biased. This occurs because the method does not correctly account for the relationship between the squared or interaction term and the outcome in the imputation model. This is the situation when the *passive* option is used in the Stata command *ice*. von Hippel concludes that the 'transform, then impute' method is the only appropriate method.

2.4.1.4 'Multiple imputation, then deletion'

von Hippel (123) suggests a modification to multiple imputation, which he calls 'multiple imputation, then deletion' (MID). This involves deleting all the cases which have imputed outcome values after the imputation but before the analysis. One advantage of MID is efficiency; "MID tends to give less variable estimates, more accurate standard-error estimates, and narrower confidence intervals with equal or higher coverage rates" (123). MID is also more robust to errors in the imputation model, as problems with imputing the outcome cannot affect the estimates. This is especially true in situations with a large proportion of missingness in the dependent variable, say 20-50% (123).

MID works because cases with imputed Y contain no information, but still add estimation error. Information is an alternative name for the log likelihood, which is equal to zero for cases where Y was imputed. MID is used in situations where there is missing outcome data as well as missing data in the independent variables. MID increases the precision of estimates despite increasing the standard error *within* each imputed data set (due to the smaller sample size), as the standard error *between* the datasets is smaller (123).

von Hippel also considers the situation of applying MID to repeated measures analyses. For newer methods, such as multilevel growth models, he concludes that it is appropriate, whereas for older methods such as repeated-measures MANOVA he concludes that it may be advisable to use partly imputed outcome data (i.e. cases which have some observed and some imputed outcome data). This justification is valid when Y gained no additional information through the imputation process. One way Y may gain extra information is by including additional variables in the imputation model that are not in the analysis of interest. When auxiliary variables do improve the imputation of Y, asymptotically (with an infinite number of imputations) MI will be more efficient than MID. However in practice we do have only a finite number of imputations.

von Hippel (123) carried out a simulation study to compare MI and MID using 2, 5 and 10 imputations, when adding a variable to the imputation model which was correlated to the dependent variable in the model of interest, but not the independent variables. When the correlation was low (0.1) and 2 imputations were used, the confidence intervals for MID were narrower and provided better coverage. However, when the correlation was 0.9 and 10 imputations were used, MI performed better. The experiment found that the 'tipping point' for 10 imputations occurred with a correlation of 0.5, with better results for MID for lower correlations and for MI for higher correlations. However in a real data set the additional variable may also contain some missing data, decreasing the information that it adds to the imputation of Y, making MI less advantageous.

Alternatively, if the missing data are non-ignorable, Y may gain extra information via the imputation process. This is because assumptions could be made about the data based on the fact that it is missing, and could be adjusted after regular MI was carried out. For example, if weight was missing for some individuals, and it was suspected that non-respondents had higher weights, then a fixed value could be added to the imputed weights. In these situations MID is inappropriate (123).

2.4.1.5 How many imputations are necessary?

The common view is that only a small number of imputations, namely 5-10, is necessary unless there is an unusually high amount of missing data (124). However, Kenward and Carpenter (112) found that in some instances many more imputations, up to 100-200, can be required before the results are adequately accurate in clinical trials. Bodner (125) carried out simulation studies investigating the impact of different numbers of imputations on p-values, confidence interval half-widths and estimated fractions of missing data. Bodner found that the greater the number of imputations, the less variance was observed in these measurements. Bodner tabulated (Table 3 in (125)) the number of imputations required for given fractions of missing information and confidence interval half-widths. As the examples were based on psychological research, where sample sizes are much smaller than in epidemiological studies, a sample size of 100 was used to calculate how many imputations would be required to calculate the confidence interval half-widths and estimated fractions of missing data within a given level of accuracy.

Rubin (117) showed that the efficiency of an estimate based on d imputations is approximately $\left(1+\frac{\gamma}{d}\right)^{-1}$, where γ is the fraction of missing information for the parameter being estimated, (see section 2.4.1.6). Graham et al (126) carried out a simulation study to investigate how many imputations were needed, allowing for varying fractions of missing information. They found that within each fraction of missing information, decreasing the number of imputations (i) increased the values of the mean square error and standard error; (ii) reduced the power; (iii) resulted in a less accurate estimate of γ, and (iv) increased the variability of γ. Table 2.1 shows the efficiencies based on the fraction of missing information and number of imputations.

	Fraction of missing information (γ)				
d	0.1	0.3	0.5	0.7	0.9
3	97	91	86	81	77
5	98	94	91	88	85
10	99	97	95	93	92
20	100	99	98	97	96

Table 2.1: Efficiencies for different fractions of missing information

2.4.1.6 The analysis of data completed by multiple imputation

To analyse data created by multiple imputation, each dataset is first analysed individually using the same methods as for a complete-case analysis. Using notation from Little and Rubin (111), let $\hat{\theta}_d$ and W_d , d = 1,...,D, be D complete-data estimates of parameter θ and their associated variances, calculated from D imputed data sets obtained under the same imputation model. The combined estimate is then obtained as follows:

$$\bar{\theta}_D = \frac{1}{D}\sum_{d=1}^{D}\hat{\theta}_d \ .$$

The variability associated with this estimate has two components: the average within-imputation variance,

$$\bar{W}_D = \frac{1}{D}\sum_{d=1}^{D} W_d \ ,$$

and the between-imputation component,

$$B_D = \frac{1}{D-1}\sum_{d=1}^{D}\left(\hat{\theta}_d - \bar{\theta}_D\right)^2 .$$

The total variance associated with $\bar{\theta}_D$ is

$$T_D = \bar{W}_D + \frac{D+1}{D} B_D ,$$

where $\frac{D+1}{D}$ is an adjustment for finite D.

The fraction of information about θ missing due to nonresponse can be estimated by $\hat{\gamma}_D = \left(1+\frac{1}{D}\right)\frac{B_D}{T_D}$, which measures the relative increase in variance due to the missing data (124).

2.4.1.7 Limitations of multiple imputation

The assumption underlying the use of multiple imputation is that the missing values are 'missing at random' (see section 2.2). However this assumption may not always be plausible, for example if a participant becomes cognitively impaired and is thus unable to participate further in the study. It is also possible to mis-specify the imputation model, which may lead to incorrect conclusions (116); for example Allison (127) points out that non-linear relationships may be missed. Inappropriate assumptions of normality may result in unrealistic imputations, such as negative heights, or dummy variables not equal to zero or one. When ‘fixing’ these unrealistic imputations, adjusting them so they are realistic values, biases may be introduced (123).

2.4.2 Heckman selection models

There are two main methods of accounting for missing data which are MNAR; sensitivity analyses and joint modelling. Sensitivity analyses normally involve fitting a range of plausible models for dropout, such as assuming everyone who has dropped out has the minimum value, or the maximum value; if the results are reasonably robust to the different dropout models this is often used as a justification for ignoring missing data. However further investigation is required if different results are found.

There are two methods of joint modelling, pattern mixture models and selection models. In pattern mixture models the data are stratified according to which observations are missing, and the distribution of the full data is treated as a mixture of distributions over these missing data ‘patterns’ – so a separate model is fitted to each stratum. However in studies with many phases of data collection this method is not practical. In a study with three phases there are $2^3 = 8$ possible strata; for the NSHD, with 22 phases of data collected there are 2^{22}, more than 4 million possible strata.

Selection models assume that the participants with incomplete data are random samples from the same distribution as those with complete data, and that a selection of the participants drop out according to their response values. Different selection models are appropriate for different types of selection. For example, a censored regression or tobit model is appropriate when the dependent variable is censored at some bound as a result of how the data are collected. An example of this is the work by Mare and Chen (128), investigating the effect of parents’ socioeconomic characteristics on the years of school their children completed, with the variable censored for individuals with more than 12 years of schooling.

The standard sample selection model generalizes the tobit model to allow selection (whether the variable is observed or not) to depend on the values of other variables. One of the most widely used selection models is the Heckman selection model, which is widely available in statistical packages. Heckman initially derived the model when considering wages; only people who are working receive a wage, and workers are not a random selection of the population. The Heckman selection model is appropriate when an underlying regression relationship of interest exists (Equation 2.1), but the dependent variable is not always observed; rather the dependent variable is only observed if the selection inequality (Inequality 2.2) holds.

Equation 2.1: Regression equation

$$y = x\beta + u_1$$

where *y* is the dependent variable of the analysis of interest, *x* is the vector of covariates, β is the vector of coefficients and u_1 is the error term.

Inequality 2.2: Selection inequality

$$z\gamma + u_2 > 0$$

where *z* is a vector of regression variables, γ is a vector of coefficients and u_2 is the error term.

$$u_1 \sim N(0, \sigma^2)$$

$$u_2 \sim N(0,1)$$

$$\text{corr}(u_1, u_2) = \rho.$$

The classic example of the application of the Heckman selection model is the prediction of wages in women, as women choose whether to work, and this decision is not random, but is related to the wage that would be earned if they chose to work. The wage is then only observed if the condition (2.2) holds.

A solution to the Heckman model exists if there is at least one variable which strongly predicts selection into the model, but not the outcome of interest in Equation 2.1 (129). The Heckman selection model is strongly dependent on the model being correct (129), however the Heckman selection model offers no guidance as to how to select the covariates which should be included in the selection model (130).

The following description of how the Heckman selection model works draws heavily on the explanation by Winship and Mare (131). Equation 2.1 can be rewritten as Equation 2.3.

Equation 2.3: Equation 2.1 rewritten

$$y = x\beta + E(u_1 | u_2 > -z\gamma) + \eta$$

The Heckman estimator involves estimating the selection model, calculating the expected error $\hat{u}_2 = E(u_2 \mid u_2 > -z\gamma)$, for each observation using the estimated γ, and using the estimated error as a regressor in the regression model of interest.

The assumption that u_1 and u_2 are bivariate normal is required, and using this assumption, along with var(u_2) = 1, then E(u_1 | u_2)=$\sigma\rho u_2$ and

$$E(u_1 \mid u_2 > -z\gamma) = \frac{\sigma\rho\phi(-z\gamma)}{1-\Phi(-z\gamma)} = \sigma\rho\lambda(-z\gamma),$$

where φ and Φ are the standardised normal density and distribution functions respectively. The ratio $\lambda(-z\gamma)$ is the inverse Mills ratio, which can then be substituted into the regression equation of interest:

$$y = x\beta + \sigma\rho\lambda(-z\gamma) + \eta$$

where η is uncorrelated with both x and $\lambda(-z\gamma)$. The assumption that u_1 and u_2 are bivariate normal is used to obtain a linear relationship between u_1 and u_2 and to get a marginally normal error u_2, which produces the Mills ratio formula.

The precision of the estimates in the regression model of interest is sensitive to the variance of λ and collinearity between x and λ (131). The variance of λ is influenced by how accurately the selection model predicts selection into the sample; the better the prediction, the larger the variance, and the better the precision of the estimators. Collinearity will partly be determined by how many variables are shared by the selection model and the regression model of interest (131).

The probit selection model assumes that the errors u_2 are homoskedastic; when this assumption does not hold the Heckman procedure gives inconsistent estimates (131).

Many studies have reported that Heckman model estimates were sensitive to the selection model (130;132;133). When choosing the variables to make up the selection model, Lalonde (132) suggests that the variables in the selection model should contain at least one variable which is related to selection, but is not an independent variable in the regression model of interest; this advice is also given in the Stata Heckman manual (129). However a stricter condition is considered necessary by some (134), requiring a variable which is associated with selection, but not associated with the outcome in the analysis of interest. Whilst these variables should be chosen via a theoretical basis, this is not always possible in observational studies. There are many possible criteria for choosing a ‘best fitting’ selection model; one is to consider the pseudo R^2 value, another is to use likelihood ratio tests, but this is only possible when the models are nested. A third method considers the model with the largest proportion of statistically significant coefficient estimates to be the best fitting model; this is potentially important as the inverse Mills ratio is calculated from the estimated coefficients in the selection model, regardless of whether they are significant (130). Another method of comparing selection models is to consider the proportion of participants who are predicted to be in the correct category.

2.5 How missing data fits into this research

As Carpenter and Plewis (113) mention, it is rare that ‘quantitative social science investigations’ explicitly examine the effect of missing data on the results, a problem which is magnified in analyses using longitudinal data due to attrition between phases. This thesis aims to carry out a thorough investigation of the effect of missing data under each of the missing data mechanisms on the relationship between SEP and cognitive function/cognitive decline as set out in the aims and objectives of this thesis (see section 1.5). It is unlikely that the missing data in either of the longitudinal studies used in this research is MCAR. It is possible that the missing data is MNAR, which would be the situation if a study member could not participate in the study due to their low level of cognitive function or high level of cognitive decline. Alternatively it may not be the cognitive function directly which stopped a study member from participating, but ill health, which is also associated with cognitive decline.

The aim of addressing missing data in this thesis is to assess the sensitivity of the results to the missing data methodology (and corresponding underlying assumptions), with the simulation study in Chapter 5 allowing more detailed conclusions to be drawn. This will enable practical advice to be provided for other researchers working with longitudinal datasets with missing data.

Chapter 3: Introduction to the data

Data from two longitudinal studies, the Medical Research Council (MRC) National Survey of Health and Development and the Whitehall II study of British civil servants, are used to address the aims outlined above (section 1.5). First some background to each study is given, after which the relevant variables available in each study are described. Finally, the advantages of the two datasets are discussed.

3.1 The National Survey of Health and Development

3.1.1 Study design

The MRC National Survey of Health and Development (NSHD), also known as the 1946 British birth cohort study, was started in 1946 to look at the availability and effectiveness of ante-natal and maternity services in Britain. The idea was that the results from this study would be used for comparison after the National Health Service (NHS) was set up in 1948 (135). The study was continued, and the 22nd wave of data collection on the full cohort took place in 2006-2011.

A total of 16,695 births were registered in England, Scotland and Wales in one week in March 1946. The maternity study collected information from 13,687 of these. Multiple births (n=180) were excluded from the study, as were the 672 children born to non-married women, who were not followed up due to the high levels of adoption, which made them impossible to trace (136). A stratified random sample of 5,362 was selected for follow-up; all single births to married women whose husbands were in non-manual and agricultural occupations were selected, along with one in four of the single births to married women with husbands in manual occupations.

As this was a stratified sample with different probabilities of selection in the different groups, and there is reason to believe that some outcomes may differ depending on which of the groups the survey member comes from, analyses should be weighted. Since 1 in 4 of those born to fathers in manual occupations were selected, they have a weighting of 4, compared to those born to non-manual or agricultural fathers, who have a weighting of 1. This is important as otherwise the results would not be representative of the population from which the sample were selected.

Until cohort members were age 15 years data were collected from the mother and/or the cohort member, after that data were only collected from cohort members. Data were collected 11 times from birth to age 15. In adulthood data collection became less frequent due to the slower rate of biological and cognitive change, and due to cost (137).

In childhood the research interest focused on socioeconomic differences in infant growth and development; during school years this was extended to cover educational experience and attainment. In early adulthood, occupation and income were investigated as outcomes of education. In mid-life the study's health data collection was refocused to measure physical and mental function, and to allow the study of pathways leading to these outcomes (137). At ages 36, 43 and 53 years teams of trained research nurses visited study members in their own homes to carry out interviews and measurements. Table 3.1 (adapted from (136)) gives more detail on the data collection process.

Ethical approval for the data collection at age 53 years was issued by North Thames Multi-centre Research Ethics Committee (MREC 98/1/121). Earlier waves of data collection had appropriate ethical approval.

3.1.2 Response rate and representativeness of the cohort

By 1999, when the participants were aged 53 years, 469 (8.7%) of the participants had died, of whom 230 (49%) had died in infancy (below the age of 5). By age 53 there were 640 (11.9%) permanent refusals who will not be contacted in any further waves of data collection (137). The population successfully contacted from those who had not died, emigrated, permanently refused or were temporarily abroad has remained high throughout the study, and was 83% (3035/3673) at age 53 (Table 3.1).

The cohort aims to be representative of all single births born to married mothers in England, Scotland and Wales, who have remained living in England, Scotland or Wales. No additions have been made to the sample to adjust for the immigration that has occurred since 1946. Data from the responding sample at ages 43 (138) and 53 (136) were compared with relevant census data to assess the representativeness of the remaining participants. They were found to be fairly nationally representative when age-specific data were available. Compared with national data, more men and women in the

NSHD were employed full time at both ages 43 and 53, and less men and women in the NSHD were single at both ages 43 and 53 (136).

Many childhood characteristics were associated with increased ‘avoidable loss’ from the study by age 53, defined as losses through permanent or temporary refusal, or those who could not be traced, such as paternal manual social class, low cognitive test scores and frequent problems with discipline, disobedience and aggression. However, this was partly balanced out by the ‘brain drain’ emigration of those study members living abroad, who were more likely to come from non-manual families, and had above average cognitive scores (136).

Table 3.1: Follow up of NSHD cohort members (adapted from Wadsworth et al)

Year	Age	Respondent	Data collector [a]	Location	Target sample [b]	Achieved sample (% achieved)
1946	8 weeks	Mother	HV, M, O	Home	5,362	5,362 (100)
1948	2	Mother	HV	Home	4,993	4,698 (94)
1950	4	Mother	HV	Home	4,900	4,700 (96)
1952	6	Mother and cohort member	SD	School	4,858	4,603 (95)
1953	7	Mother and cohort member	SN or HV	School	4,838	4,480 (93)
1954	8	Mother and cohort member	SN or HV & T	School	4,826	4,435 (92)
1955	9	Mother and cohort member	SN or HV & T	School	4,807	4,181 (87)
1956	10	Cohort member	T	School	4,811	4,077 (85)
1957	11	Mother and cohort member	SN or HV SD T	School	4,799	4,281 (89)
1959	13	Cohort member	T	School	4,794	4,127 (86)
1961	15	Mother and cohort member	SN or HV & T	School	4,790	4,247 (89)
1965	19	Cohort member	HV	Home	4,741	3,561 (75)
1966	20	Cohort member	P	Home	4,715	3,899 (83)
1968	22	Cohort member	P	Home	4,638	3,885 (84)
1969	23	Cohort member	P	Home	4,518	3,026 (67)
1971	25	Cohort member	P	Home	4,446	3,307 (74)
1972	26	Cohort member	I	Home	4,410	3,750 (85)
1977	31	Cohort member	P	Home	4,293	3,340 (78)
1982	36	Cohort member	RN	Home	3,863	3,322 (86)
1989	43	Cohort member	RN	Home	3,839	3,262 (87)
1999	53	Cohort member	RN	Home	3,673	3,035 (83)

[a] HV=health visitor, M=midwife; O=obstetrician, SD=school doctor, SN=school nurse, T=teacher, P=postal questionnaire, I=interviewer, RN=research nurse.

[b] Excludes the dead, those living abroad, and permanent refusals.

3.1.3 Variables of interest

A wide variety of data were collected, covering socioeconomic, medical and psychological variables. Although this thesis focuses on the relationship between SEP and cognitive function, other variables were also required for the analyses.

3.1.3.1 Socioeconomic Position (SEP)

SEP covers a wide range of variables (section 1.2.9), which may have different relationships with each outcome. As the thesis investigates the effect of life course SEP, SEP variables were required from each stage of the life course.

Childhood SEP

Father's occupational SEP

The participant's father's occupational SEP was collected at age 4. All occupational social classes were defined according to the Registrar General's classification, which divides occupations into six classes. For some of the analyses in this thesis the 6 categories are divided into those which are non-manual (I: professional, II: managerial and technical, IIINM: skilled non-manual) and those which are manual (IIIM: skilled manual, IV: partly-skilled manual, V: unskilled manual). The Registrar General classification is viewed as a hierarchy, with all non-manual occupations considered to represent a 'higher' SEP than manual occupations (98).

This classification scheme was devised by T. H. C. Stevenson in 1913, and revised into its current form in 1921, when more emphasis was given to skill and the 'standing within the community' of the various occupations. In 1980 this was changed so that 'social class was equated directly with occupational skill' (139). Other changes have occurred between 1913 and 2011, with the same occupation falling into different categories at different time points, for example, postmen were moved from class II to class IV in 1961 (139).

Childhood material deprivation

The measure of childhood material deprivation was the number of amenities in the home that the study members were lacking at age 2. The three amenities considered were running hot water, their family having access to their own kitchen and their family having access to their own bathroom. This variable was considered as a categorical variable scored 0-3.

Early adulthood SEP

Educational qualifications

The highest level of educational qualification achieved by age 26 has been recorded. This variable was used in the current analyses as a categorical variable with 4 categories; none attempted/vocational course (proficiency only), sub GCE/sub Burnham C/GCE O level/Burnham C, GCE A level/Burnham B/Burnham A2, and 1st degree/masters degree/doctorate degree.

Occupational SEP

The participant's own occupational SEP at age 26 was collected and grouped using the Registrar General classification.

Midlife Adult SEP

Own occupational SEP

The participant's own occupational SEP at ages 36, 43 and 53 were collected. In these analyses a variable indicating SEP at age 43 years was used. Where an individual participated at age 43 but did not have an occupational SEP recorded, SEP from age 36 years was used. Occupation was again grouped using the Registrar General classification.

Head of household occupational SEP

The head of household's occupational SEP at age 43 was derived from information collected on the participant's occupational SEP and their partner's occupational SEP. For women the male partner's SEP was recorded if available, otherwise the woman's SEP was recorded. For men their own occupational SEP was recorded.

3.1.3.2 Cognitive function

For cognitive function, data are available at ages 8, 11, 15, 43 and 53 years. Scores at age 8 were selected to represent childhood cognitive function in these analyses, as it is the earliest available measure, and is therefore the least affected by schooling. At age 8, four cognitive tests were taken: a 60-item non-verbal picture test published by the National Foundation for Educational Research (N.F.E.R), a 35-item reading comprehension test published by the N.F.E.R., a 50-item mechanical word reading test and a 50-item vocabulary test. During school years (ages 5-15) the cognitive tests were sat at school and supervised by the children's teachers. The tests were marked by the supervising teacher, and all were re-marked at the National Foundation.

The adult cognitive measures cover crystallized and fluid cognition. At age 53, the National Adult Reading Test (NART) was considered as the measure of crystallized cognitive function. It is effectively a test of knowledge acquisition, although it correlates with full-scale IQ (140). The NART is a pronunciation test, where participants are given a list of words which are pronounced differently to how they appear, and are therefore not likely to be pronounced correctly unless the participant is familiar with the word (81). The NART was administrated by research nurses at home visits (136).

As there were only two repeat measures of the fluid cognitive variables (43 and 53 years), trajectories of fluid cognitive function were not analysed using the NSHD. The second study used in this thesis, the Whitehall II study, has more repeated cognitive data and was therefore used to study trajectories (Chapter 8). These cognitive function variables were, however, considered for the imputation model and the selection model in the missing data analyses (section 4.5.1).

3.1.4 Descriptive results

3.1.4.1 Socioeconomic position

Childhood SEP

Father's occupational SEP

The distribution of father's social class in 1950 can be seen in Figure 3.1. The skilled manual category (Registrar General category IIIM) had the largest frequency, making up 32% of those with a recorded occupational SEP.

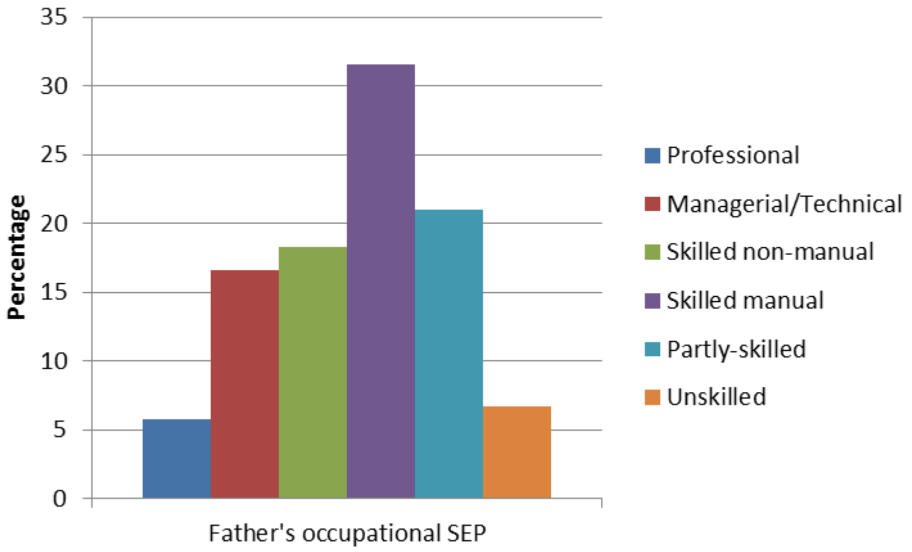


Figure 3.1: Bar chart of father's occupational social class in 1950

Childhood material deprivation

The modal group was those participants who had access to all three amenities at age 2 (47%), with only 5% not having access to any of the amenities (Figure 3.2).

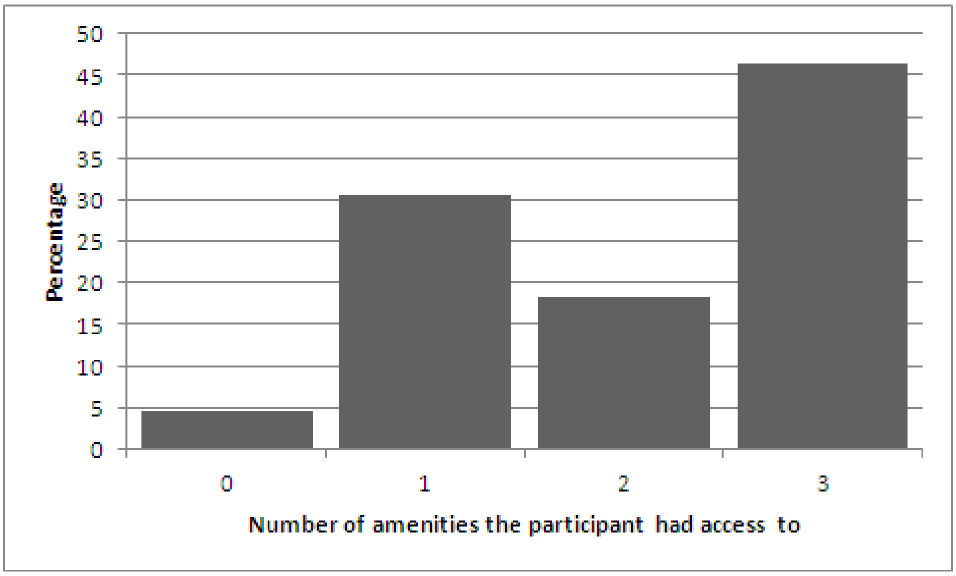


Figure 3.2: Bar chart of access to amenities at age 2

Early adulthood SEP

Educational qualifications

The modal group for educational qualifications achieved by age 26 was none attempted for both men (42%) and women (45%) (Figure 3.3). More than twice as many men as women had a degree (13% vs. 5%).

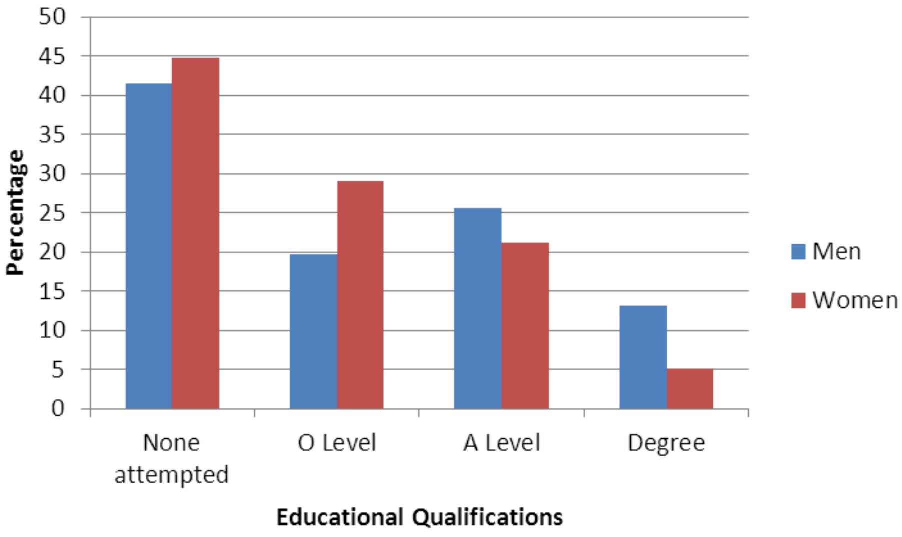


Figure 3.3: Bar chart of educational qualifications attained by age 26, by gender

<u>Occupational SEP</u>

Participant's own occupational SEP at age 26 is displayed in Figure 3.4. The modal group was skilled manual (IIIM) for men and skilled non-manual (IIINM) for women. There was a much higher proportion of men (11%) than women (1%) in the professional (I) category. 29% of women were in a manual category, compared to 51% of men.

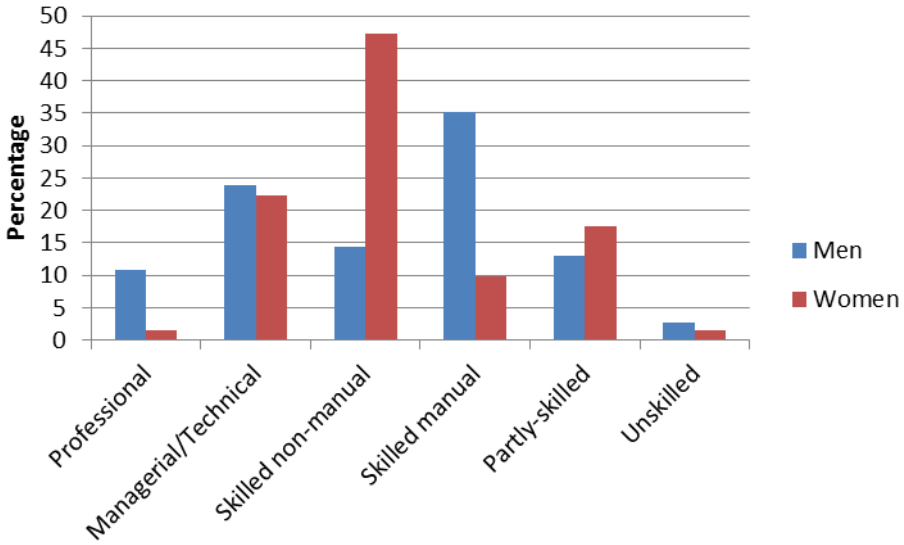


Figure 3.4: Bar chart of own occupational SEP at age 26

<u>Adult SEP</u>

<u>Own occupational SEP at age 43</u>

Frequencies in the 6 categories of participant's own occupational SEP at age 43 are displayed in Figure 3.5. The modal group was managerial/technical (II) for men and skilled non-manual (IIINM) closely followed by managerial/technical (II) for women.

There was a much higher proportion of men (11%) than women (1%) in the professional (I) category. 27% of women were in manual occupations, compared to 38% of men.

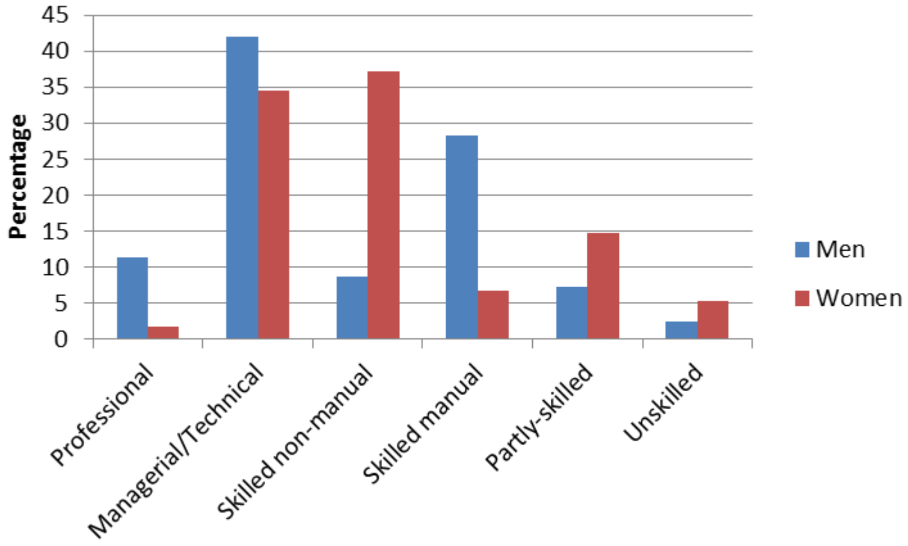


Figure 3.5: Bar chart of own occupational SEP at age 43

<u>Head of household occupational SEP at age 43</u>

The distribution of head of household occupational social class compared to the participant's own can be seen in Figure 3.6 for women. Data were available from an additional 306 women when head of household social class was used. As head of household SEP was defined as own occupational SEP for men, there was no difference for men. For women, a larger proportion were classified as IIIM (skilled manual) when head of household SEP was used, and many fewer were classified as IIINM (Skilled non-manual). Additionally a larger proportion were classified as I (professional) and II (managerial/technical), and fewer as IV (partly-skilled) and V (unskilled).

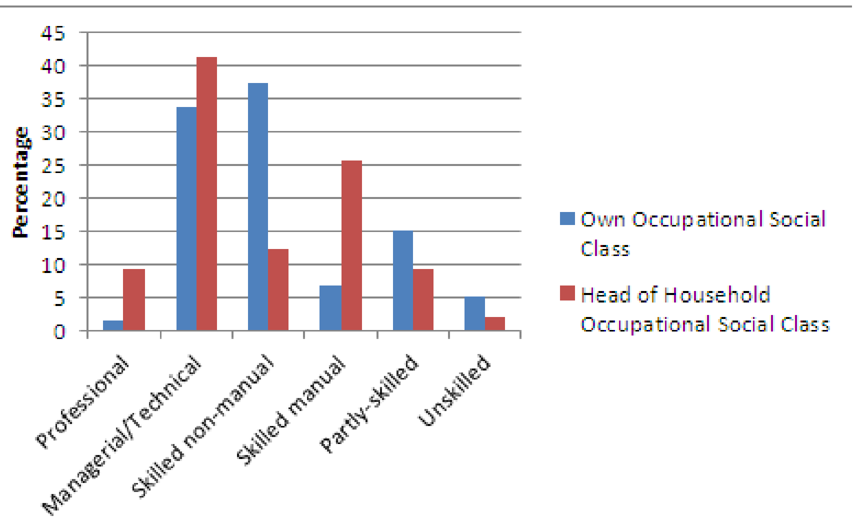


Figure 3.6: Bar chart of age 43 own social class and head of household social class, for women

3.1.4.2 Cognitive function

Complete information on all 4 cognitive tests taken at 8 years was available for 4,256 of the 5,362 original participants (79%). The picture intelligence scores were negatively skewed (Table 3.2 and Figure 3.8), whereas the reading (Figure 3.9) and vocabulary scores (Figure 3.10) were positively skewed. Correlations between each pair of the four tests were all strong at over 0.5 (Table 3.3). The highest correlation was 0.87 between reading and sentence completion. When the cognitive scores were summed the total score was normally distributed (Figure 3.11).

Table 3.2: Descriptive statistics for age 8 cognitive variables

Cognitive Variable	N	Weighted Mean* (95% C.I.)	Skewness
Reading	4259	15.66 (15.32, 16.00)	0.317
Picture Intelligence	4266	39.25 (38.90, 39.59)	-0.717
Sentence Completion	4259	13.15 (12.89, 13.42)	0.037
Vocabulary	4259	15.44 (15.24, 15.64)	0.184

* Weighted to allow for the stratified sample

Table 3.3: Correlations between age 8 cognitive variables

Age 8	Reading	Sentence Completion	Picture Intelligence	Vocabulary
Reading	1			
Sentence Completion	0.868	1		
Picture Intelligence	0.524	0.566	1	
Vocabulary	0.689	0.675	0.577	1

As childhood cognitive function squared was also required for the analyses, a standardised childhood cognitive function variable was created, as the range of the summed cognitive scores squared would be 0-30,976, with a large proportion of the values unattainable. The sample was first restricted to those participants who had a score recorded for each of the four cognitive tests taken at age 8. Each of the four tests was then standardised; then the four standardised variables were summed, and the resulting variable was standardised. This variable was then squared to form a (childhood cognitive function)^2 variable.

'The Home and The School' (135), the book which describes the early data collections in the study, discussed whether to treat the test scores individually or whether to combine them to give an average score for each age. The conclusion reached was that the circumstances which are associated with a drop in the school achievement tests are also associated with a drop in mental ability tests, so an average of the test scores should be used. By standardising each variable the tests are assigned equal weightings when they are combined.

Douglas, who initiated and directed the NSHD for 33 years (137), also discusses a possible source of bias in this study; that the family's attitude towards the things they are asked about may be altered through the process of observing the participants as children, and talking to their parents about their education,. To test whether this was true for cognitive tests in childhood, the 11+ results were compared with those of one third of the manual workers' children who, despite being born in the survey week, were not selected for the study. No significant difference was found, so the surveyed group was considered representative in terms of cognition of all children taking the 11+ that year.

Figure 3.7: Bar chart of sentence completion scores at age 8

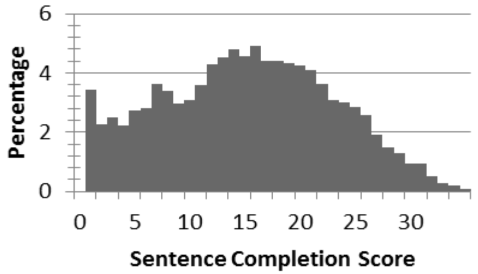


Figure 3.8: Bar chart of picture intelligence scores at age 8

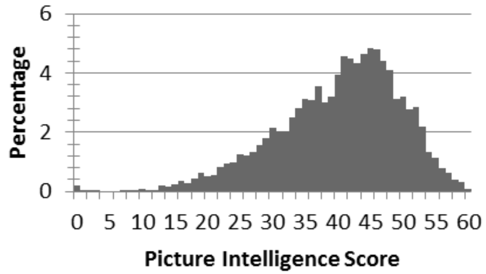


Figure 3.9: Bar chart of reading scores at age 8

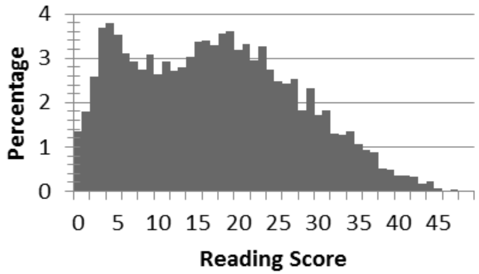


Figure 3.10: Bar chart of vocabulary scores at age 8

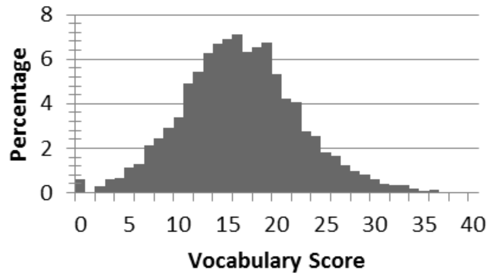
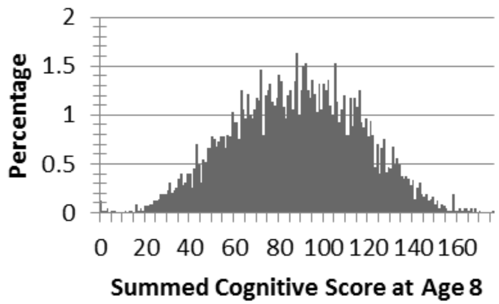


Figure 3.11: Bar chart of summed cognitive scores at age 8



The NART (2,825 (53%) participants completed) was slightly negatively skewed (-0.682) (Figure 3.12). A variety of transformations were assessed, including square, cubic, square root and log, but none significantly improved the normality of the distribution. A normal probability plot showed the data not to be far from a normal distribution (Figure 3.13).

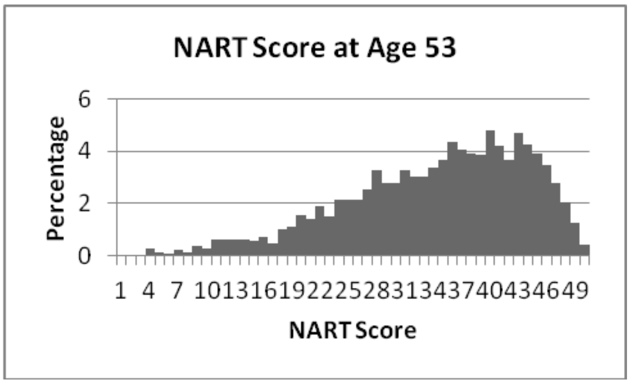


Figure 3.12: Bar chart of NART score at age 53

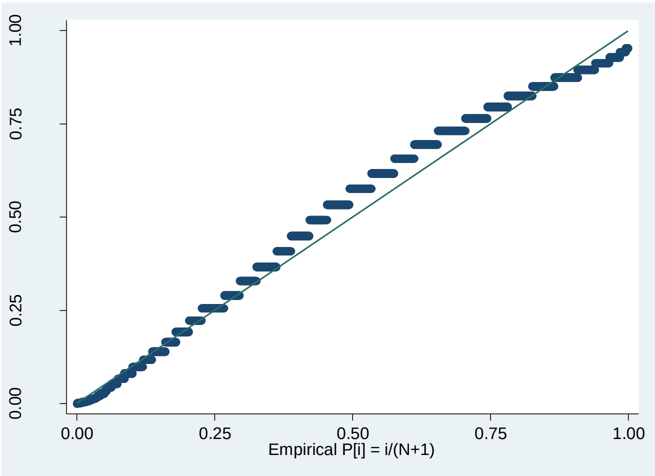


Figure 3.13: Standardised normal probability plot of the NART

3.1.5 Benefits of the NSHD

The NSHD has the benefits of regular longitudinal studies, namely the ability to study changes in individuals over time, so any differences observed are less likely to be due to cultural differences across generations. They also allow the development of exposure variables over time to be related to the development of outcome variables.

The additional advantages of the NSHD are that the data collection started at birth, and was nationally representative. The prospective nature of the study ensures the reliability of the data, without the worry of recall bias. Repeat data has been collected on many variables, allowing development and decline to be measured and modelled. A wide range of variables were collected, covering social, biological and psychological data.

Cognitive data are available from childhood, which allows for investigation into whether any relationship between childhood SEP and adult cognitive function acts through childhood cognitive function. Considering the length of time the study has been running, the levels of follow-up are high. The regularity of the data collections helps maintain contact with the participants, as well as the birthday cards sent to the participants every year, which maintains a relationship with the participants, reducing the number lost due to loss of contact. The regularity of data collection also ensures that the data covers all periods of the life course.

3.2 The Whitehall II study

3.2.1 Study design

The original Whitehall study was set up by Donald Reid and Geoffrey Rose in 1967, following 18,000 middle-aged men employed in the British civil service. One of the initial conclusions from the study was that an inverse relationship existed between employment grade and mortality, with only a third of the mortality risk differential between clerical and Unified Grades 1-6 grades explained by conventional risk factors (141;142). In 1985 the Whitehall II study was set up to investigate factors other than SEP which might contribute to this inequality in mortality. The baseline survey took place in 1985-1988, and recruited 10,308 non-industrial civil servants aged 35-55 from the London-based offices of 20 civil service departments. Ten phases of data have been collected, with phase 10 completed in 2011, and phase 11 is currently underway. Five of the completed phases have included a medical screening and self-completion questionnaire, with alternate phases involving a postal questionnaire. In order to provide estimates of reliability for the various measures, repeat data were collected from a subsample of the participants 3 months after phases 3, 7 and 9.

Ethical approval for the Whitehall II study was obtained from the University College London Medical School Committee on the ethics of human research.

3.2.2 Response rate and representativeness of the cohort

The target population for the Whitehall II study was all London-based office staff working in 20 civil servant departments in 1985 (n = 14,121), and there was a response rate of 73%. Anyone who did not respond to the baseline survey was not followed up. Response rates to the follow-up phases are therefore expressed as a percentage of the 10,308 (6,895 men, 3,413 women) baseline respondents. As can be seen from Table 3.4

(143), 66% of phase 1 respondents participated in the study at phase 9. Of those who did not participate in the study at phase 9, 737 (7.1% of the baseline respondents) had died by 31st July 2006 (144). Ferrie et al (144) investigated whether non-response was associated with mortality up to the end of phase 5, and concluded that both non-response to baseline and non-response to follow-up were associated with double the mortality hazard, after adjusting for age and sex, with the highest hazard ratio found for those who were baseline non-responders (HR=2.77, compared to those who had responded at all phases). They found that this association was not affected by adjustment for SEP (HR=2.52). This indicates that there are differences between those people who participated in the first phase of the study, and in further phases, and those who did not. It is therefore important to properly account for the missing data and the missing data mechanism when analysing data from the Whitehall II study.

Table 3.4: Data collection in Whitehall II

Phase	Dates	Type	Number of participants	Response Rate
1	1985-1988	Screening / questionnaire	10,308	73% of those invited
2	1989-1990	Questionnaire	8,132	79% of phase 1 responders
3	1991-1994	Screening / questionnaire	8,815	86% of phase 1 responders
4	1995-1996	Questionnaire	8,628	84% of phase 1 responders
5	1997-1999	Screening / questionnaire	7,870	76% of phase 1 responders
6	2001	Questionnaire	7,355	71% of phase 1 responders
7	2002-2004	Screening / questionnaire	6,967	68% of phase 1 responders
8	2006	Questionnaire	7,173	70% of phase 1 responders
9	2007-2009	Screening / questionnaire	6,761	66% of phase 1 responders

3.2.3 Variables of interest

Data collected included socioeconomic, physical health, mental health, health behaviour, biological and cognitive data, as well as retrospective data about the participants' childhood.

3.2.3.1 Socioeconomic Position (SEP)

Information on current SEP has been collected at each phase; additionally, measures of childhood SEP were collected, as well as educational qualifications attained at the end of full time education.

Childhood SEP

Father's occupational SEP

Father's occupational SEP during childhood was collected retrospectively at phase 1, using the Registrar General classification.

Childhood material deprivation

The questions used in this study to form a childhood material disadvantage variable came from the EPIC Health and Life Experiences Questionnaire and the Childhood Experience of Care and Abuse interview (145). The questions were:

"Did any of the following things happen during your childhood (that is, until you were 16)?

Your father/mother were unemployed when they wanted to be working Yes No"

And

"Did you experience any of the following circumstances during your childhood (that is, until you were 16)?

Your family had continuing financial problems	Yes	No
Your family/household did not have an inside toilet	Yes	No
Your family/household owned a car	Yes	No"

Factor analysis was used to reduce the number of variables representing childhood material deprivation, and to investigate whether they could be summarized by one factor. Tetrachoric factor analysis was used to account for the binary nature of the variables, as "factor analysis applied to dichotomous variables leads to artificial results" (146). For material deprivation the first factor had an eigenvalue of 1.45, whereas the second factor had an eigenvalue of 0.14. Therefore one factor was chosen as the optimum number of factors, which indicated that the four variables all loaded onto the same underlying factor, which was designated childhood material deprivation. The material deprivation score was then created using the resulting weightings. In the life course analyses (chapter 6) where a binary variable was required, the number of deprivations experienced was used, with a cut-off of ≥3.

Early adulthood SEP

Educational qualifications

At phase 5 the participants were asked about the highest level of educational qualifications they had attained. This variable was then categorised into 4 categories; 1: no academic qualifications, 2: school certificate/matriculation, 3: ‘O’ Level/ GCSE/‘A’ Level/SCE higher/‘S’ Level/National Diploma/Certificate and 4: BA/BSc/ University or CNAA Higher degree. These are slightly different to the categories used in the NSHD, reflecting the lower levels of educational qualifications in earlier cohorts.

Adult SEP

Occupational SEP

At each phase in the study the participant’s current grade was recorded if they still worked in the Civil Service. Employment grade in the civil service relates to income and associated living conditions (147). A variable was derived to record the last recorded grade of those participants who were no longer working in the civil service, but were still participating in the Whitehall II study, and the last recorded occupational SEP at phase 7 was used in these analyses.

3.2.3.2 Cognitive function

Cognitive data were available at phases 3 (age 39-64), 5 (age 44-69), 7 (age 50-74) and 9 (age 55-80). The cognitive function tests were introduced part of the way through phase 3, with only 39.9% of phase 3 participants completing any of the cognitive tests. Crystallized cognitive function was assessed via the Mill Hill test (148), a vocabulary test which assessed the participants’ understanding of words. It has a multiple choice format, with one point given for each correct answer, to a maximum possible score of 33. The measure of fluid cognitive function used in this thesis was verbal memory, which was assessed using a 20-word short-term verbal memory test, where participants listened to a tape recording of 20 words at 2-second intervals and were then asked to write down as many as they could (70).

In order to take account of practice effects, the number of times the cognitive tests had been taken up to and including phase 9 was derived from the data. As cognitive function tests were included in the repeat samples that took place after phases 3 and 7, the number of times the tests had been taken ranged from 0 to 5. Practice effects are discussed in more detail in section 4.3.1.

3.2.3.3 Other variables

Some variables which do not fit under the headings of SEP or cognitive function were also used in this thesis. Only the variable whose derivation requires explanation is described in this section; the remaining variables are mentioned in section 4.5.2, where the imputation and selection models are developed.

A childhood emotional deprivation variable was derived for consideration in the imputation model and selection model. It was derived using tetrachoric factor analysis, from four binary variables collected retrospectively at phase 5: parents divorced during childhood, parents argued during childhood, parents mentally ill/drunk during childhood and went to an orphanage during childhood. Similarly to the material deprivation factor analysis, the first factor of the emotion deprivation factor analysis had an eigenvalue of 1.36, and the second factor had an eigenvalue of 0.21; therefore one factor was chosen as the optimum number of factors. The weightings resulted in possible range for the childhood emotional deprivations score was 0 – 2.2459, with both the extreme values observed, and a mean of 0.21 (sd = 0.41).

3.2.4 Descriptive results

3.2.4.1 Socioeconomic position

Childhood SEP

Childhood material deprivation

The distributions of the four variables which make up the childhood material deprivation variable can be seen in Table 3.5. Parental unemployment occurred least frequently, with only 11% of participants recalling this, and not owning a car occurred most frequently with 58% of participants recalling this.

Table 3.5: Childhood material deprivation measures

	Yes	No
Family did not own a car	4,052 (58%)	2,876 (42%)
No inside toilet	1,649 (24%)	5,238 (76%)
Financial problems	1,997 (29%)	4,903 (71%)
Parental unemployment	733 (11%)	6,119 (89%)

Of the 6,750 participants who had data for all four variables, 186 (2.8%) were in the less advantaged SEP category for all four variables (Figure 3.14), and 2,046 (30.1%) were in the more advantaged SEP category for all four variables. The modal group was being in the more advantaged SEP category for three of the four variables (34.1%). The possible range for the childhood material deprivation score was 0 – 2.3744, with both the extreme values observed, and a mean of 1.68 (sd = 0.64).

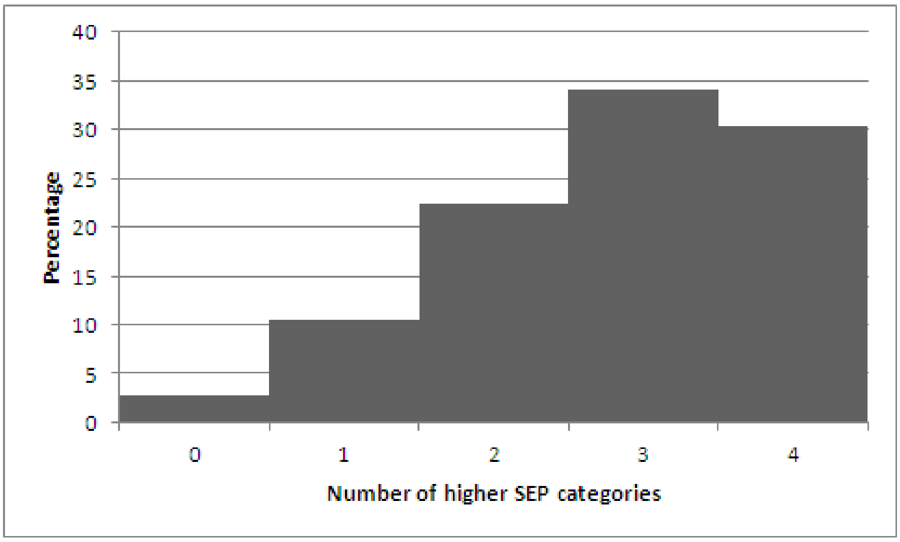


Figure 3.14: Bar chart of childhood SEP distribution: Number of higher SEP categories

<u>Father's occupational SEP</u>

The modal group for father's occupational SEP during childhood was skilled manual (33.2%), closely followed by managerial/technical (30.5%) (Figure 3.15).

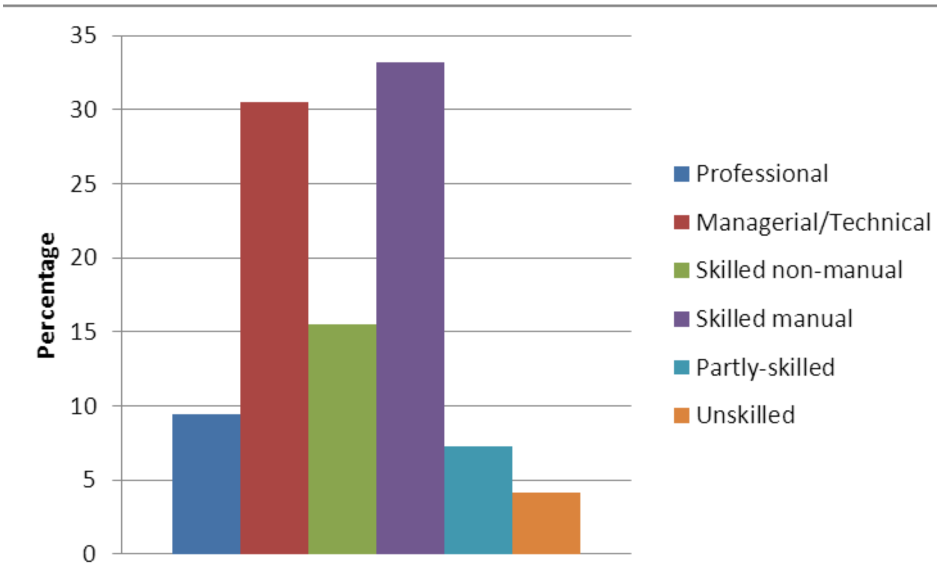


Figure 3.15: Bar chart of father's occupational SEP in childhood

Educational qualifications

The frequency of the different levels of educational qualifications can be seen in Figure 3.16. The modal category, with 47%, was the O-Level/GCSE/A-Level category, followed by degree (36%).

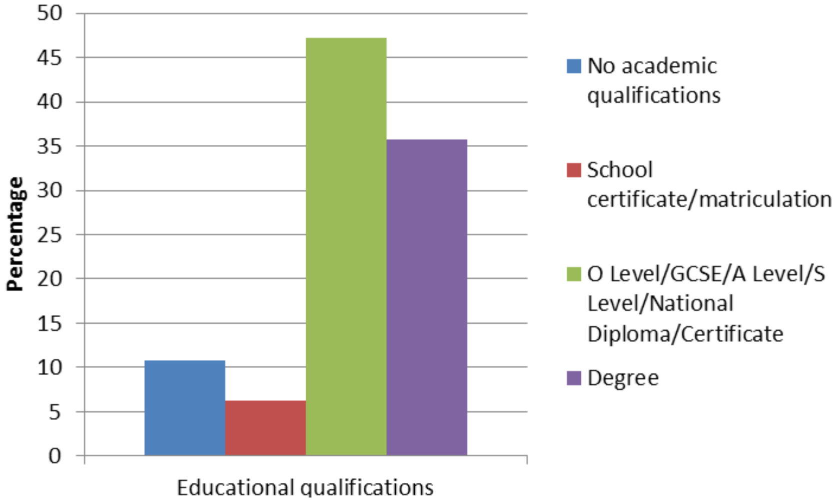


Figure 3.16: Bar chart of highest educational qualifications achieved by age 26

Occupational SEP

At phase 9 only 11.1% of participants remained in the civil service; with 19.1% working outside of the civil service (Figure 3.17). The majority, 65.4% of the participants, had retired, with the remaining participants either out of work or long-term sick.

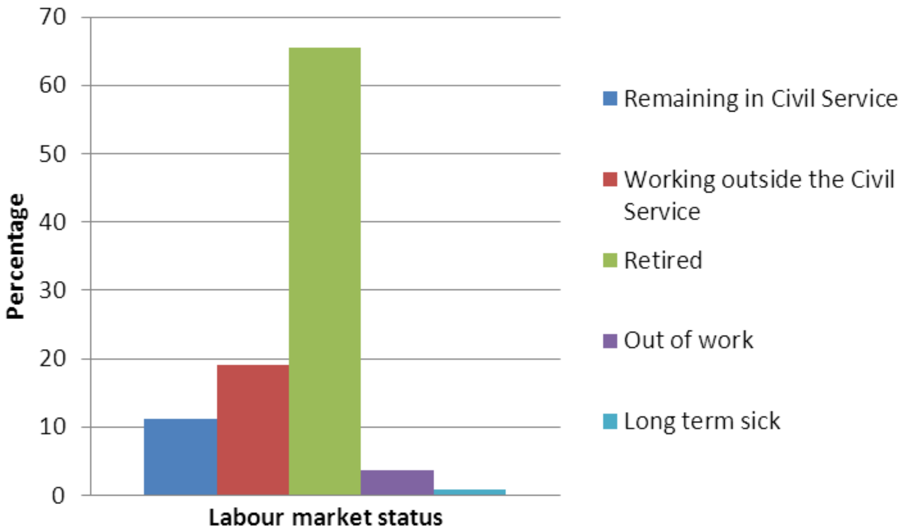


Figure 3.17: Bar chart of labour market status at phase 9

Occupational grade was categorised into three categories, which broadly correspond to Registrar General social classes I, II and III non-manual. Whilst using the last recorded occupational grade is not a perfect system, as those who leave the civil service at a younger age may have progressed to a higher grade if they had remained at the civil service, the stability of grade is fairly high for those who remained in the civil service (Table 3.6).

Figure 3.18 shows that for phase 9 participants, the modal occupational grade at phase 7 for men was Unified Grades 1-6 (55%), followed by Senior and Higher Executive Officers (41%). For women the modal category was Senior and Higher Executive Officers (48%), followed by Clerical (30%).

		phase 3			
		Unified Grades 1-6	Senior/Higher Exec	Clerical	Total
phase 1	Unified Grades 1-6	2,245 (99.6%)	10 (0.4%)	0 (0.0%)	2,255
	Senior/Higher Exec	561 (15.0%)	3,141 (84.2%)	30 (0.8%)	3,732
	Clerical	1 (0.1%)	246 (16.0%)	1,295 (84.0%)	1,542
	Total	2,807	3,397	1,325	7,529

		phase 5			
		Unified Grades 1-6	Senior/Higher Exec	Clerical	Total
phase 3	Unified Grades 1-6	1,267 (97.7%)	29 (2.2%)	1 (0.1%)	1,297
	Senior/Higher Exec	156 (9.6%)	1,428 (87.8%)	42 (2.6%)	1,626
	Clerical	5 (0.9%)	64 (11.9%)	467 (87.1%)	536
	Total	1,428	1,521	510	3,459

		phase 7			
		Unified Grades 1-6	Senior/Higher Exec	Clerical	Total
phase 5	Unified Grades 1-6	750 (93.6%)	49 (6.1%)	2 (0.2%)	801
	Senior/Higher Exec	116 (13.0%)	763 (85.3%)	15 (1.7%)	894
	Clerical	10 (4.2%)	47 (19.6%)	183 (76.3%)	240
	Total	876	859	200	1,935

		phase 9			
		Unified Grades 1-6	Senior/Higher Exec	Clerical	Total
phase 7	Unified Grades 1-6	283 (95.3%)	8 (2.7%)	6 (2.0%)	297
	Senior/Higher Exec	22 (6.7%)	293 (89.6%)	12 (3.7%)	327
	Clerical	1 (1.5%)	4 (5.9%)	63 (92.6%)	68
	Total	306	305	81	692

Table 3.6: Stability of current occupational grade

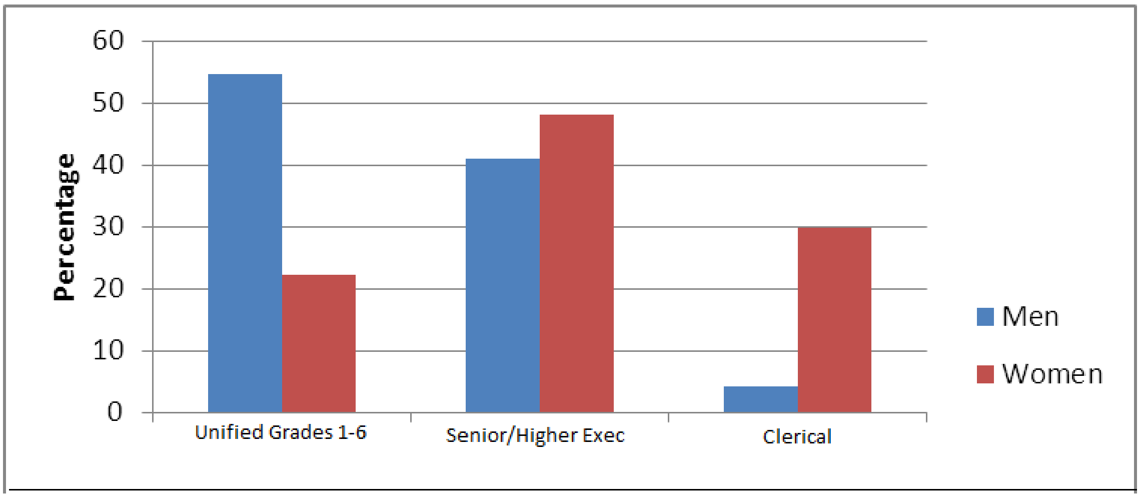


Figure 3.18: Bar chart of last recorded occupational grade at phase 7 for phase 9 participants

3.2.4.2 Cognitive function

The cognitive variables used as outcome measures were the Mill Hill Test at phase 9, and the memory scores. The Mill Hill Test score was slightly negatively skewed (Figure 3.19). A variety of transformations were assessed, including square, cubic, square root and log, but none substantially improved the normality of the distribution. The mean score was 25.2 (standard deviation: 4.33), ranging from 2 to 33. As in section 3.1.4.2, the standardised normal probability plot was not far off normally distributed (Figure 3.20).

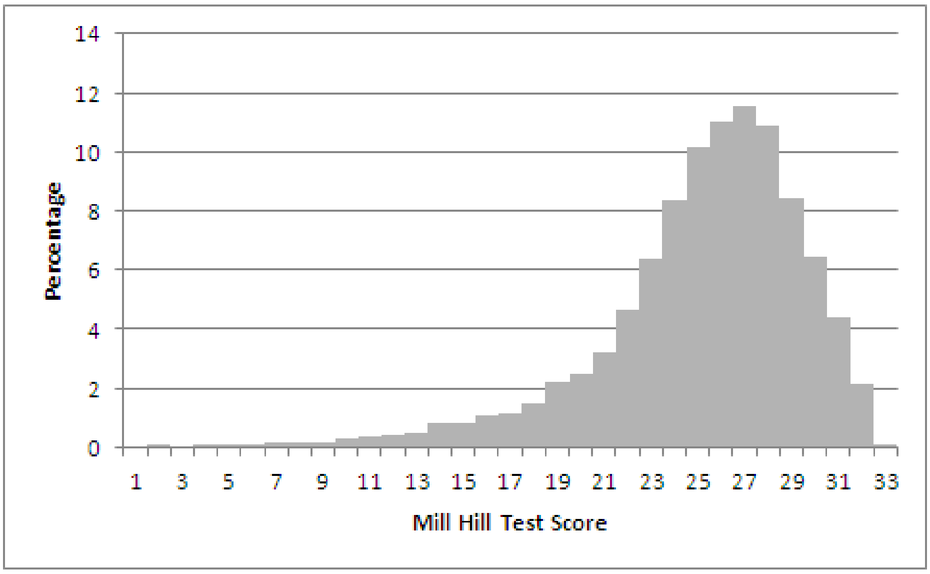


Figure 3.19: Mill Hill Test score distribution at phase 9

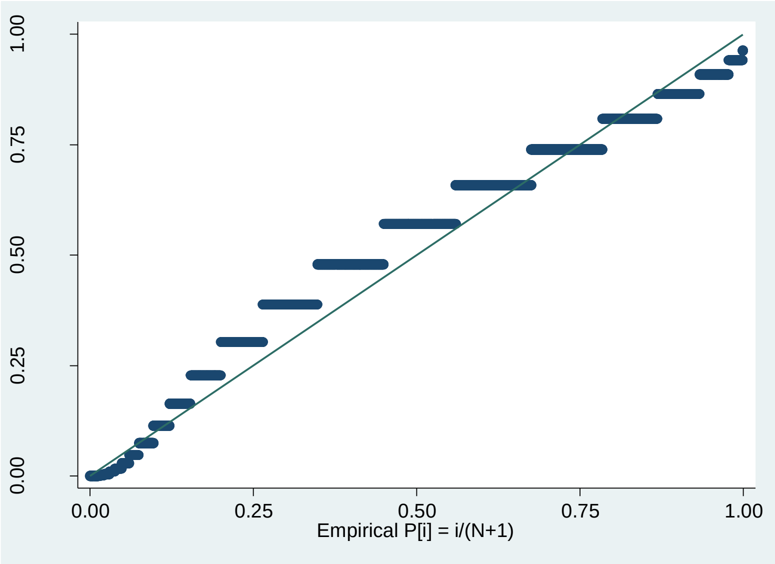


Figure 3.20: Standardised normal probability plot of the Mill Hill test

Scores for the memory tests followed approximate normal distributions, and were similar for each of phases 3, 5, 7 and 9 (Figure 3.21). The mean memory score increased between phase 3 and phase 5, then remained similar at phase 7 before dropping at phase 9 (Table 3.7).

Table 3.7: Descriptive statistics of memory scores

Memory Score	N	Mean	Std. Dev.	Min	Max
Phase 3	3430	5.86	2.25	1	16
Phase 5	6017	6.86	2.45	0	18
Phase 7	6349	6.79	2.44	1	18
Phase 9	6060	6.21	2.29	1	20

Only 3,430 participants took the memory test at phase 3. In theory this missingness could be missing completely at random, however in the Whitehall II study certain characteristics, including age, educational qualifications and occupational grade, predicted who attended the clinic earlier in the study phase and who had to be reminded, leading them to participate later. The memory scores increased from phase 3 to 5, which may be due to practice effects, as well as the limited sample at phase 3. This is discussed in section 4.3.1.

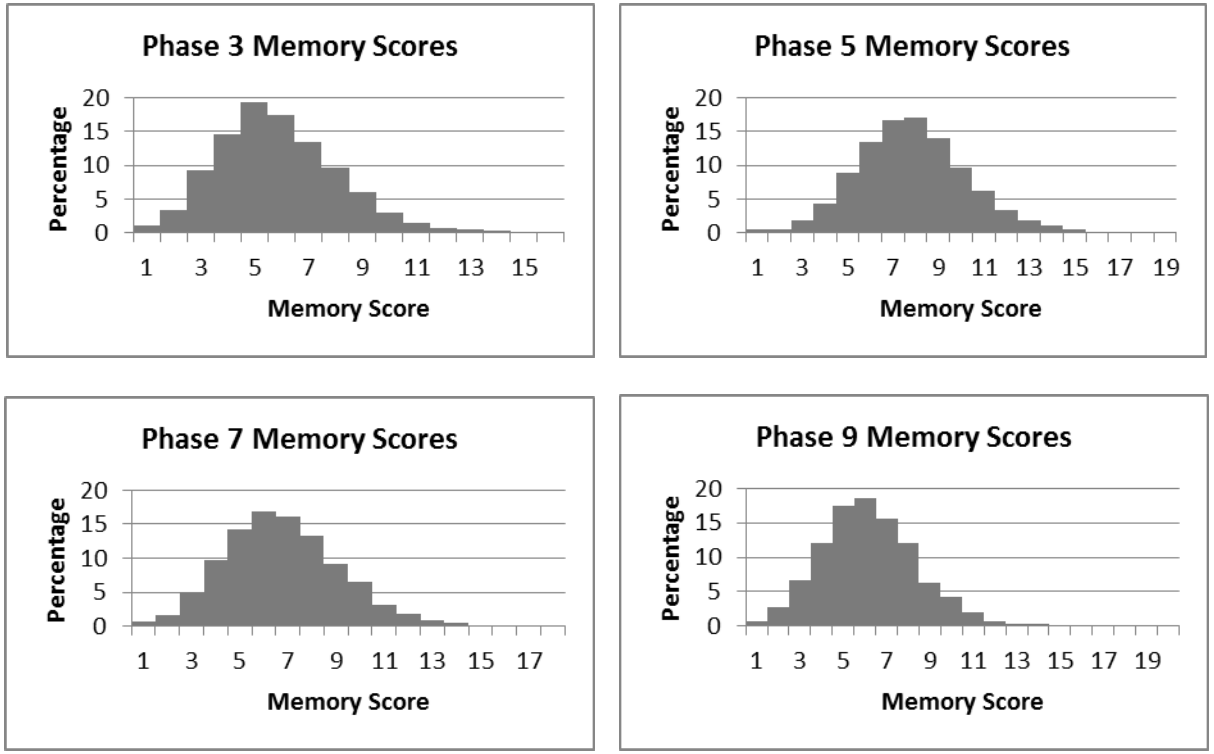


Figure 3.21: Memory score distributions at phases 3, 5, 7 and 9

3.2.5 Benefits of the Whitehall II study

The Whitehall II study is a large (N=10,308) longitudinal dataset, with 10 phases of data on social, economic, psychological and biological measures. Many of the measures have been repeated, including four repeats of the cognitive measures, allowing cognitive decline to be investigated within individuals, over a period of 15 years. The participants' ages cover 20 years at each phase, enabling cohort effects to be investigated, as well as cognitive decline over a wide age range, from 39 to 79. Although data collection started when the participants were ages 35-55, retrospective data are available on childhood characteristics.

3.3 The advantages of using both the NSHD and Whitehall II study

There are different advantages to each of the two studies, so using both studies enhances analyses. Both datasets are used to investigate the majority of aim 1; however the NSHD has data on childhood cognition which allows the effect of childhood SEP to be investigated after adjusting for childhood cognition (Aim 1 Objective 1:ii). The childhood data were collected prospectively, which is more accurate than recall data (149). However the Whitehall II dataset has four waves of repeated cognitive variables covering a wide range of ages, making it more appropriate than the NSHD for investigating aim 2, the effect of SEP on cognitive decline. Both the NSHD and Whitehall II have a wide range of variables available for imputation models and selection models. The two datasets are samples from different populations, which allow the results for aim 1 to be compared with respect to the relevant populations.

Chapter 4: The effect of childhood SEP on adult crystallized cognitive function, adjusting for later life SEP

4.1 Introduction

This chapter investigates whether there is an effect of childhood SEP on adult crystallized cognitive function, after adjusting for later life SEP, and applies missing data techniques to the analyses (Aim 1, Objective 1).

This was carried out by investigating the relationship between SEP variables at three stages of the life course and crystallized cognitive function in adulthood, in both the NSHD and Whitehall II datasets. Two childhood SEP variables were considered; father's occupational SEP and childhood material deprivation.

First the categorisation of variables for this Chapter is described, if different to the categorisations described in Chapter 3 (section 4.2). The methodology is discussed, and the topic of practice effects investigated (section 4.3), before describing the complete case analysis results for the NSHD (section 4.4.1) and Whitehall II (section 4.4.2). The issue of missing data is then addressed, first through multiple imputation (section 4.5), then Heckman selection (section 4.6), and the results compared (sections 4.7-4.9). The results are then discussed in section 4.10.

4.2 Use of variables

4.2.1 Methods

As described in Chapter 3, in the NSHD two SEP variables were selected for each of the three stages of the life course; childhood, early adulthood and mid adulthood. Occupational SEP using the Registrar General classification was used for at least one measure at each of the three stages of the life course.

Having the six categories in the model increases the complexity of the model, and leads to a smaller sample within each category, leading to larger standard errors and wider confidence intervals. However this must be balanced with the loss of information in order to simplify the models. In order to decide how many categories of each occupational SEP variable to use in the analyses, the R^2 values (Equation 4.1) were compared when the NART at age 53 was regressed on each occupational SEP variable with the six Registrar General categories and dichotomised into manual versus non-manual. If the R^2 value dropped by more than 25% between using six categories and two categories, then using three categories was considered; combining I and II, IIINM and IIM, and IV and V. If the R^2 value between the six category and the three category model dropped by more than 25%, using four categories was considered; I and II, IIINM, IIIM, and IV and V.

Equation 4.1:

$$R^2 = 1 - \frac{Residual\ Sum\ of\ Squares}{Total\ Sum\ of\ Squares} = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2}$$

where y_i is the observed outcome for individual *I*, and f_i is the fitted value of y_i.

4.2.2 Results

The R^2 values when the NART score was regressed on each of the occupational SEP variables in turn, for men and women, are reported in Table 4.1. For father's occupational SEP the percentage drop appears high, but this is due to the low R^2 value when there were 6 categories. For own occupational SEP at age 26, analysis with 4 categories resulted in very little loss of information (Table 4.2); four categories were therefore used in the following analyses. When own occupational SEP at age 43 was dichotomised, more information was lost for women than men, but this was less than 25% for both. The dichotomised variable for head of household occupational SEP had an R^2 0.02 lower than the six-category variable, but as the R^2s are both low, this represents an 18% drop.

Table 4.1: R^2 values for each model when the NART was regressed on each occupational SEP variable in turn, first with six categories, then in dichotomised form

R^2	Men			Women		
	6 categories	Dichoto-missed	% drop in R^2	6 categories	Dichoto-mised	% drop in R^2
Father's occupational SEP at age 4	0.0826	0.0679	17.80	0.1037	0.0842	18.80
Own occupational SEP at age 26	0.2231	0.2128	4.62	0.1983	0.1350	31.92
Own occupational SEP at age 43	0.2291	0.1990	13.14	0.1568	0.1213	22.64
Head of household occupational SEP at age 43	n/a	n/a	n/a	0.1176	0.0963	18.11

Table 4.2: R^2 values for each model when the NART was regressed on own occupational SEP at age 26, with three and four categories

R^2	6 categories	3 categories	% drop in R^2	6 categories	4 categories	% drop in R^2
	Men					
Own occupational SEP at age 26	0.2231	0.1538	31.06	0.2231	0.2165	2.96
	Women					
Own occupational SEP at age 26	0.1983	0.1727	12.91	0.1983	0.1974	0.45

4.3 General statistical methodology

Linear regression models were used to investigate the relationship between childhood SEP and adult crystallized cognitive function. As the likelihood ratio test should not be implemented when robust variances are used (129), which is the situation in the NSHD due to the sampling weights, the models were compared using the BIC (defined in Equation 4.2).

Equation 4.2:

BIC = *-2*ln(likelihood) + ln(N)*k* where *k* is the number of parameters estimated, and *N* is the number of observations.

The analyses were carried out separately for each gender to allow the different relationships between the SEP variables and cognitive function to be easily identified without requiring numerous interaction terms, which can make the results complicated to interpret. When a model containing SEP at each of the three time points was run for men and women together with gender interactions, using the NSHD, there were significant gender interactions with level of educational qualification (p=0.002), occupational SEP (p=0.044) and childhood cognitive function (p<0.001).

For both the NSHD and Whitehall II analyses, Model 1 contained a childhood SEP variable, Model 2 added an early adulthood SEP variable, and Model 3 an adult SEP variable. By comparing the results from Model 1 with those from Model 2 and Model 3 it was possible to see whether the earlier SEP variables had a direct effect on crystallized cognitive function after adjusting for later life SEP.

Very few additional variables were included in the models, as the aim was to investigate whether there remained an effect of childhood SEP after adjusting for measures of later life SEP. However childhood cognitive function was additionally adjusted for in the NSHD, to examine whether an effect of childhood SEP on adult crystallized cognitive function remained after adjusting for both later life SEP and childhood cognitive function, which would indicate that an effect remained that did not act through either later life SEP or childhood cognitive function. The linearity of the relationship between childhood cognitive function and crystallized cognitive function was investigated by including (standardised cognitive ability at age 8)^2 in the model.

Due to the wide age range in the Whitehall II study, each of the models was adjusted for age at phase 9. Practice effects were also investigated (section 4.3.1), as the participants had not all taken the cognitive tests the same number of times.

4.3.1 Practice effects

It is possible that practice effects (also known as learning or re-test effects) influence the scores of cognitive tests when they are repeated, as the earlier exposure to the tests may have interfered with the normal cognitive development that is being measured (150). The concern is that the improvements due to practice may hide any decline due to ageing. Between-person differences would be expected in the size of the practice effect, and this may be related to the level of cognitive function when the tests were originally taken, and the ability to absorb new information and skills. The use of parallel tests does not resolve the problem of practice effects; instead it could even complicate matters as more cognitively able and younger members improved more rapidly on the new version of the task in one study (151).

Another factor influencing practice effects is the time between the repeated tests. It is expected that practice effects would be smaller the longer the intervals between the tests, yet Rabbitt et al. (151) showed that practice effects still existed when the intervals were 2-3 years.

One option is to think of practice effects as an 'intrinsic process associated with change in processes of interest, which as such cannot be disentangled from "true change" associated with aging' (152). This would allow comparisons between individuals who have taken the same tests at the same ages. This is done in childhood development studies, where intellectual development is not separated from practice, and it has been suggested that a similar approach could also be applied in ageing studies (152).

Practice effects are only an issue in the Whitehall II study as participants repeated the same cognitive tests at phases 3, 5, 7 and 9, at approximately 5 year intervals. Repeat data, including cognitive data, were also collected on a subsample 3 months after phases 3, 7 and 9. Further, the participants had taken the tests a different number of times, which may affect the scores, and this must also be accounted for.

An individual was defined as having taken the cognitive tests if they had a score for any of the five cognitive tests at that phase. Table 4.3 shows the frequency of the number of times the cognitive tests had been taken. Of the 50 participants who took the phase 3 repeat cognitive tests, only 6 had taken the original phase 3 cognitive tests. Of the 556 who took the repeat cognitive tests at phase 7, 554 had taken the original phase 7 cognitive tests.

Table 4.3: Frequency of taking the cognitive tests in the Whitehall II dataset

Number of Times Taken Cognitive Tests	Frequency	Percentage
0	2,303	22.34
1	1,225	11.88
2	1,447	14.04
3	3,074	29.82
4	1,990	19.31
5	269	2.61
Total	10,308	100.00

The most common pattern was to have taken the cognitive tests at phases 5, 7 and 9 (24%), followed by not having taken the tests at all (22%) and having taken the tests at phases 3, 5, 7 and 9 (17%).

Initially a simple linear regression model was carried out to investigate the relationship between the number of times the cognitive tests had been taken and the Mill Hill Test score at phase 9, adjusting for age and sex, treating the number of times the cognitive tests had been taken as a linear variable (Model 1). The linearity of the number of times the cognitive tests had been taken was investigated by fitting a model first with a squared term (Model 2), then treating it as a categorical variable (Model 3).

It is possible that factors may interact with the number of times the tests had been taken, for example there may be a larger advantage to having taken the tests before for younger participants, or it may be of greater benefit to those with a better memory or higher level of educational qualifications. Therefore for the outcome Mill Hill test score at phase 9, interactions were tested between the number of times the cognitive tests had been taken and age at phase 9, sex, childhood material deprivation, educational qualifications, Mill Hill test score at phase 7, memory score at phase 9 and grouped civil service grade at phase 7 or last recorded.

In an unadjusted model (Table 4.4, Model 1), a higher number of times taking the tests predicted a higher Mill Hill test score. There was no evidence of a deviation from linearity (Model 2). The BIC was lowest for Model 1 (Table 4.5), indicating the linear relationship provided the best fit to the data.

Table 4.4: Investigating the linearity of practice effects on Mill Hill test score at phase 9 in Whitehall II, unadjusted models

Model 1 (N=6044)	**Coefficient (s.e.)**	**p-value**
Practice effect	0.42 (0.07)	<0.001
Constant	23.87 (0.22)	<0.001

Model 2 (N=6044)	**Coefficient (s.e.)**	**p-value**
Practice effect	1.00 (0.34)	0.003
Practice effect sq	-0.09 (0.05)	0.080
Constant	23.04 (0.52)	<0.001

Model 3 (N=6044)	**Coefficient (s.e.)**	**p-value**
Practice effect		<0.001
Practice effect_2*	0.83 (0.35)	0.018
Practice effect_3*	1.41 (0.33)	<0.001
Practice effect_4*	1.62 (0.33)	<0.001
Practice effect_5*	2.11 (0.41)	<0.001
Constant	23.84 (0.32)	<0.001

* Practice effect_j represents having taken the cognitive tests for the *j*th time at phase 9.

Table 4.5: Comparison of BIC for Models 1-3

	BIC
Model 1	34831.39
Model 2	34837.04
Model 3	34852.36

The effect of having taken the cognitive tests more times was reduced for those with higher Mill Hill test scores at phase 7. However this may represent a ceiling effect. The interaction with phase 7 Mill Hill test score remained significant after adjustment for age, sex, father's occupational SEP during childhood, childhood material deprivation, educational qualifications and last recorded occupational SEP at phase 7. None of the other interactions tested (using the number of times the cognitive tests had been taken) were significant.

Duff et al (153) investigated practice effects by calculating the difference between the baseline test and the one-week retest score. This variable was then included in a regression model, with the one-year score as the dependent variable, and the baseline score as an independent variable. The significance of the practice effect was then examined.

Only 6 participants had repeat cognitive data at phase 3 and phase 3 repeat, so to investigate practice effects using the methodology of Duff et al (153), repeat cognitive data were used from phase 7. The impact of both the original Mill Hill test score at phase 7 and the difference, defined as the practice effect, between the repeat score (taken three months later) and original score were used as independent variables, with phase 9 Mill Hill test score used as the dependent variable. The correlation between the practice effect and the original score was -0.273, suggesting that those who originally had higher scores had more negative differences between their phase 7 and phase 7 repeat scores. The practice effects were significantly associated with the phase 9 Mill Hill test. The interaction between Mill Hill test score at phase 7 and the practice effect was also significant, with a small negative coefficient. This implies that practice effects do exist (Table 4.6), both before adjustment (Model 1) and after adjusting for sex, age, father's occupational SEP, childhood material deprivation, educational qualifications and employment grade (Model 2).

Table 4.6: Modelling practice effects with phase 9 Mill Hill score as outcome in the Whitehall II dataset

N=378	Model 1 Coefficient (s.e.)	Model 1 p-value	Model 2 Coefficient (s.e.)	Model 2 p-value
Phase 7 Mill Hill Test Score	0.93 (0.03)	<0.001	0.89 (0.03)	<0.001
Phase 7 Practice Effect	1.06 (0.22)	<0.001	0.98 (0.22)	<0.001
Phase 7 MH*Practice Effect	-0.02 (0.01)	0.018	-0.02 (0.01)	0.044
Female			-0.02 (0.35)	0.948
Phase 9 Age			-0.00 (0.02)	0.829
Father's occupational SEP			0.13 (0.20)	0.503
Childhood material deprivation			-0.04 (0.15)	0.819
Educational Qualifications:				
Baseline – no qualifications				
School certificate			-0.06 (0.58)	0.917
O-Level/A-Level			0.31 (0.42)	0.467
Degree			0.38 (0.44)	0.392
Phase 7 Occupation:				
Baseline – Unified Grades 1-6				
Senior and Higher Executive Officers			-0.23 (0.22)	0.286
Clerical			-0.53 (0.48)	0.276
Constant	1.75 (0.72)	0.015	2.76 (1.56)	0.079

Ideally a practice effect would be calculated for everyone and included as a covariate in the analyses (as in Duff et al (153)). However due to the study design it was not possible to calculate a practice effect for everyone, and restricting the sample to those who have a practice effect would considerably reduce the size of the dataset. The practice effect was, therefore, partially accounted for in the Whitehall II analyses by adjusting for the number of times the cognitive tests had previously been taken.

4.4 Complete case results

The main area of interest was whether an effect of childhood SEP on crystallized cognitive function in adulthood existed after adjusting for later life SEP. Adjustment for later life SEP may be considered an overadjustment by some (154), and this is discussed in section 4.10. The tables for results using father's occupational SEP as the measure of childhood SEP, educational qualifications as the measure of early adulthood SEP and own occupational SEP at age 43 can be found below; the remaining tables investigating other measures of SEP are in Appendix 7. Each model of the model development process within a table contains the same sample.

4.4.1 NSHD

4.4.1.1 Childhood SEP measured using father's occupational SEP

In the crude model, Model 1, childhood SEP had a similar effect size for women (6.41 (95% CI: 5.10, 7.22)) and men (6.16 (95% CI: 4.77, 7.56)). When adjusting for educational qualifications (Table 4.7 (men) and Table 4.8 (women)), childhood SEP remained a significant predictor of NART score at age 53. Additional adjustment for own occupation SEP at age 43 (Model 3) further attenuated the effect size for men (from 2.89 (95% CI: 1.58, 4.21) to 2.09 (95% CI: 0.80, 3.37)), whereas for women the effect size was virtually unchanged (from 2.33 (95% CI: 1.16, 3.51) to 2.31 (95% CI: 1.15, 3.47)).

For men, further adjustment for childhood cognitive function (Model 4) resulted in childhood SEP becoming non-significant. There was no evidence on non-linearity between cognitive function and the NART (Model 5).

For women, childhood SEP remained a significant predictor of NART score after adjusting for childhood cognitive function (Models 4 and 5); however own occupational SEP at age 43 was fully attenuated. As mentioned above (section 4.3), it was not possible to carry out a likelihood ratio test as robust standard errors were used, due to the sample weights. As the quadratic term was significant in Model 5, Model 5 was accepted as the final model, despite the slightly increased BIC value.

When occupational SEP was used as the measure of early adulthood SEP, the results were very similar for men (Appendix 7: Table A7.3), whereas for women the effect of childhood SEP was attenuated to a lesser extent than when educational qualifications were used (Appendix 7: Table A7.4). Own occupational SEP at age 43 also remained a significant predictor of NART score for women in Model 5, unlike when educational qualifications were used. For women the effect of childhood SEP was only slightly different when using head of household occupational SEP as the measure of adult SEP (Appendix 7: Table A7.5 and Appendix 7: Table A7.6).

4.4.1.2 Childhood SEP measured using lack of household amenities at age 2

In the unadjusted model, Model 1, having access to each additional amenity increased the NART score at age 53 for men and women. For both men and women, adjusting for educational qualifications fully attenuated the effect of childhood SEP (Model 2), but both educational qualifications and own occupational SEP were significant in Model 3. As when father's occupational SEP was used as the measure of childhood SEP, own occupational SEP at age 43 was fully attenuated for women by adjustment for childhood cognitive function (Model 4). Childhood cognitive function was significant in Model 4, but there was no evidence that the relationship was not linear for men (Model 5). For women there was a non-linear relationship between childhood cognitive function and the NART score (Model 5), as when father's occupational SEP was used as the measure of childhood SEP.

When occupational SEP was used as the measure of early adulthood SEP, childhood SEP remained a significant predictor of NART score after adjusting for early adulthood SEP (Model 2), although the effect was fully attenuated by the additional adjustment for adult SEP (Model 3) (Appendix 7: Table A7.7). For women, adult SEP was a significant predictor of NART score (Appendix 7: Table A.7.8), unlike when educational qualifications were used as the early adulthood measure of SEP. Head of household occupational SEP was significant for both measures of early adult SEP (Appendix 7: Table A7.9 and Appendix 7: Table A7.10).

4.4.2 Whitehall II

4.4.2.1 Childhood SEP measured using father's occupational SEP

In Model 1 childhood SEP was a significant predictor of Mill Hill test score at phase 9 for both men (Table 4.9) (0.91 (95% CI: 0.62, 1.20)) and women (Table 4.10) (2.06 (95% CI: 1.41, 2.71), after adjusting for age at phase 9 and the number of times the cognitive tests had been taken.

When educational qualifications were adjusted for childhood SEP was partially attenuated for men (non-manual childhood SEP: 0.51 (95% CI: 0.24, 0.78)) and fully attenuated for women (non-manual childhood SEP: 0.51 (95% CI: -0.14, 1.16)).

The coefficient for men was slightly further attenuated by the addition of adult SEP to the model (Model 3), but an effect of childhood SEP remained on Mill Hill test score at phase 9. All three SEP measures were significant predictors of Mill Hill test score at phase 9 for men, and educational qualifications and adult occupational SEP were significant predictors for women.

4.4.2.2 Childhood SEP measured using childhood material deprivation

In Model 1, when only childhood SEP, age and the number of times the cognitive tests had been taken were included in the model, childhood SEP was a significant predictor for both men (Appendix 7: Table A7.11) (0.43 (95% CI: 0.19, 0.67)) and women (Appendix 7: Table A7.12) (0.77 (95% CI: 0.51, 1.28)).

When educational qualifications were adjusted for (Model 2), childhood SEP was fully attenuated for women (0.19 (95% CI: -0.28, 0.66)), but remained significant for men (0.25 (95% CI: 0.03, 0.47)). Educational qualifications were significant for both men and women, with similar coefficients to when father's occupational SEP was used as the measure of childhood SEP, and similar R^2 values. There remained an effect of childhood SEP for men in Model 3, after additionally adjusting for occupational SEP, with very little difference between the results when father's occupational SEP was used and when childhood material deprivation was used.

To summarize, the same overall conclusions were drawn for both measures of childhood SEP. For men there remained an effect of childhood SEP on adult crystallized cognitive function after adjusting for later life SEP, however for women the effect was fully attenuated by the addition of educational qualifications.

Table 4.7: NSHD complete case model development for men, with father's occupational SEP as the measure of childhood SEP, and the outcome NART at age 53

MEN (N=893)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Father's occupational SEP - manual										
Father's occ. SEP - non-manual	6.16 (0.71)	<0.001	2.89 (0.67)	<0.001	2.09 (0.65)	0.001	0.61 (0.62)	0.321	0.64 (0.62)	0.300
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			5.75 (0.98)	<0.001	4.46 (1.03)	<0.001	2.48 (0.88)	0.005	2.36 (0.89)	0.008
Education - GCE 'A'-Level			7.58 (0.87)	<0.001	5.98 (0.94)	<0.001	3.26 (0.83)	<0.001	3.22 (0.83)	<0.001
Education - Degree			13.47 (0.84)	<0.001	10.94 (0.97)	<0.001	6.24 (0.95)	<0.001	6.35 (0.94)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					4.29 (0.84)	<0.001	2.94 (0.73)	<0.001	2.90 (0.73)	<0.001
Standardised age 8 cognitive score							4.34 (0.34)	<0.001	4.24 (0.31)	<0.001
Standardised age 8 cognitive score squared									-0.29 (0.23)	0.205
Constant	31.83 (0.47)	<0.001	27.73 (0.63)	<0.001	26.49 (0.67)	<0.001	30.08 (0.63)	<0.001	30.40 (0.70)	<0.001
Model fit										
R-squared	0.0796		0.2735		0.3104		0.4553		0.4566	
BIC	6513.041		6322.211		6282.421		6078.640		6083.1662	

Table 4.8: NSHD complete case model development for women, with father's occupational SEP as the measure of childhood SEP, and the outcome NART at age 53

WOMEN (N=955)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Father's occupational SEP - manual										
Father's occ. SEP - non-manual	6.41 (0.67)	<0.001	2.33 (0.60)	<0.001	2.31 (0.59)	<0.001	1.08 (0.49)	0.027	1.11 (0.48)	0.021
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			7.85 (0.70)	<0.001	7.04 (0.70)	<0.001	4.03 (0.64)	<0.001	3.95 (0.65)	<0.001
Education - GCE 'A'-Level			12.34 (0.74)	<0.001	11.12 (0.77)	<0.001	6.75 (0.70)	<0.001	6.80 (0.70)	<0.001
Education - Degree			17.15 (1.05)	<0.001	15.79 (1.09)	<0.001	7.83 (1.19)	<0.001	8.53 (1.16)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.78 (0.74)	<0.001	0.98 (0.65)	0.132	0.89 (0.66)	0.177
Standardised age 8 cognitive score							4.86 (0.31)	<0.001	4.73 (0.28)	<0.001
Standardised age 8 cognitive score squared									-0.43 (0.18)	0.018
Constant	30.88 (0.44)	<0.001	26.60 (0.53)	<0.001	25.19 (0.66)	<0.001	29.50 (0.61)	<0.001	29.95 (0.67)	<0.001
Model fit										
R-squared	0.0858		0.3607		0.3756		0.5398		0.5426	
BIC	6946.363		6625.456		6609.713		6325.201		6326.124	

Table 4.9: Whitehall II complete case model development for men, with father's occupational SEP as the measure of childhood SEP, and the outcome Mill Hill score at phase 9

Men (N=2,440)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e)	p-value
Baseline: Father's occupational SEP in childhood – manual						
Father's occ. SEP - non-manual	0.91 (0.15)	<0.001	0.51 (0.14)	<0.001	0.43 (0.13)	0.001
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			2.66 (0.44)	<0.001	2.04 (0.42)	<0.001
Education: O-Level/A-Level/National diploma/Certificate			3.15 (0.30)	<0.001	2.41 (0.30)	<0.001
Education: University degree			4.85 (0.31)	<0.001	3.47 (0.31)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					2.63 (0.38)	<0.001
Last occ. grade - Unified Grades 1-6					4.34 (0.38)	<0.001
Age (phase 9)	-0.00 (0.01)	0.921	0.03 (0.01)	0.017	0.01 (0.01)	0.258
No. of times taken cognitive tests	0.38 (0.09)	<0.001	0.25 (0.09)	0.004	0.20 (0.08)	0.017
Constant	23.92 (0.89)	<0.001	19.00 (0.93)	<0.001	17.73 (0.93)	<0.001
Model fit						
R-squared	0.0212		0.1293		0.2018	
BIC	13144.44		12879.33		12680.94	

Table 4.10: Whitehall II complete case model development for women, with father's occupational SEP as the measure of childhood SEP, and the outcome Mill Hill score at phase 9

Women (N=826)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e)	p-value
Baseline: Father's occupational SEP in childhood – manual						
Father's occ. SEP - non-manual	2.06 (0.33)	<0.001	0.51 (0.33)	0.123	0.13 (0.31)	0.668
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			0.84 (0.72)	0.239	0.19 (0.67)	0.772
Education: O-Level/A-Level/National diploma/Certificate			3.43 (0.44)	<0.001	2.02 (0.43)	<0.001
Education: University degree			6.08 (0.52)	<0.001	3.38 (0.54)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					3.18 (0.37)	<0.001
Last occ. grade - Unified Grades 1-6					5.54 (0.48)	<0.001
Age (phase 9)	-0.14 (0.03)	<0.001	-0.02 (0.03)	0.448	0.02 (0.03)	0.554
No. of times taken cognitive tests	0.15 (0.24)	0.525	0.35 (0.22)	0.105	0.27 (0.20)	0.178
Constant	31.29 (2.06)	<0.001	20.57 (2.13)	<0.001	17.20 (2.00)	<0.001
Model fit						
R-squared	0.0778		0.2127		0.3235	
BIC	4914.89		4801.33		4687.51	

4.5 Multiple imputation methodology

The next stage was to carry out the analysis using multiple imputation, which allows for MAR missingness. The general methodology to carry out multiple imputation was described in section 2.4.1. The main decision required for multiple imputation analyses involves choosing the variables to include in the imputation model. This section provides specific details of the process used to choose the variables. As the analyses were carried out separately for men and women, and gender was fully observed, multiple imputation was carried out separately for each gender. This meant that the different relationships between the variables for men and women could be captured without including a large number of sex interactions. The same variables were considered for inclusion in the imputation model for each sex, but the final imputation models were not required to be the same.

The first step was to include all the variables which appeared in the model of interest, identified above. In the NSHD and Whitehall II analyses there is more than one model of interest, allowing for the different SEP variables at the three time points, but the same imputation model was used, including the variables from all the models of interest.

The second step involved identifying the variables which were predictive of missing data. This was done by creating a binary variable indicating whether the outcome measure in the analysis of interest was missing or observed, and then carrying out logistic regression to investigate which variables were predictive of not having an observed outcome measure. Similar analyses were repeated to identify variables which predicted missingness of SEP at each of the three timepoints, and any other variables included in the model of interest. To be included in the group of variables to be considered at step 3, it was only necessary to be predictive of missingness for one variable. Only variables from earlier phases were tested to investigate whether they were predictive of missingness, as the majority of the missingness in these analyses was monotone (72% NSHD, 57% Whitehall II).

Although there are numerous studies that have investigated predictors of dropout, most of these studies focus on specific populations, such as those in treatment for depression (155) or alcoholism (156). However in longitudinal cohort studies low cognitive scores have been predictive of increased dropout (157), therefore cognitive measures from earlier in the study were considered for the imputation model. Another birth cohort study found that those from a working class background were more likely to drop out (158), so those childhood SEP variables that were not included in the model of interest were also considered for the imputation model. The same study also found that more hostile and withdrawn behaviour were associated with dropout, therefore measures such as truancy, whether the participant was difficult to discipline and whether the teacher considered the participant to be a daredevil were considered for the imputation model in the NSHD, where childhood variables were available. As the variables were all tested before they were included in the imputation model, a large range of other variables was considered where there were plausible relationships between the variable and dropout, for example diagnosis of cancer, job involving travel away from home, state of health and isolation score. The full list of variables considered for each study can be found in Table 4.11 (NSHD) and Appendix 8: Table A8.1 (Whitehall II).

The third step involved checking whether the variables identified in step two were associated with the variables in the model of interest. This was done by carrying out regression models to investigate the association between each variable identified in step two and the variables in the model of interest. Linear regression was used for the continuous variables in the model of interest, logistic regression for the binary variables, and ordered logistic regression for the ordered categorical variables. Only variables from step two that were significant predictors of at least one of the variables in the model of interest were retained in the imputation model. The variables identified in the second step were allowed to predict the values of variables from earlier waves, even though they were not allowed to predict missingness at earlier waves in the second step.

Step four aimed to identify auxiliary variables that may be associated with any of the variables in the imputation model, in order to improve the precision of the imputations. As there were already a large number of variables in the imputation model, the variables in step four were only considered if there was a reason to think they would improve the imputation of specific variables. Therefore, only those variables where a relationship was hypothesised were investigated in this step.

4.5.1 Developing an imputation model for the NSHD

The variables that were included in the models of interest, and were therefore also included in the imputation model, were NART, father’s occupational SEP, household amenities at age 2, educational qualifications achieved by age 26, own occupational SEP at age 26, own occupational SEP at age 43, and, for women, head of household SEP at age 43. The weighting variable was also included.

Standardised cognitive function at age 8 and its squared term were in the models of interest. Cognitive function at age 8 was calculated from 4 cognitive function variables, as described in section 3.1.4.2. Some participants had scores for some of these tests but not others, so they did not have a standardised cognitive function score. Therefore the four cognitive test scores were imputed rather than the standardised cognitive function score. The standardised cognitive score was then calculated in the same way as in the complete case analyses, so that the same values for the standardized cognitive scores would be recorded for those individuals in both the complete case dataset and multiple imputation dataset. The standardized variables were then summed, and the complete case mean was subtracted, and divided by the complete case standard deviation, to ensure the mean and standard deviation of the complete case data were zero; this would not necessarily be the case with the imputed data though. The (standardised cognitive function)^2 term was included in the imputation model; had it not been included, this would assume that there was no relationship between (standardised cognitive function)^2 and the outcome of interest, the NART. As the quadratic term was included in the imputation model, the (standardised cognitive function)^2 variable may not equal the squared value of the standardised cognitive function score for those values that were imputed; however as described in section 2.4.1.3, constraining the squared variable to be equal to the squared value may introduce bias.

The variables which were considered for the NSHD imputation model are in Table 4.11. The results from the tests of whether variables were associated with the missingness of variables in the model of interest can be seen in Table 4.11, while Table 4.12 and Table 4.13 show which of the variables from Table 4.11 were predictive of the values of the variables in the model of interest. Very few of the variables that were predictive of missingness were not associated with any of the variables in the model of interest: for women only the teacher’s assessment of whether the child was a daredevil at age 13; and for men attending church or religious activities, the Present State Examination total score and how confident you feel in yourself, all at age 36.

The auxiliary variables considered and the variables that were hypothesised to be associated with the auxiliary variables are in Table 4.14, along with whether they were significant predictors. This led to the final imputation models for women (Table 4.15) and men (Table 4.16).

Once the main imputation model had been developed, other imputation models were compared, to investigate the influence of the imputation model on the findings for the effect of childhood SEP on adult crystallized cognitive function. Two smaller imputation models were compared to the full imputation model (imputation model 1) for women in the NSHD. Imputation model 2 consisted of the variables in imputation model 1, without the auxiliary variables in Table 4.15. Imputation model 3 contained a much smaller group of variables, which were either contained in the model of interest or considered to be strongly related to at least one of the included variables (verbal memory at age 43, crowding at age 4, occupational social class at age 36, cognitive tests at age 11 and marital status at age 36).

The regression results for comparing imputation models are in Table 4.17, for Model 5 from Table 4.8. The results of the three models were very similar, with the same variables significant for each of the imputation models, at a similar level of significance, with similar effect sizes. The largest imputation was used, as including additional variables in an imputation model is ‘at worst neutral, and at best extremely beneficial’ (121), implying that it does not make the imputations worse, but may improve them.

The effect of the number of imputations was also investigated, using the largest imputation model. As there was no reason to expect the effect of the number of imputations used to differ between men and women, this was only carried out for women. The results from 5, 10 and 20 imputations were compared, as 5-10 imputations are often recommended (124).

The regression results comparing the number of imputations are in Table 4.18. The main difference between the numbers of imputations occurs for adult occupational SEP, where adult occupational SEP was significant with 5 imputations, but not 10 or 20 imputations. The result using more imputations was assumed to be more accurate.

Twenty imputations were used in all future multiple imputation analyses, given these differences in results, as using the larger number of imputations increases the precision of the results (125).

Table 4.11: Step 2 of choosing an imputation model: Identifying variables for the NSHD imputation model

Age variable collected		Significant predictor of				
		missing NART score		missing own occupational SEP at age 43		missing head of household occupational SEP at age 43
		men	**women**	**men**	**women**	**women**
4	Age of dwelling	x	x	x	✔	x
4	Crowding	x	✔	✔	✔	x
4	Clothes repair	✔	x	x	x	x
4	Yard or garden	✔	x	✔	✔	x
4	Cleanliness of child	✔	✔	✔	✔	✔
4	Cleanliness of house	✔	✔	✔	✔	✔
4	Mother's management and understanding of the child	✔	✔	✔	✔	✔
4	Child's shoes	✔	✔	✔	x	x
4	Child has own bed	x	x	x	x	x
4	Repair of dwelling	x	x	x	x	x
10	Difficult to discipline	x	x	x	x	✔
11	Father's occupational SEP	✔	x	x	x	x
11	Reading score	✔	✔	✔	✔	✔
11	Non-verbal reasoning score	✔	✔	✔	✔	✔
11	Arithmetic score	✔	✔	✔	✔	✔
11	Verbal reasoning score	✔	✔	✔	✔	✔
11	Vocabulary score	✔	✔	✔	✔	✔
11	Home amenities	x	x	x	x	x
13	Have to stick up for myself	✔	x	x	x	x
13	Daredevil (teacher)	✔	x	x	✔	x
13	Get angry about nothing	x	x	x	x	x

Table 4.15 continued

Age variable collected		Significant predictor of				
		missing NART score		missing own occupational SEP at age 43		missing head of household occupational SEP at age 43
		men	**women**	**men**	**women**	**women**
13	It's usually safer to do things alone	x	x	x	x	x
15	Truancy during past year	✔	x	✔	x	x
15	Reading score	✔	✔	✔	✔	✔
15	Non-verbal reasoning score	✔	✔	✔	✔	✔
15	Mathematics score	✔	✔	✔	✔	✔
15	Verbal reasoning score	✔	✔	✔	✔	✔
16	Is there anything which causes a lot of worry	x	✔	x	✔	✔
26	Brought up in any faith or religious denomination	✔	x	✔	x	x
26	Currently have a religion	✔	x	x	x	x
26	Job stress	✔	✔	✔	x	x
26	Happiness	x	x	x	x	x
26	Parents ever divorced or separated	x	x	x	x	x
31	Any hospital admissions	x	x	x	✔	✔
36	How confident do you feel in yourself	✔	x	✔	x	x
36	Present State Examination total score	✔	x	✔	x	✔
36	Attend church or religious activities	x	x	✔	x	x
36	Marital status	✔	x	✔	✔	✔
36	Ever been unemployed	x	x	✔	✔	x

Table 4.15 continued

year variable collected		Significant predictor of				
		missing NART score		missing own occupational SEP at age 43		missing head of household occupational SEP at age 43
		men	**women**	**men**	**Women**	**women**
36	Current smoking status	✔	✔	✔	✔	✔
36	Does religious upbringing have effect on life now	✔	✔	x	x	✔
36	Ever lived abroad	x	x	x	x	x
36	Ever had cancer	x	x	x	omitted*	omitted*
36	Any hospital admissions	x	x	x	x	x
36	Felt depressed in the last year	x	x	x	x	x
36	Any bad news in the past year	x	x	x	x	x
36	Blood pressure	x	x	x	✔	✔
43	Do you think that you have friends or neighbours or relatives who would help you?	✔	x			
43	How many friends or relatives could you visit at any time without waiting for an invite	✔	x			
43	Verbal memory score	✔	✔			
43	Ever had cancer	x	x			
43	Serious illness in last year	x	x			
43	Search speed score	x	x			
43	Been in hospital since last time asked	x	x			

* all cases had the same outcome

Table 4.15 continued

year variable collected		Significant predictor of				Significant predictor of	
		missing own occupational SEP at age 26		educational qualifications attained by age 26		missing cognitive score at age 8	
		men	**women**	**men**	**women**	**men**	**women**
4	Age of dwelling	x	x	x	x	x	x
4	Crowding	✔	x	x	x	x	✔
4	Clothes repair	x	x	x	x	x	x
4	Yard or garden	x	x	✔	x	x	✔
4	Cleanliness of child	x	x	x	x	x	x
4	Cleanliness of house	x	x	x	x	x	x
4	Mother's management and understanding of the child	✔	x	✔	x	x	x
4	Child's shoes	x	x	x	x	x	x
4	Repair of dwelling	x	x	x	x	x	x
10	Difficult to discipline	x	✔	x	x		
11	Father's occupational SEP	x	x	x	x		
11	Reading score	x	x	x	x		
11	Non-verbal reasoning score	x	✔	x	✔		
11	Arithmetic score	✔	x	✔	✔		
11	Verbal reasoning score	x	x	x	✔		
11	Vocabulary score	x	x	x	x		
11	Home amenities	x	x	x	x		
13	Have to stick up for myself	x	x	x	x		
13	Daredevil (teacher)	x	x	x	x		
13	Get angry about nothing	x	x	x	x		
13	It's usually safer to do things alone	x	x	x	x		

Table 4.15 continued

year variable collected		Significant predictor of			
		missing own occupational SEP at age 26		educational qualifications attained by age 26	
		men	**women**	**men**	**women**
15	Truancy during past year	x	x	x	x
15	Reading score	x	✔	x	x
15	Non-verbal reasoning score	x	x	x	x
15	Mathematics score	x	✔	✔	x
15	Verbal reasoning score	x	✔	x	x
16	Is there anything which causes a lot of worry	x	x	x	x

Table 4.12: Step 3 of identifying variables for the NSHD imputation model - men

		Significantly associated with						
Age		NART	Cognitive score at age 8	Educational qualifications attained by age 26	Household amenities at age 2	Own occupational SEP at age 26	Own occupational SEP at age 43	Father's occupational SEP at age 4
4	Crowding	✔	✔	✔	✔	✔	✔	✔
4	Clothes repair	✔	✔	✔	✔	✔	x	✔
4	Yard or garden	x	x	✔	✔	x	✔	✔
4	Cleanliness of child	✔	✔	✔	✔	✔	✔	✔
4	Cleanliness of house	✔	✔	✔	✔	✔	✔	✔
4	Mother's management and understanding of the child	✔	✔	✔	✔	✔	✔	✔
4	Child's shoes	✔	✔	✔	✔	✔	x	✔
4	Child has own bed	✔	✔	✔	✔	✔	✔	✔
10	Difficult to discipline	✔	x	✔	x	x	✔	x
11	Father's occupational SEP	✔	✔	✔	✔	✔	✔	✔
11	Reading score	✔	✔	✔	✔	✔	✔	✔

Table 4.12 continued

Age		NART	Cognitive score at age 8	Educational qualifications attained by age 26	Household amenities at age 2	Own occupational SEP at age 26	Own occupational SEP at age 43	Father's occupational SEP at age 4
11	Non-verbal reasoning score	✔	✔	✔	✔	✔	✔	✔
11	Arithmetic score	✔	✔	✔	✔	✔	✔	✔
11	Verbal reasoning score	✔	✔	✔	✔	✔	✔	✔
11	Vocabulary score	✔	✔	✔	✔	✔	✔	✔
13	Daredevil (teacher)	✔	x	✔	x	x	x	x
13	Have to stick up for myself	✔	✔	✔	✔	✔	✔	✔
15	Truancy during past year	x	✔	✔	✔	✔	x	✔
15	Reading score	✔	✔	✔	✔	✔	✔	✔
15	Non-verbal reasoning score	✔	✔	✔	✔	✔	✔	✔
15	Mathematics score	✔	✔	✔	✔	✔	✔	✔
15	Verbal reasoning score	✔	✔	✔	✔	✔	✔	✔
26	Brought up in any faith or religious denomination	✔	✔	✔	x	✔	✔	✔

Table 4.12 continued

Age		NART	Cognitive score at age 8	Educational qualifications attained by age 26	Household amenities at age 2	Own occupational SEP at age 26	Own occupational SEP at age 43	Father's occupational SEP at age 4
26	Currently have a religion	✔	x	x	x	x	✔	✔
26	Job stress	✔	✔	✔	x	✔	✔	✔
36	How confident do you feel in yourself	x	x	x	x	x	x	x
36	Present State Examination total score	x	x	x	x	x	x	x
36	Attend church or religious activities	x	x	x	x	x	x	x
36	Marital status	x	x	✔	x	✔	x	x
36	Ever been unemployed	✔	✔	✔	x	✔	✔	✔
36	Current smoking status	✔	x	✔	✔	✔	✔	✔
36	Does religious upbringing have effect on life now	✔	✔	✔	✔	✔	✔	✔

Table 4.12 continued

Age		NART	Cognitive score at age 8	Educational qualifications attained by age 26	Household amenities at age 2	Own occupational SEP at age 26	Own occupational SEP at age 43	Father's occupational SEP at age 4
43	How many friends or relatives could you visit at any time without waiting for an invite	x	x	x	✔	x	x	x
43	Do you think that you have friends or neighbours or relatives who would help you?	x	x	x	x	x	x	x
43	Verbal memory score	✔	✔	✔	✔	✔	✔	✔

Table 4.13: Step 3 of identifying variables for the NSHD imputation model - women

Age		NART	Cognitive score at age 8	Educational qualifications attained by age 26	Household amenities at age 2	Own occupational SEP at age 26	Own occupational SEP at age 43	Head of household occupational SEP at age 43	Father's occupational SEP at age 4
4	Age of dwelling	x	x	✔	✔	✔	x	x	✔
4	Crowding	✔	✔	✔	✔	✔	✔	✔	✔
4	Yard or garden	✔	x	✔	✔	✔	✔	✔	✔
4	Cleanliness of child	✔	✔	✔	✔	✔	✔	✔	✔
4	Cleanliness of house	✔	✔	✔	✔	✔	✔	✔	✔
4	Mother's management and understanding of the child	✔	✔	✔	✔	✔	✔	✔	✔
4	Child's shoes	✔	✔	✔	x	✔	✔	✔	✔
4	Child has own bed	✔	✔	✔	✔	✔	✔	✔	✔
10	Difficult to discipline	x	x	✔	x	✔	x	x	x
11	Reading score	✔	✔	✔	✔	✔	✔	✔	✔
11	Non-verbal reasoning score	✔	✔	✔	✔	✔	✔	✔	✔
11	Arithmetic score	✔	✔	✔	✔	✔	✔	✔	✔
11	Verbal reasoning score	✔	✔	✔	✔	✔	✔	✔	✔
11	Vocabulary score	✔	✔	✔	✔	✔	✔	✔	✔
13	Daredevil (teacher)	x	x	x	x	X	x	x	x

Table 4.13 continued

Age		NART	Cognitive score at age 8	Educational qualifications attained by age 26	Household amenities at age 2	Own occupational SEP at age 26	Own occupational SEP at age 43	Head of household occupational SEP at age 43	Father's occupational SEP at age 4
15	Reading score	✓	✓	✓	✓	✓	✓	✓	✓
15	Non-verbal reasoning score	✓	✓	✓	✓	✓	✓	✓	✓
15	Mathematics score	✓	✓	✓	✓	✓	✓	✓	✓
15	Verbal reasoning score	✓	✓	✓	✓	✓	✓	✓	✓
16	Is there anything which causes a lot of worry	✓	x	✓	x	✓	x	x	x
26	Job stress	✓	✓	✓	✓	✓	x	✓	✓
31	Any hospital admissions	✓	✓	x	x	x	x	x	x
36	Blood pressure	x	x	x	✓	x	x	✓	✓
36	Ever been unemployed	x	✓	x	x	x	x	x	x
36	Current smoking status	✓	✓	✓	x	✓	✓	✓	✓
36	Does religious upbringing have effect on life now	✓	✓	✓	x	✓	✓	✓	✓
36	Marital status	✓	✓	✓	✓	✓	x	x	✓
43	Verbal memory score	✓	✓	✓	✓	✓	✓	✓	✓

Table 4.14: Step 4 of identifying variables for the imputation model: Potential auxiliary variables for the NSHD imputation models

Potential auxiliary variable	Hypothesised to improve the fit of:	Men	Women
Child has own bed (age 4)	Crowding (age 4)	-	Yes
Clothes repair (age 4)	Child's shoes (age 4)	-	Yes
Repair of dwelling (age 4)	Household amenities (age 2)	Yes	Yes
Difficult to discipline (age 10)	Educational qualifications (age 26)	Yes	-
Household amenities (age 11)	Household amenities (age 2)	Yes	Yes
Father's occupational SEP (age 11)	Father's occupational SEP (age 4)	-	Yes
Have to stick up for myself (age 13)	Reading score (age 15)	-	Yes
Truancy during past year (age 15)	Reading score (age 15)	-	Yes
Happiness (age 26)	Job stress (age 26)	Yes	Yes
Brought up in any faith or religious denomination (age 26)	Does religious upbringing have effect on life now (age 36)	-	Yes
Currently have a religion (age 26)	Does religious upbringing have effect on life now (age 36)	-	Yes
How confident do you feel in yourself (age 36)	Present State Examination total score (age 36)	-	Yes
Ever had cancer (age 36)	Ever been unemployed (age 36)	Yes	No
Verbal fluency score (age 53)	NART score (age 53)	Yes	Yes

Table 4.15: Final NSHD imputation model for women

Age	Predictive of missingness and observed value	Age	Included in model of interest	Age	Auxiliary variables
4	Crowding	0	Father's occupational SEP	4	Child has own bed
4	Age of dwelling	2	Household amenities	4	Clothes repair
4	Yard or garden	4	Father's occupational SEP	4	Repair of dwelling
4	Child's shoes	8	Reading score	11	Household amenities
4	Cleanliness of house	8	Sentence completion score	11	Father's occupational SEP
4	Cleanliness of child	8	Picture intelligence score	13	Have to stick up for myself
4	Mother's management and understanding of the child	8	Vocabulary score	15	Truancy during past year
10	Difficult to discipline	8	Cognitive score at age 8 squared	26	Happiness
11	Non-verbal reasoning score	26	Own occupational SEP	26	Brought up in any faith or religious denomination
11	Reading score	26	Educational qualifications	26	Currently have a religion
11	Verbal reasoning score	43	Own occupational SEP	36	How confident do you feel in yourself
11	Arithmetic score	43	Head of household occupational SEP	53	Verbal fluency score
11	Vocabulary score	36	Own occupational SEP		
15	Reading score				
15	Non-verbal reasoning score				
15	Mathematics score				
15	Verbal reasoning score				
26	Job stress				
31	Any hospital admissions				
36	Ever been unemployed				
36	Blood pressure				
36	Does religious upbringing have effect on life now				
36	Marital status				
36	Current smoking status				
43	Verbal memory score				

Table 4.16: Final NSHD imputation model for men

Age	Predictive of missingness and observed value
4	Child has own bed
4	Clothes repair
4	Yard or garden
4	Child's shoes
4	Cleanliness of house
4	Cleanliness of child
4	Mother's management and understanding of the child
4	Crowding
11	Non-verbal reasoning score
11	Reading score
11	Verbal reasoning score
11	Arithmetic score
11	Vocabulary score
11	Father's occupational SEP
13	Daredevil (teacher)
13	Have to stick up for myself
15	Reading score
15	Non-verbal reasoning score
15	Mathematics score
15	Verbal reasoning score
15	Truancy during past year
26	Brought up in any faith or religious denomination
26	Currently have a religion
26	Job stress
36	Marital status
36	Current smoking status
36	Ever been unemployed
36	Does religious upbringing have effect on life now
43	How many friends or relatives could you visit at any time without waiting for an invite
43	Verbal memory score

Age	Included in model of interest
0	Father's occupational SEP
2	Household amenities
4	Father's occupational SEP
8	Reading score
8	Sentence completion score
8	Vocabulary score
8	Cognitive score at age 8 squared
26	Own occupational SEP
26	Educational qualifications
36	Own occupational SEP
43	Own occupational SEP

Age	Auxiliary variables
4	Repair of dwelling
10	Difficult to discipline
11	Household amenities
26	Happiness
36	Ever had cancer
53	Verbal fluency score

Table 4.17: Comparison of restricted imputation models with full imputation models, for NSHD women, with the outcome NART

WOMEN (N=2,547)	Imputation model 1		Imputation model 2		Imputation model 3	
	coef (se)	p-value	coef (se)	p-value	coef (se)	p-value
Baseline: Childhood SEP - manual						
Childhood SEP - non-manual	1.07 (0.43)	0.015	1.07 (0.41)	0.011	1.18 (0.49)	0.019
Baseline: no qualifications/proficiency only		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level	4.36 (0.54)	<0.001	4.20 (0.47)	<0.001	4.33 (0.56)	<0.001
Education - GCE 'A'-Level	6.64 (0.65)	<0.001	6.61 (0.59)	<0.001	6.56 (0.63)	<0.001
Education – Degree	7.26 (1.12)	<0.001	7.38 (1.04)	<0.001	7.31 (1.15)	<0.001
Baseline: Age 43 own SEP - manual						
Age 43 own SEP - non-manual	0.93 (0.50)	0.067	0.95 (0.58)	0.107	0.99 (0.53)	0.067
Standardised age 8 cognitive score	6.13 (0.29)	<0.001	5.99 (0.32)	<0.001	5.78 (0.32)	<0.001
Standardised age 8 cognitive score squared	-0.68 (0.25)	0.010	-0.63 (0.23)	0.011	-0.49 (0.20)	0.017
Constant	29.24 (0.50)	<0.001	29.35 (0.62)	<0.001	29.21 (0.57)	<0.001

<u>Imputation model 1</u>: Full imputation model in Table 4.16
<u>Imputation model 2</u>: Imputation model 1 without the auxiliary variables (see Table 4.16)
<u>Imputation model 3</u>: NART at age 53, father’s occupational SEP at birth and age 4, verbal memory at age 43, crowding at age 4, occupational SEP at ages 26, 36 and 43, head of household occupational SEP at age 43, cognitive tests at ages 8 and 11, cognitive function at age 8 squared, educational qualifications at age 26, marital status at age 36, household amenities at age 2.

Table 4.18: Comparison of number of imputations, for NSHD women, with the outcome NART

WOMEN (N=2,547)	5 imputations		10 imputations		20 imputations	
	coef (se)	p-value	coef (se)	p-value	coef (se)	p-value
Baseline: Childhood SEP - manual						
Childhood SEP - non-manual	1.01 (0.41)	0.019	0.86 (0.43)	0.049	1.07 (0.43)	0.015
Baseline: no qualifications/proficiency only		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level	4.38 (0.48)	<0.001	4.37 (0.50)	<0.001	4.36 (0.54)	<0.001
Education - GCE 'A'-Level	6.42 (0.58)	<0.001	6.53 (0.54)	<0.001	6.64 (0.65)	<0.001
Education – Degree	7.18 (1.22)	<0.001	7.26 (1.24)	<0.001	7.26 (1.12)	<0.001
Baseline: Age 43 own SEP - manual						
Age 43 own SEP - non-manual	1.03 (0.43)	0.020	0.86 (0.50)	0.096	0.93 (0.50)	0.067
Standardised age 8 cognitive score	6.12 (0.24)	<0.001	6.02 (0.30)	<0.001	6.13 (0.29)	<0.001
Standardised age 8 cognitive score squared	-0.53 (0.16)	0.003	-0.53 (0.17)	0.003	-0.68 (0.25)	0.010
Constant	29.09 (0.40)	<0.001	29.35 (0.50)	<0.001	29.24 (0.50)	<0.001

4.5.2 Developing an imputation model for Whitehall II

The variables that were considered for the imputation model are in Appendix 8: Table A8.1. As in the NSHD, only variables from earlier phases were tested to investigate whether they were predictive of missingness. One exception to this rule was the retrospective measure of childhood emotional deprivation, which was collected at phase 5, but was a measure from the participants' childhoods. This was therefore allowed to predict missingness of other variables at phase 5.

The variables that were included in the models of interest (analysing the association of childhood SEP with adult cognitive function), and were therefore included in the imputation model, were: Mill Hill test score at phase 9; father's occupational SEP in childhood; childhood material deprivation; educational qualifications; last recorded occupational grade at phase 7; age at phase 9; and the number of times the participant had taken the cognitive tests at phase 9. As the childhood material deprivation variable was created from four binary variables, these four variables were included in the imputation model: whether the participant's parents were unemployed during the participant's childhood; whether the family had financial problems during the participant's childhood; whether the house had an outside toilet during the participant's childhood; and whether the family owned a car during the participant's childhood.

The results from the tests of whether variables were associated with the missingness of variables in the model of interest can be seen in Appendix 8: Table A8.1, and Appendix 8: Table A8.2 (men) and Appendix 8: Table A8.3 (women) show which of the variables from Appendix 8 Table A8.1 were predictive of the values of the variables in the model of interest. For women, all variables that were predictive of missingness were also associated with all of the variables in the model of interest. For men, only whether the participant had ever been diagnosed with cancer was associated with missingness but not with any of the variables in the model of interest. As childhood emotional deprivation was associated with both missingness and the variables of interest for both men and women, the four variables that make up the variable were included in the imputation model.

The auxiliary variables considered for Whitehall II are in Appendix 8: Table A8.4, along with whether they were significant predictors of other variables already in the imputation model. This led to the final imputation models for men (Table 4.19) and women (Table 4.20).

Table 4.19: Final imputation model for Whitehall II men

Phase	Predictive of missingness and observed value
1	Year of birth
1	Age finished full time education
1	Accommodation type
1	State of health in the last year
1	Any longstanding illnesses?
1	Smoking status
1	Usually pressed for time
1	Grouped occupational grade
1	Isolation score
1	Believe no one cares much about you
3	Marital status
3	Memory score
3	AH4 score
3	Mill Hill test score
3	Verbal fluency - S words
3	Verbal fluency – animals
3	Job involves travel away from home
3	Last recorded occupational grade
4	Ever told had depression
4	Ever told had anxiety
5	Ever told high blood pressure
5	Deprivation score
5	Memory score
5	AH4 score
5	Mill Hill test score
5	Verbal fluency - S words
5	Verbal fluency – animals
5	To what extent do you feel you might as well give up because you can't make things better for yourself
5	How financially secure do you feel in next 10 years
5	Last recorded occupational grade
7	General health
7	Health stops you from doing what you want to do
7	Clinic or home visit
7	MMSE score
7	Still at civil service
7	Memory score
7	AH4 score
7	Mill Hill test score
7	Verbal fluency - S words
7	Verbal fluency - animals
7	CASP score

Phase	Included in model of interest
1	Father's occupational SEP in childhood
5	Childhood material deprivation - 4 raw variables
5	Educational qualifications
7	Last recorded occupational grade
9	Mill Hill test score
9	Age
9	Number of times taken cognitive tests

Phase	Auxiliary variables
1	Age mother finished full time education
5	Childhood emotional deprivation - 4 raw variables
9	Memory score
9	AH4 score
9	Verbal fluency - S words
9	Verbal fluency - animals
9	MMSE score
9	General health
9	Difficulty paying bills
9	Marital status
9	Last recorded occupational grade

Table 4.20: Final imputation model for Whitehall II women

Phase	Predictive of missingness and observed value	Phase	Included in model of interest
1	Year of birth	1	Father's occupational SEP in childhood
1	Age finished full time education	5	Childhood material deprivation
1	Accommodation type	5	Educational qualifications
1	Age mother finished full time education	7	Last recorded occupational grade
1	State of health in the last year	9	Mill Hill test score
1	Any longstanding illnesses?	9	Age
1	Smoking status	9	Number of times taken cognitive tests
1	Usually pressed for time		
1	Grouped occupational grade		
1	Isolation score		
3	Marital status	**Phase**	**Auxiliary variables**
3	Memory score	1	Job satisfaction
3	AH4 score	5	To what extent do you feel you might as well give up because you can't make things better for yourself
3	Mill Hill test score	9	Memory score
3	Verbal fluency - S words	9	AH4 score
3	Verbal fluency – animals	9	Verbal fluency - S words
3	Job involves travel away from home	9	Verbal fluency - animals
3	Last recorded occupational grade		MMSE score
5	Ever told high blood pressure	9	General health
5	Deprivation score	9	Difficulty paying bills
5	Memory score	9	Marital status
5	AH4 score	9	Last recorded occupational grade
5	Mill Hill test score		
5	Verbal fluency - S words		
5	Verbal fluency – animals		
5	Childhood emotional deprivation		
5	Last recorded occupational grade		
7	General health		
7	Health stops you from doing what you want to do		
7	Clinic or home visit		
7	MMSE score		
7	Still at civil service		
7	Memory score		
7	AH4 score		
7	Mill Hill test score		
7	Verbal fluency - S words		
7	Verbal fluency – animals		

4.6 Heckman selection model methodology

As outlined in Chapter 2, there are many possible methods of choosing which variables to include in a Heckman selection model. The approach used here was the proportion of statistically significant coefficient estimates, since pseudo R^2s can only be compared across models using the same sample, which would result in the loss of some data.

As explained in Chapter 2, it is necessary for the selection model to contain at least one variable that is associated with missingness but is not included in the model of interest. There are also those who consider a stricter condition to be necessary: for the selection model to contain a variable that is associated with missingness, but not associated with the outcome of interest (134). A list of variables that were associated with missingness of the outcome variable, but not included in the analysis of interest was developed. The potential variables were then checked to see whether they were associated with the outcome of interest.

The variables that were predictive of not having an observed outcome were identified whilst selecting variables for the imputation model. For the selection model variables that were *not* associated with the variables of interest were identified, unlike in the imputation model, although variables that were associated with the outcome of interest were also permitted. Backwards selection was carried out on the variables that were predictive of missingness of the outcome variable, until all the variables remaining in the selection model were significant predictors of selection, giving a proportion of 1 for significant variables. This was carried out first according to the stricter conditions, constraining a variable that was predictive of selection and not associated with the outcome of interest to be in the selection model, and then for the less strict conditions, where only a variable that was predictive of selection but not included in the model of interest was required to be in the selection model.

4.6.1 Developing a Heckman selection model for the NSHD

The variables that predict not having an observed NART score are presented in Table 4.11. However the model would not converge when all of these variables were included in the selection model. The selection model, which has the binary outcome measure observed or unobserved outcome variable in the model of interest, uses only complete cases; therefore variables from ages 43 and 53 were excluded from the selection model.

For men, 33 variables were identified in section 4.5.1 as being predictive of not having an observed NART score, although four of these variables were collected at age 43 or 53. Six of the 33 variables were not associated with the NART score; at least one of these six variables was required to be in the selection model. When backwards selection was carried out, the final model contained one of the six variables not associated with the NART score, Present State Examination total score (age 36), as well as vocabulary score at age 11, non-verbal reasoning score at age 15 and reading score at age 15, with all four variables significant predictors of missingness. The output from the selection models with the top three proportions is compared in Table 4.21.

For women, 22 variables were identified in section 4.5.1 as being predictive of not having an observed NART score, of which two were collected at age 43 or 53, so were excluded from the selection model. All of the remaining 20 variables were associated with the observed NART score. As being a daredevil at age 13 was a significant predictor of having an observed NART score at a 10% level of significance, but not significantly associated with the observed NART score, this was included in the list of variables on which backwards selection was carried out on. The backwards selection model containing all these variables had a sample size of 258, only 10.1% of the females in the study. Therefore the two variables with the largest amount of missing data for women, the gated question ‘something causing a lot of worry’ at age 16, which was only asked to participants who were in employment, and own occupational SEP at age 36 were dropped from the model, leaving a sample size of 641.

Backwards selection was carried out, with being a daredevil at age 13 constrained to be in the model. An alternative model was run, which did not constrain being a daredevil at age 13 to be in the model. When being a daredevil at age 13 was forced into the model the final model contained: being a daredevil at age 13; arithmetic score at age 11; reading score at age 11; crowding at age 4; smoking status at age 36; cleanliness of the house at age 4; cleanliness of the child at age 4; and job stress at age 26. However when being a daredevil at age 13 was not constrained to remain in the model the final model contained only smoking status at age 36 and reading score at age 15. This model had a proportion of significant variables of 1.

The results from the men's selection models can be found in Table 4.21. However the women's selection models containing daredevil at age 13 would not converge, so the selection models with the next highest proportions of significant variables were tested until a model was found that converged. Daredevil at age 13 was then added to these models.

The specific form of the selection model did not have a large influence on the results for the effect of childhood SEP on crystallized cognitive function in men (Table 4.21). The same variables were significant using each selection model, at similar levels of significance. As there were no major differences, selection model 1 was chosen in order to maintain the largest sample size. The results for women (Table 4.22) were much more dependent on the selection model used. When crowding at age 4 was included in the selection model childhood SEP was a significant predictor (selection models 2 and 4). However when it was not included in the model (selection models 1 and 3), childhood SEP narrowly missed significance at a 5% level. Similarly in selection model 1 (standardised cognitive function at age $8)^2$ was not significant, whereas in selection models 2-4 the squared term was significant. When choosing which of the selection models to use in the main analyses, it was decided to keep daredevil at age 13 in the model as the best variable available that was predictive of missingness (the only variable with $p<0.10$) but not associated with the outcome of interest. The choice was then between selection models 3 and 4. When a probit model was run containing the variables in selection model 4, crowding at age 4 was not a significant predictor of whether a NART score was observed at age 53. Hence model 3 was chosen as the final selection model, containing daredevil at age 13, reading score at age 15 and smoking status at age 36.

When the more lenient conditions were used, the final selection model for men included vocabulary score at age 11, non-verbal reasoning score at age 15, reading score at age 15 and smoking status at age 36, and the final selection model for women contained reading score at age 15 and smoking status at age 36. These selection models are very similar to those chosen using the stricter conditions.

Table 4.21: Comparing Heckman selection models, with outcome NART (NSHD men)

MEN	Selection model 1 (N=1,204)		Selection model 2 (N=1,202)		Selection model 3 (N=1,012)	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP – manual						
Childhood SEP - non-manual	0.27 (0.56)	0.625	0.22 (0.56)	0.695	0.18 (0.61)	0.772
Baseline: no qualifications/proficiency only		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level	1.14 (0.82)	0.163	1.24 (0.82)	0.129	1.04 (0.91)	0.251
Education - GCE 'A'-Level	2.89 (0.80)	<0.001	2.86 (0.81)	<0.001	2.86 (0.84)	0.001
Education – Degree	4.88 (0.93)	<0.001	4.83 (0.93)	<0.001	4.98 (1.02)	<0.001
Baseline: Age 43 own SEP – manual						
Age 43 own SEP - non-manual	2.18 (0.66)	0.001	2.21 (0.66)	0.001	2.09 (0.70)	0.003
Standardised age 8 cognitive score	4.28 (0.38)	<0.001	4.28 (0.38)	<0.001	4.53 (0.41)	<0.001
Standardised age 8 cognitive score squared	-0.63 (0.27)	0.020	-0.66 (0.28)	0.020	-0.64 (0.31)	0.036
Constant	34.24 (0.69)	<0.001	34.22 (0.69)	<0.001	33.98 (0.78)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	51.86	<0.001	54.58	<0.001	37.43	<0.001

Selection model 1: vocabulary score at age 11, non-verbal reasoning score at age 15, reading score at age 15 and Present State Examination total score.
Selection model 2: Selection model 1 plus smoking status at age 36
Selection model 3: Selection model 2 plus mother's management and understanding of the child at age 4, cleanliness of house at age 4 and have a religion at age 26

Table 4.22: Comparing Heckman selection models, with outcome NART (NSHD women)

WOMEN	Model 1 (N=1,236)		Model 2 (N=1,236)		Model 3 (N=1,196)		Model 4 (N=1,196)	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP – manual								
Childhood SEP - non-manual	0.94 (0.48)	0.053	1.02 (0.49)	0.036	0.96 (0.49)	0.052	1.04 (0.50)	0.037
Baseline: no qualifications/proficiency only		<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level	2.97 (0.62)	<0.001	2.84 (0.59)	<0.001	3.01 (0.62)	<0.001	2.88 (0.59)	<0.001
Education - GCE 'A'-Level	5.45 (0.68)	<0.001	5.24 (0.67)	<0.001	5.35 (0.69)	<0.001	5.15 (0.68)	<0.001
Education – Degree	6.59 (1.02)	<0.001	6.40 (1.01)	<0.001	6.31 (1.09)	<0.001	6.11 (1.08)	<0.001
Baseline: Age 43 own SEP – manual								
Age 43 own SEP - non-manual	0.77 (0.60)	0.202	0.94 (0.59)	0.109	0.93 (0.61)	0.124	1.09 (0.59)	0.066
Standardised age 8 cognitive score	4.77 (0.36)	<0.001	4.76 (0.35)	<0.001	4.83 (0.37)	<0.001	4.81 (0.36)	<0.001
Standardised age 8 cognitive score squared	-0.46 (0.24)	0.058	-0.49 (0.24)	0.040	-0.55 (0.25)	0.024	-0.57 (0.25)	0.020
Constant	32.21 (0.64)	<0.001	32.22 (0.63)	<0.001	32.12 (0.65)	<0.001	32.14 (0.64)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	54.54	<0.001	59.11	<0.001	49.38	<0.001	53.84	<0.001

Selection model 1: smoking status at age 36 and reading score at age 15
Selection model 2: selection model 1 plus crowding at age 4
Selection model 3: selection model 1 plus daredevil at age 13
Selection model 4: selection model 2 plus daredevil at age 13

4.6.2 Developing a Heckman selection model for Whitehall II

The variables that predict not having an observed Mill Hill test score at phase 9 can be found in Appendix 8: Table A8.1, however as in the NSHD they could not all be put into the selection model as the model would not converge. The selection model only uses complete cases; therefore variables from phase 7 were excluded from the selection model as including them would reduce the sample size.

For men, 42 variables were identified in section 4.5.2 as being predictive of not having an observed Mill Hill test score. However 12 of these variables were collected at phase 7, so were excluded from the selection model. Of the remaining 30 variables five were not associated with the Mill Hill test score; year of birth (phase 1), isolation score (phase 1), whether the participant had ever been diagnosed with cancer (phase 5), whether the participant had ever been diagnosed with depression (phase 4), and whether the participant had ever been diagnosed with anxiety (phase 4).

When backwards selection was carried out, the final model contained year of birth (phase 1), accommodation type (phase 1), isolation score (phase 1), smoking status (phase 1), usually pressed for time (phase 1), ever told high blood pressure (phase 5), and extent feel might as well give up (phase 5). All these variables were significant predictors of missingness, resulting in a proportion of significant variables of 1. The results from this selection model (selection model 1) were compared with a model that only included phase 1 variables from selection model 1 in order to maintain a larger sample size (selection model 2). The proportion of significant variables was also 1 in selection model 2.

For women, 36 variables were identified in section 4.5.2 as being predictive of not having an observed Mill Hill test score. However 10 of these variables were collected at phase 7, so were excluded from the selection model. Of the remaining 26 variables two were not associated with the Mill Hill test score; having a longstanding illness (phase 1) and isolation score (phase 1).

Backwards selection was carried out. Having a longstanding illness (phase 1) was the third variable to be dropped from the model, and isolation score (phase 1), the other variable which was associated with missingness but not Mill Hill test score at phase 9, was the seventh variable to be dropped. Therefore two selection models were compared; one which forced isolation score to remain in the selection model (selection model 1), and one which allowed isolation score to drop out of the selection model (selection model 2) in order to compare the results when there was not a variable which was not associated with the Mill Hill test score in the selection model. Selection model 1 had a proportion of 0.83 of significant variables, and selection model 2 had a proportion of 1.

In men, similar results for coefficients and significance levels were obtained for selection models 1 and 2 (Table 4.23). As there were no major differences selection model 2 was chosen in order to maintain the largest sample size. The results are also similar for women (Table 4.24), therefore selection model 1 was chosen as it contained the variable isolation score, which was a significant predictor of missingness but not Mill Hill test score at phase 9.

When the more lenient conditions were used, the final selection model for men contained: year of birth (phase 1); isolation score (phase 1); state of health in the last year (phase 1); occupational grade (phase 1); smoking status (phase 1); ever told high blood pressure (phase 5); and how financially secure the participant felt for the next ten years (phase 5). The final selection model for women contained: year of birth (phase 1); smoking status (phase 1); educational qualifications (phase 5); AH4 score (phase 5); and usually pressed for time (phase 1). Although educational qualifications is included in the model of interest, one overlapping variable between the model of interest and selection model should not cause problems due to collinearity.

Table 4.23: Comparing Heckman selection models for men in Whitehall II, with outcome Mill Hill test

Men	Selection model 1 (N=3,507)		Selection model 2 (N=5,032)	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value
Baseline: Father's occupational SEP in childhood - manual				
Father's occ. SEP - non-manual	0.52 (0.13)	<0.001	0.45 (0.12)	<0.001
Baseline: no educational qualifications		<0.001		<0.001
Education: School certificate/matriculation	1.69 (0.37)	<0.001	1.73 (0.37)	<0.001
Education: O-Level/A-Level/National diploma/Certificate	2.24 (0.26)	<0.001	2.15 (0.26)	<0.001
Education: University degree	3.41 (0.27)	<0.001	3.32 (0.27)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical		<0.001		<0.001
Last occ. grade - Senior and Higher Executive Officers	0.72 (0.26)	0.006	0.58 (0.26)	0.025
Last occ. grade – Unified Grades 1-6	1.75 (0.27)	<0.001	1.61 (0.27)	<0.001
Age (phase 9)	0.08 (0.01)	<0.001	0.07 (0.01)	<0.001
No. of times taken cognitive tests	0.07 (0.08)	0.377	0.09 (0.07)	0.205
Constant	18.07 (0.97)	<0.001	20.15 (0.99)	<0.001

Model fit	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	146.16	<0.001	163.73	<0.001

Selection model 1: year of birth (phase 1), accommodation type (phase 1), isolation score (phase 1), smoking status (phase 1), usually pressed for time (phase 1), ever told high blood pressure (phase 5), extent feel might as well give up (phase 5).

Selection model 2: year of birth (phase 1), accommodation type (phase 1), isolation score (phase 1), smoking status (phase 1), usually pressed for time (phase 1).

Table 4.24: Comparing Heckman selection models for women in Whitehall II, with outcome Mill Hill test

Women	Selection model 1 (N=1,152)		Selection Model 2 (N=1,170)	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value
Baseline: Father's occupational SEP in childhood – manual				
Father's occ. SEP - non-manual	0.78 (0.27)	0.003	0.77 (0.27)	0.005
Baseline: no educational qualifications		<0.001		<0.001
Education: School certificate/matriculation	-0.17 (0.52)	0.746	-0.20 (0.54)	0.704
Education: O-Level/A-Level/National diploma/Certificate	1.28 (0.37)	0.001	1.30 (0.38)	0.001
Education: University degree	2.92 (0.50)	<0.001	3.09 (0.51)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical		<0.001		<0.001
Last occ. grade - Senior and Higher Executive Officers	1.62 (0.31)	<0.001	1.77 (0.31)	<0.001
Last occ. grade – Unified Grades 1-6	2.51 (0.42)	<0.001	2.72 (0.43)	<0.001
Age (phase 9)	0.12 (0.03)	<0.001	0.09 (0.03)	0.001
No. of times taken cognitive tests	0.29 (0.22)	0.186	0.32 (0.22)	0.141
Constant	13.86 (2.21)	<0.001	15.21 (2.20)	<0.001

Model fit	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	158.96	<0.001	142.55	<0.001

Selection model 1: smoking status (phase 1), usually pressed for time (phase 1), year of birth (phase 1), memory score (phase 5), AH4 score (phase 5), isolation score (phase 1).
Selection model 2: smoking status (phase 1), usually pressed for time (phase 1), year of birth (phase 1), memory score (phase 5), AH4 score (phase 5)

4.7 Multiple imputation results

4.7.1 Comparing results for MI and MID in the NSHD

This section compares the results of carrying out the same model development process as in the complete case analysis (section 4.4) using first, multiple imputation (MI) where the imputed NART scores are kept, and second, multiple imputation, then deletion (MID), where the imputed NART scores are deleted (see Appendix 10 for tables).

For all of the models for men (Appendix 10) SEP at ages 26 and 43 had a smaller effect in the MID models (example using father's occupational SEP and educational qualifications, Model 5: age 26 (education up to GCSE/O-Level): MI: 3.20 (95% CI: 1.75, 4.65), MID: 2.69 (95% CI: 1.22, 4.16); age 43: MI: 3.57 (95% CI: 2.49, 4.65), MID: 3.33 (95% CI: 2.10, 4.56)), but childhood SEP had a larger effect size (MI: 0.15 (95% CI: -0.77, 1.07), MID: 0.46 (95% CI: -0.54, 1.46)), although the differences are not statistically significant. When household amenities was used as the childhood SEP variable childhood cognitive function had a linear effect in the MI models, but the coefficient for the quadratic term was larger and significant in the MID models.

For women, father's occupational SEP remained significant in all the models for both MI and MID. When only SEP variables were included in the model the effect size of childhood SEP was slightly larger in the MI analyses than the MID analyses (MI: 6.94 (95% CI: 5.90, 7.98), MID: 6.25 (95% CI: 5.11, 7.39)). However, when childhood cognitive function was added to the model (Model 4), the effect size of childhood SEP became slightly larger in the MID models than the MI models (MI: 1.01 (95% CI: 0.17, 1.85), MID: 1.07 (95% CI: 0.21, 1.93)), with the coefficient for childhood cognitive function higher in the MI models.

There were some differences in the significance of SEP variables between MI and MID models. For example when household amenities was used as the childhood SEP measure, adding occupational SEP at age 26 to the model (Model 2) fully attenuated childhood SEP in the MID analyses ($p=0.087$), but not in the MI analyses ($p=0.024$). In the MI analyses childhood SEP was not fully attenuated until Model 4, when adult SEP and child cognitive function were included in the model.

For all the women's models, cognitive function squared was significant in the MI models but not the MID models. This was due a difference in the coefficients, and was therefore not merely a result of the smaller sample size.

4.7.2 Discussion of the comparison of results for MI and MID

Overall, the conclusions drawn were fairly similar for the MI and MID analyses. However for men childhood SEP generally had a larger effect size in MID analyses, whereas early adulthood and adult SEP had lower effect sizes in MID analyses. The differences between the MI and MID analyses are due to the different sample sizes and compositions (Table 4.25). The MID dataset had higher mean cognitive scores and higher proportions in the non-manual occupational SEP categories, as well as higher levels of education and access to more amenities in childhood.

Although von Hippel claims that MID tends to give shorter confidence intervals than MI, the standard errors were smaller in the MI results than the MID results (123), perhaps due to the larger sample size. For men the sample size was 2,815 in the MI analyses, the size of the original sample, whereas for MID the sample was 1,370, 49% of the initial sample. For women equivalent sample sizes were 2,547 and 1,455 (57%). This means that 49% of men and 57% of women had a NART score recorded.

As there were important differences in the results, it was important to use the method that is most theoretically correct. Although additional variables were added to the imputation model beyond the variables in the model of interest, von Hippel (123) concluded that the additional information needs to be 'quite good' before the extra information 'trumps the extra variation by using a finite number of randomly imputed' dependent variables, and makes the point that many of the variables added to the imputation model will themselves contain missing data, limiting the amount of information added. For example, in the NSHD variables for women from the early childhood period of the life course, around 15-20% of the data are missing (for example, whether the child shared a bed at age 4: 18% missing, mother's management and understanding of the child: 16% missing); for the early adulthood period of the life course around 25-30% of the data are missing (29% of father's social class at age 15); for early middle age, around 35% of data were missing (smoking status at age 36: 35% missing, marital status at 36: 35% missing), and at the older ages over 40% of the data were missing (verbal memory at age 53: 42% missing). Therefore the results from the MID analyses were used in future comparisons.

Table 4.25: Summary of variables under MI and MID

	Men Mean (s.d.)		Women Mean (s.d.)	
	MI (N=2,815)	MID (N=1,370)	MI (N=2,547)	MID (N=1,455)
NART	32.66 (10.62)	34.41 (9.66)	33.07 (10.11)	34.21 (9.42)
Standardised cognitive function (age 8)	-0.00 (0.91)	0.11 (0.87)	0.05 (0.90)	0.13 (0.86)
	Proportion (s.d.)		Proportion (s.d.)	
Father's occ. SEP non-manual (age 4)	0.40 (0.49)	0.43 (0.49)	0.42 (0.49)	0.44 (0.50)
Own occ. SEP non-manual (age 43)	0.60 (0.49)	0.64 (0.48)	0.70 (0.46)	0.74 (0.44)
HoH occ. SEP non-manual (age 43)	-	-	0.61 (0.49)	0.65 (0.49)
Access to household amenities	%	%	%	%
0	4.75	4.63	4.59	5.02
1	30.14	28.46	30.83	29.12
2	19.59	18.09	17.11	16.73
3	45.52	48.82	47.46	49.13
Educational qualifications	%	%	%	%
No qualifications/ proficiency only	42.44	35.39	45.05	39.56
Up to GCE 'O'-Level	19.64	19.89	29.48	31.11
GCE 'A'-Level	25.68	29.34	20.73	24.17
Degree	12.23	15.38	4.73	5.16
Own occupational SEP (age 26)	%	%	%	%
RG: IV and V	15.73	13.01	20.65	17.35
RG: IIIM	36.55	34.59	10.07	8.95
RG: IIINM	14.23	15.32	46.60	48.04
RG: I and II	33.49	37.08	22.68	25.67

4.8 Heckman selection results

The results from the two sets of conditions for choosing a selection model are compared. In this section they are described as the stricter (required a variable which was predictive of selection but not associated with the outcome of interest) and less strict (required a variable which was predictive of missingness and not an independent variable in the model of interest) sets of conditions.

The results using the stricter conditions are in Appendix 11 (NSHD) and Appendix 12 (Whitehall II), and using the less strict conditions are in Appendix 13 (NSHD) and Appendix 14 (Whitehall II).

Overall the conclusions reached were very similar in the NSHD. The only difference in significance of the SEP variables between the results from the two selection models was for men when using father's occupational SEP, educational qualifications and own occupational SEP, childhood SEP was a significant predictor of NART score using the stricter conditions (p=0.045) (Appendix 11: Table A11.1) but not using the less strict conditions (p=0.062) (Appendix 13: Table A13.1). In general the coefficients for the SEP variables were slightly higher using the stricter selection model for men, although the SEP coefficients were more similar for women. However for women there was a difference in whether the relationship between childhood cognitive function and NART score was linear; using the less strict conditions the quadratic term was not significant in any of the models, whereas using the stricter conditions it was significant in all but one of the models.

There were no differences in the significance of the variables in the Whitehall II analyses. The adult SEP coefficients were lower in the analyses using the less strict selection model, especially for men, although the early adulthood SEP coefficients were higher.

For the comparison with the results from the complete case and multiple imputation analyses the results from the stricter conditions will be used, as they are the more stringent conditions.

4.9 Comparing results from the complete case, MID and Heckman selection analyses

4.9.1 NSHD

The main comparison of results is now presented, comparing the preferred multiple imputation and Heckman selection models with the complete case analysis. The results for the model development process are discussed below, for the complete case (Appendix 7), MID (Appendix 10) and the Heckman selection analyses using the stricter conditions (Appendix 11).

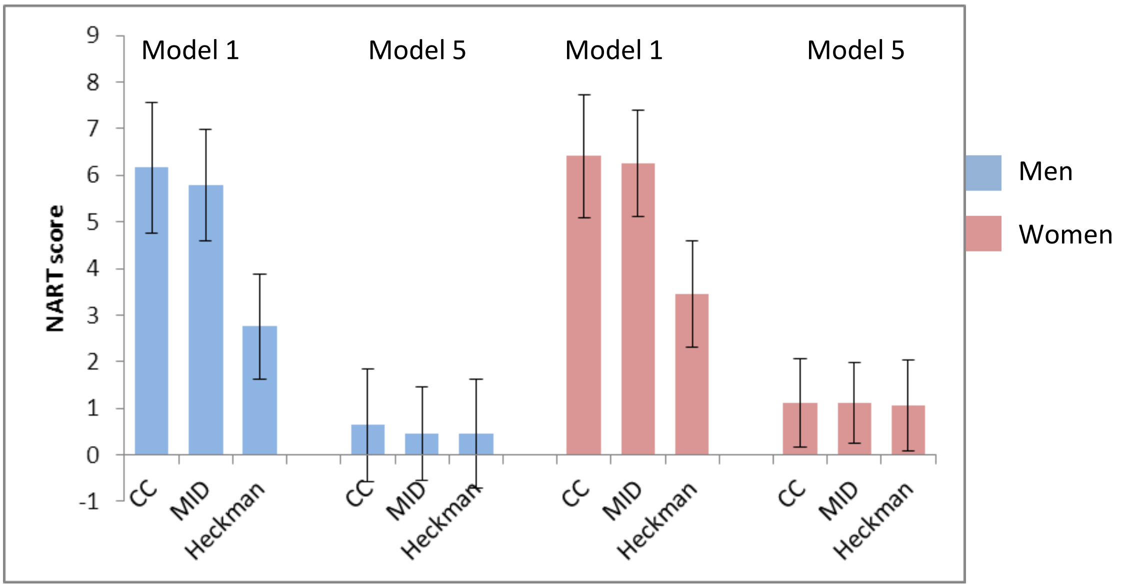


Figure 4.1: Coefficient for father's occupational SEP (non-manual) in the NSHD, with outcome NART score at age 53 (Model 1: father's occupational SEP, Model 5: Father's occupational SEP, educational qualifications, occupational SEP, cognitive function at age 8, (cognitive function at age 8)^2)

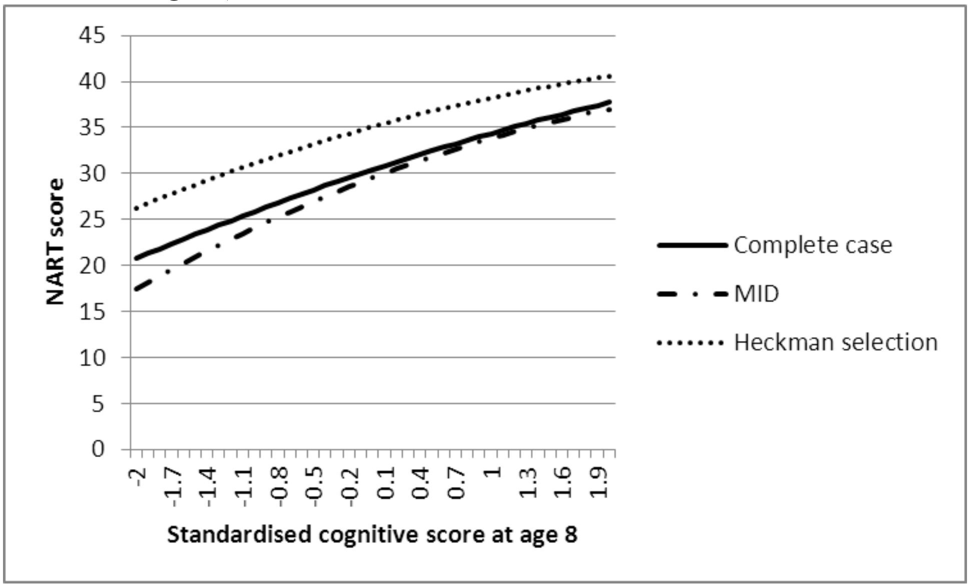
Men:

When father's occupational SEP was used as the measure of childhood SEP, for each of the complete case, MID and Heckman selection analyses childhood SEP remained significant when early adulthood SEP and adult SEP were added to the model, but was fully attenuated when childhood cognitive function was added to the model (Figure 4.1). In all four of the models for men, the coefficient for childhood SEP was the largest in the complete case analyses, both when it was significant and when it was fully attenuated. When household amenities was used as the childhood SEP variable, childhood SEP was not a significant predictor of NART score in the unadjusted Heckman model, whereas in the unadjusted complete case and MID analyses childhood SEP was significant.

When father's occupational SEP was used as the measure of childhood SEP, (cognitive function at age 8)^2 was significant in the MID analyses, but not the complete case or Heckman selection analyses (Figure 4.2).

There was no consistent order of coefficient sizes between the three analyses, but there were differences between the three sets of results. When the three SEP variables were father's occupational SEP, educational qualifications and own occupational SEP at age 43, the largest coefficients were generally found in the MID analyses, except for childhood SEP. When own occupational SEP was used as the measure of early adulthood SEP, early adulthood SEP was not significant in the Heckman analyses, although it was significant in the complete case and MID analyses. This is due to the much smaller coefficients rather than larger standard errors.

Figure 4.2: NART score at age 53 under complete case, MID and Heckman selection, for baseline SEP conditions (manual father's occupational SEP, no educational qualifications and manual own occupational SEP at age 43), for men



<u>Women:</u>

Overall, the significance levels of the variables were very similar for the complete case, MID and Heckman selection model results. One difference was when only household amenities in childhood was in the model; as in the men's analysis, in the Heckman selection model this variable was not significant, whereas it was in the complete case and MID analyses. Another situation was in the models containing own occupational SEP at ages 26 and 43, where own occupational SEP at age 43 was not significant after adjusting for childhood cognitive function in the Heckman selection analyses, but was in the complete case and the MID analyses. In all the models except the model containing household amenities, own occupational SEP at age 26 and own occupational SEP at age 43, $(\text{childhood cognitive function})^2$ was not significant in the MID analyses, but was significant in the complete case and Heckman selection models.

4.9.2 Whitehall II

There were some differences between the results for the complete case, MID (Appendix 9) and Heckman selection analyses (Appendix 12). The differences in the significance of childhood SEP are focussed on (Figure 4.3).

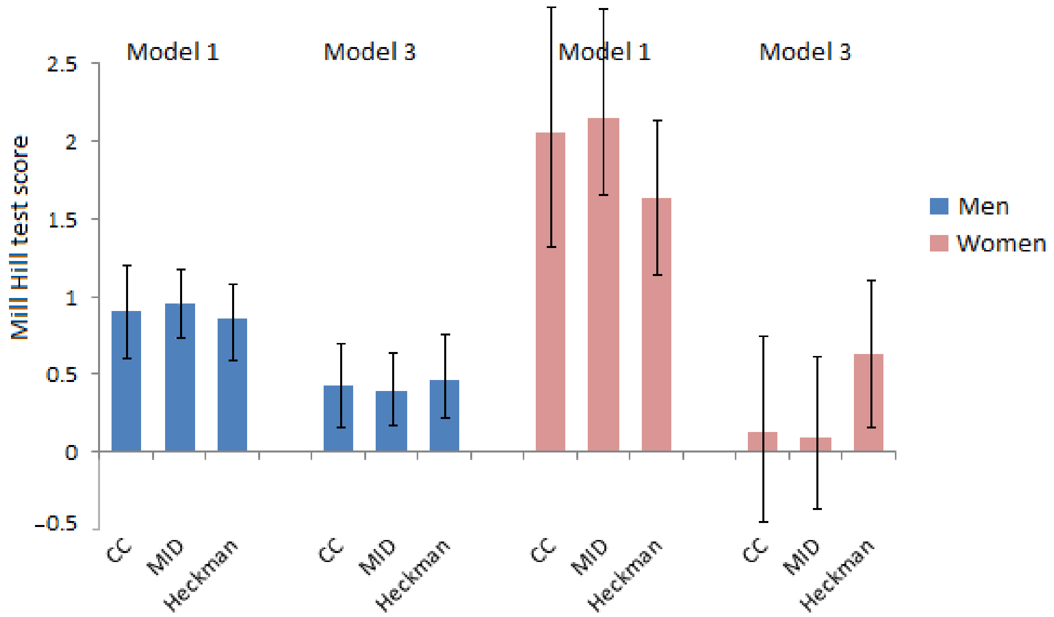


Figure 4.3: Coefficient for father's occupational SEP (non-manual) in the Whitehall II, with outcome Mill Hill test score at phase 9 (Model 1: father's occupational SEP, Model 3: Father's occupational SEP, educational qualifications, occupational SEP)

Men

Childhood material deprivation was fully attenuated by the addition of educational qualifications in the Heckman selection analyses, whereas childhood material deprivation remained a significant predictor in all three models for the complete case and MID analyses. In the models using childhood material deprivation the SEP coefficients were generally largest in the MID analyses, and smallest in the Heckman analyses.

Women

Childhood material deprivation was significant in the unadjusted model for the complete case, MID and Heckman selection analyses. However when father's occupational SEP was the childhood SEP measure, childhood SEP was only significant in Model 1 for the complete case and MID analyses, but remained significant in all three models in the Heckman selection models. The childhood SEP coefficient was smallest in the Heckman selection analyses in Model 1, but it was attenuated to a much smaller extent when adjusting for the other SEP variables.

4.10 Discussion

Main findings

In the NSHD whether childhood SEP remained a significant predictor of adult cognitive function after adjustment for later life SEP depended on the childhood SEP measure used. In the NSHD father's occupational SEP remained a significant predictor after adjusting for educational qualifications and occupational SEP in adulthood, but childhood household amenities was fully attenuated. However in the Whitehall II analyses childhood SEP remained significant for men but not women for both of the childhood SEP measures.

The results from the Heckman selection analyses sometimes differed from those for the complete case and multiple imputation analyses, particularly in the unadjusted childhood material deprivation/household amenities analyses.

Comparison with other studies

In the analyses carried out in this chapter, childhood SEP was almost always significant in the unadjusted analyses; this agrees with the vast majority of the studies discussed in section 1.2.6, the only exception being certain measures of childhood SEP in the study by Kaplan et al (73), who found no association between father's education or mother's occupation and adult cognitive function.

As in the analyses carried out above, Kaplan et al (73) found that whether there was a significant effect of childhood SEP on adult cognitive function after adjusting for later life SEP depended on the childhood SEP measure used. In this study an effect remained for father's occupation in the NSHD, whereas Kaplan et al found that the effect of father's occupation was fully attenuated. Richards and Sacker (81) investigated the effect of father's occupational SEP in the NSHD using structural equation modelling and found that a significant but 'substantially unimportant' effect of childhood SEP remained. In the analyses above the effect of father's occupational SEP was fully attenuated by adjustment for childhood cognitive function for men, but not for women.

Two other studies investigated the effect of father's occupation, both collected retrospectively; Zhang et al (78) found that father's occupation remained significant, whereas Johnson et al (83) found that it was fully attenuated. The cognitive measure in the study by Zhang et al was constructed from measures of memory, verbal fluency, digit span and block design, whereas Johnson used the Moray House test, which consists of verbal, numerical and spatial reasoning items (20).

Johnson et al (83) investigated 'environmental deprivation' at age 11, which consisted of the number of rooms in their home, the number of people living in their home, indoor or outdoor toilet facilities, and the number of people sharing toilet facilities. This is a similar variable to the childhood material disadvantage variable used in the Whitehall II analyses and the household amenities variable used in the NSHD analyses. Johnson et al found no effect of environmental deprivation on cognitive function in adulthood after adjustment for later life SEP and childhood cognitive function, the same result as was found in the NSHD.

Explanation of findings

Epidemiological

Both measures of childhood SEP were significant after adjustment for later life SEP for men but not women in the Whitehall II study. The sample size was smaller for women than men, leading to larger standard errors, which is likely to be the reason for the difference when father's occupational SEP was used, as the regression coefficients were very similar for men and women in both the complete case and MID analyses after adjustment for educational qualifications. However after adjustment for educational qualifications the coefficient was lower for women than men when childhood household amenities were used, implying the smaller sample size was not the reason for the difference in significance. As the coefficient was larger in the unadjusted models, this shows that a larger proportion of the effect of childhood material deprivation acted through educational qualifications for women than men.

The fact that no effect of childhood material deprivation on adult cognitive ability after adjustment for later life SEP was found in the NSHD does not imply that it is not an important factor in determining later life cognitive function; rather it acts indirectly and influence both childhood cognitive function and later life SEP measures.

The significant independent effect of father's occupational SEP which existed even after adjustment for later life SEP demonstrates how important this effect is in shaping cognitive development. However it does not only act through cognitive development, as for women the effect remained after additional adjustment for childhood cognitive function. It is not clear why the results should differ for men and women, although adjustment for childhood cognitive function removed the significance of adult occupational SEP for women but not men. As Richards and Sacker (81) point out, this may be a period effect, and reflect the underachievement occupationally of women given their cognitive ability.

There are a variety of mechanisms through which childhood SEP may affect cognitive function in adulthood, although it is less clear why father's occupational SEP would have an effect that access to household amenities does not. It is possible that father's occupational SEP has an effect on the environment the child grows up in that material conditions do not. The adjustment for educational qualifications attenuated the effect of father's occupational SEP, demonstrating that education is on the pathway between father's occupational SEP and adult cognitive function, perhaps due to increased educational opportunities. Further attenuation was found with own occupational SEP in adulthood. As childhood material deprivation was fully attenuated by educational qualifications, this implies that all the whole effect acted through educational qualifications.

Father's occupational SEP may influence the environment in the home during childhood, including the level of cognitive stimulation (159;160). Turrell et al (77) suggest that the exposure to the stimulating environment results in 'more extensive brain development as indicated by increased cortical thickness and dendritic branching and improved communication among neuron networks'. It is also possible that the difference is due to genetics, as intelligence is highly heritable (161), and the father's level of intelligence is likely to be associated with his occupational SEP. Other potential mediators include health (162), nutrition (163) and physical activity (164).

The difference between the results in the two cohorts is likely to be due to the Whitehall II study being an occupational cohort, unlike in the general population. However there are other important differences between the two cohorts, with the Whitehall II participants born from 1930 to 1952, leading to very different childhood experiences to the NSHD participants, who were all born after the Second World War was finished, although the effects of the war were still in existence, such as rationing. Due to the unique periods at which these cohorts were born there is limited generalizability of the results, and it will be interesting to see how the results compare to those of the later birth cohort studies, when the participants are older.

Some researchers consider the adjustment of later life SEP to be an overadjustment (154) when considering the relationship between childhood SEP and an outcome later in life, and have the opinion that it is the total effect of childhood SEP that is important, as the early life factors 'set in motion a trajectory of social and behavioural exposures that persists into old age' (165). The total effect of childhood SEP was investigated in Model 1, however whether an effect of childhood SEP remained after holding later life SEP constant was also of interest; this would allow interventions or additional assistance to be focused on those participants who were at the greatest risk of needing it. Overadjustment typically biases results towards the null (166), and a finding of no significant effect of childhood SEP after adjusting for later life SEP would not lead to the conclusion that childhood SEP is not important, as it is known to be important through its effect on later life SEP. A related issue to overadjustment is that of collider bias, where the association between two variables is affected by conditioning on a common effect (167). In this situation the collider variable, adult SEP, is not influenced by the outcome variable, which should limit the impact of any collider bias that exists (168;169).

Methodological:

Different conclusions were reached between the MI and MID analyses. There is not yet much research on the relative benefits of MI and MID; and most of this work has been carried out in simulation studies. It is possible that enough additional information was included in the imputation model to add sufficient information to the imputed outcome variables to outweigh the 'extra variation introduced by imputing a finite number of randomly imputed' (123) outcome values, in which case the MI results may be more efficient.

In the imputation model the raw components of the childhood cognitive function score were included in the imputation model, so as to use all the available data, including the data for those individuals with scores for some but not all of the cognitive tests. As shown in Figures 3.7 – 3.11, the components of the cognitive score are not normally distributed, unlike the combined score, so it could be argued that the combined score would be more suitable for the imputation model, since there are only 14 individuals who have some but not all of the cognitive scores at age 8. However multiple imputation is robust to non-normality (170), so this should not have a large impact on the imputed values.

Part of the output from the Heckman selection model in Stata is a test of whether standard regression methods would yield biased results. It tests whether rho, the correlation between the error terms in the analysis model and the selection model, is equal to zero – if rho is equal to zero then standard regression methods would not give biased results, however if rho is significantly different to zero then standard regression methods would yield biased results. The results from each of the Heckman selection models fitted concluded that standard regression methods would yield biased results. As explained in section 2.2.4, there is no test between MAR and MNAR as it is not possible to test whether the missingness is related to data that were not observed.

As demonstrated when choosing the variables to include in the Heckman selection model, it can be difficult to find a variable that is associated with the missingness of the outcome variable but not the value of the outcome variable, as required using the stricter conditions. As shown in section 4.6.1, the selection model can influence the conclusions of the Heckman selection analysis, although the results were quite similar in section 4.8. Only one variable predictive of missingness but not associated with the outcome is needed for the selection model under the stricter conditions; other variables that predict missingness can also be included, although the model is unlikely to converge when it contains all the variables which predict missingness of the outcome variable. In the NSHD, for women none of the 54 variables considered for the imputation model were predictive of missing NART score but not the value of the NART, with only one variable predictive of missingness for the NART at a 10% level of significance.

Compared to choosing a selection model, choosing an imputation model was straight-forward, although a fair amount of work was required to follow the advice provided by van Buuren (120) and Carpenter & Plewis (113). It is important that variables that were predictive of the missing data were included, as multiple imputation relies on the assumption of MAR. Unlike in the selection model, further variables can be added without worrying about convergence; although it may increase the time taken for the imputation to run, the addition of further variables is ‘at worst neutral, and at best extremely beneficial’ (121).

The same conclusions were reached for childhood SEP in the complete case and MID analyses, however conclusions from the Heckman selection analyses sometimes differed. As it is not known whether the missing data mechanism is MAR or MNAR, it is not known whether the multiple imputation or Heckman selection results should be trusted, assuming that each method is the most appropriate for the type of missingness it allows for.

Limitations

In the Whitehall II analyses it is possible that the variables making up the childhood material deprivation question had different influences depending on when the participant was born. The same could be said of educational qualifications, with university education increasing over time in the UK (171). Childhood SEP was also subject to recall bias, as the variables forming the childhood material deprivation variable were not collected until phase 5 of the study, when the participants were aged 44-69. The recall bias is likely to differ by characteristics of the participants, both by adult memory and potentially by childhood material conditions.

Strengths

The main strength of this study is the attention paid to missing data. Imputation and selection models were developed, with a large range of variables considered, and two methods of choosing variables for the Heckman selection model were compared. Additionally two childhood SEP measures were compared; in previous studies often only one childhood SEP variable was considered, or a composite variable was created, combining many possible variables, without considering that they may have different effects, as found in this study. The childhood SEP data were collected prospectively in the NSHD, which removes the possibility of recall bias. The practice effects were adjusted for in the Whitehall II analyses.

Conclusions and implications

An effect of childhood SEP was found on crystallized cognitive function in adulthood, after adjusting for later life SEP. Future work should investigate the potential pathways through which childhood SEP could act on adulthood cognitive function, beyond later life SEP and childhood cognitive function.

It is not possible to test whether the missing data mechanism was MAR or MNAR, and the results differed between the Heckman selection analyses and the complete case and multiple imputation analyses. Therefore a simulation study was carried out to further investigate how well each of the missing data methods perform under each missing data mechanism.

Chapter 5: Simulation study

5.1 Introduction

There are two main strategies used to investigate the effect of different methods of accounting for missing data. The first strategy involves analysing data from a dataset with missing data, using the missing data methods, and comparing the results (as in chapter 4). However, the true results remain unknown, making it impossible to compare the results achieved using missing data methods to the true values. Hence, it is not possible to assess which of the missing data methods produce the least biased results.

The second strategy is to use complete datasets and then delete values by simulating the missingness according to the postulated missing data mechanism. The analyses of interest are initially carried out on the complete dataset; then using missing data techniques on the dataset with the simulated missing data, and these results are compared with the results from the complete data analyses. One limitation of this method is that in practice the missing data mechanism is generally unknown.

When both the true coefficients and the estimates from the various missing data methods are known (as in the second method above), the performance of different missing data methods under various missing data mechanisms can be evaluated, using measures of bias, coverage and accuracy.

Neither the NSHD nor Whitehall II have complete data, therefore it was necessary to simulate full datasets, and delete data according to each of the potential missing data mechanisms, using similar levels of missing data to those observed in the real dataset. The NSHD was chosen as the study on which to base the simulation study, as it contained prospective measures of childhood SEP, the main variable of interest. The simulation study was only carried out for men.

Each of the three missing data methods used in this thesis was then applied to the simulated datasets, and the results were compared using measures of accuracy, coverage and bias, as described in section 5.3. Collins et al (121) highlighted the importance of examining evaluation criteria beyond bias, as the results may differ across criteria. Knowing how these methods for dealing with missing data perform on simulated data, with known missing data mechanisms, allows for a better interpretation of the results in chapter 4.

5.2 Simulation

5.2.1 Protocol for carrying out a simulation study

Following the advice of Burton et al (172), the simulation study was planned in detail before being carried out. The aim of the simulation study was to compare three methods of dealing with missing data (complete case, multiple imputation and Heckman selection) under different missing data mechanisms (MCAR, MAR and MNAR).

A complete dataset was simulated, using the NSHD as the motivating example. Data were simulated to closely represent the structure of the real dataset. The three different mechanisms of missingness were then applied to the complete dataset, and the analyses carried out for each missing data method on each type of dataset (Figure 5.1). The Heckman selection analyses were run for both of the selection models developed in section 4.6. The results were then compared to the true results from the simulated complete datasets.

The number of simulations required to achieve the aim of the study was calculated using the following equation (172):

$$B = \left(\frac{Z_{1-(\alpha/2)}\sigma}{\delta} \right)^2$$

where δ is the specified level of accuracy required for the estimate ('the permissible difference from the true value β' (172)), here chosen to be 5% of the variable's coefficient, and σ^2 is the variance of the parameter of interest.

A realistic estimate of the variance was obtained from the observed NSHD data. The number of simulations required to estimate each coefficient to an accuracy level of 5% was calculated, with the highest value chosen for the number of simulations required. The highest number of simulations required was 1,428 (Table 5.1).

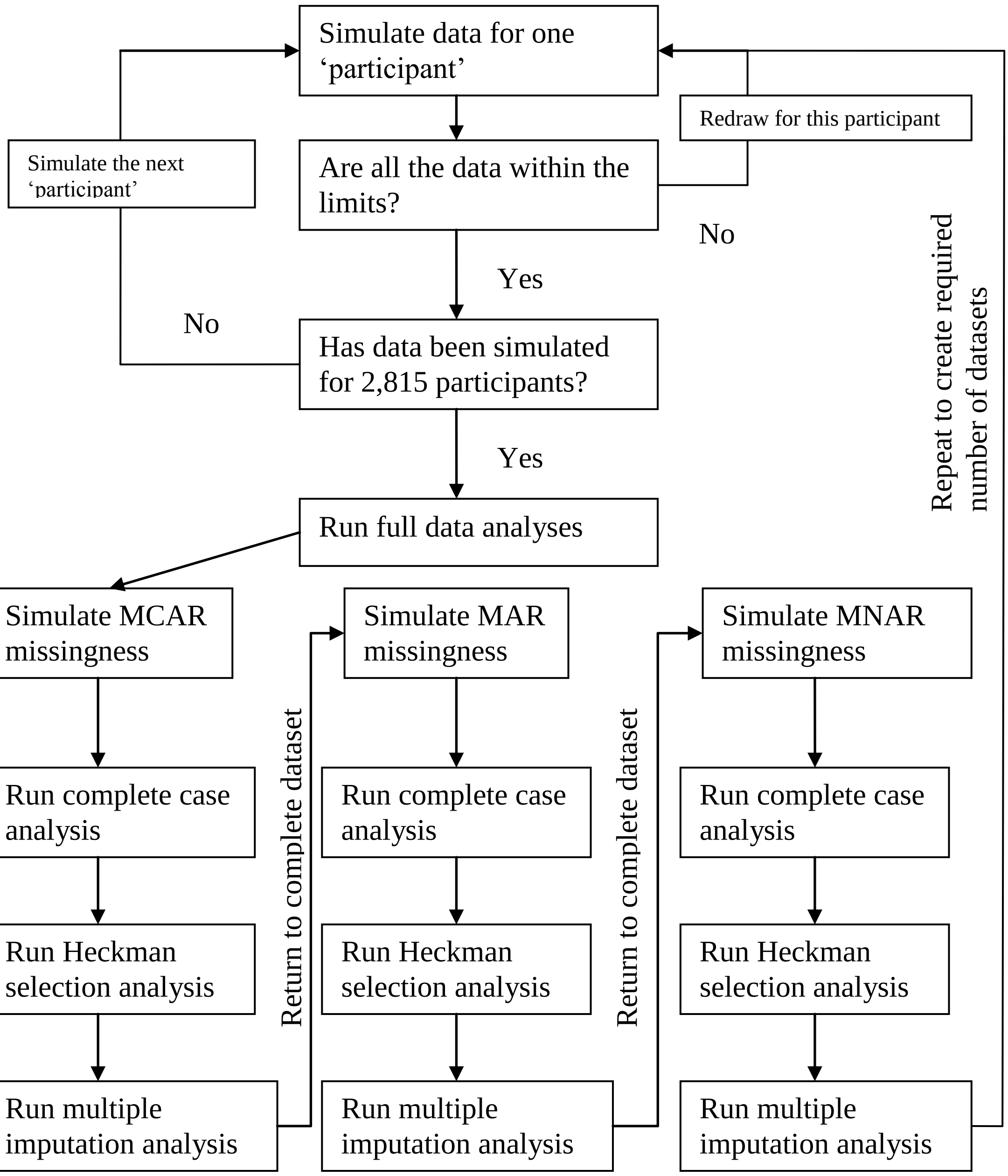


Figure 5.1: Simulation study flowchart

Table 5.1: Number of simulations required for a 5% level of accuracy

Variable	δ	σ	B
Father's Occupational SEP at Age 4	0.032	0.616	**1427.66**
Educational qualifications:			
up to GCE 'O'-Level	0.118	0.890	218.98
GCE 'A'-Level	0.161	0.826	101.04
Degree	0.318	0.939	33.58
Own Occupational SEP at Age 43	0.145	0.729	97.32
Cognitive Function at Age 8	0.212	0.308	8.08
(Cognitive Function at Age 8)^2	-0.014	0.228	954.65

5.2.2 Simulating the complete dataset

As multiple imputation and Heckman selection analyses were to be investigated, it was important to simulate all the variables which were included in the multiple imputation and Heckman selection models for men in Chapter 4. Although Chapter 4 considers 8 models, the simulations only addressed one analysis, containing father's occupational SEP, educational qualifications and own occupational SEP at age 43.

The list of variables which were simulated is in Table 5.2, by age at data collection. As each of the variables was simulated from a joint normal distribution, it was important to check whether each of the continuous variables was normally distributed. As the sample size in the NSHD is large, the standard tests for normality of the observed variables, such as the skewness-kurtosis test and the Shapiro-Wilk test, were oversensitive (129), and normality was thus assessed visually using histograms. For those variables which were possibly not normally distributed on inspection, the untransformed variable was compared to transformed variables (log, inverse, square-root and squared) via probability plots to assess whether the transformation improved the normality of the distribution. The square-root of the reading score at age 11 and square of the mathematics score at age 15 were simulated from a normal distribution, and the simulated variables were then reverse-transformed back to their original form. The categorical variables were categorised after simulating from a normal distribution. This was in order to allow all variables, categorical as well as continuous, to be correlated with each other.

Table 5.2: Variables included in the simulation study

Phase of data used in this study	Participant's Age	Variables
1	0	Father's occupational SEP
2	2	Childhood household amenities
2	4	Father's occupational SEP, child has own bed, clothes repair, yard or garden, child's shoes, cleanliness of house, cleanliness of child, mother's management and understanding of the child, crowding, repair of dwelling
3	8	Reading score, sentence completion score, picture intelligence score, vocabulary score, (cognitive score at age 8)[2]
3	10	Difficult to discipline
3	11	Father's occupational SEP, non-verbal reasoning score, reading score, verbal reasoning score, arithmetic score, vocabulary score, household amenities
3	13	Daredevil, have to stick up for myself
3	15	Non-verbal reasoning score, reading score, verbal reasoning score, mathematics score, truancy during the past year
4	26	Educational qualifications, own occupational SEP, happiness, job stress, brought up in a faith, currently have a religion
5	36	Marital status, current smoking status, ever been unemployed, own occupational SEP, ever had cancer, present state examination total, does religious upbringing effect life now
5	43	Own occupational SEP, how many friends or relatives could you visit at any time without waiting for an invite, verbal memory score
6	53	Verbal fluency score, NART

The simulations were set up using the `drawnorm` command in Stata 12. For each variable, the observed NSHD means and standard deviations for men were used as the means and standard deviations of the variables to be simulated. Similarly, the pairwise correlations, calculated from all available data for each pair of NSHD variables, were used as the correlations for the simulated variables. This enabled data to be generated from a multivariate normal distribution. To ensure the same data were not generated each time, a seed was set, with the seed dependent on the value of both the simulation number and the participant in the dataset being generated. A 'participant' in a simulated dataset was redrawn if any of the cognitive variables were outside the possible limits for that variable (see Table 5.3), for example if the simulated NART score was below 0 or above 50. To prevent the programme getting stuck in a loop when a generated dataset was redrawn, the seed increased by 1 if a generated dataset was rejected. The observed standard deviations were used for the simulations, even though this led to slightly lower simulated standard deviations.

Table 5.3: Limits for simulated variables

	Redrawn if	
Variable	Less than	More than
NART (age 53)	0	50
Verbal Fluency (age 53)	0	62
Verbal Memory (age 43)	0	45
Verbal score (age 15)	0	65
Reading score (age 15)	0	35
Non-verbal score (age 15)	0	65
Square-root maths score (age 15)	0	6.782
Verbal score (age 11)	0	40
Reading score squared (age 11)	0	2500
Non-verbal score (age 11)	0	40
Arithmetic score (age 11)	0	50
Vocabulary score (age 11)	0	49
Reading score (age 8)	0	49
Sentence completion score (age 8)	0	35
Picture intelligence score (age 8)	0	60
Vocabulary score (age 8)	0	40

The categorical variables were split into categories via the following steps:

1. The proportion of the sample in each category was calculated for the observed NSHD data.
2. The cut-off points for the simulated data were calculated from a cumulative normal distribution, using the means and standard deviations.
3. The simulated data were categorised according to these cut-off points.

This resulted in the creation of one complete simulated dataset, consisting of 2,815 individuals.

15 datasets were simulated for the trial simulations, which were carried out to check the simulated data were a suitable representation of the observed NSHD data. The observed means were used in the full simulation study, as the mean trial simulation means were within 3% of the observed mean for all the continuous variables. For the categorical

variables the proportion in each category of the binary and categorical variables was within 3%, a level considered sufficiently accurate.

5.2.3 Simulating MCAR missingness

In the observed NSHD dataset there were 955 different missing data patterns for men when considering the 49 variables which were simulated, with 730 of the patterns only observed once. Therefore some simplifications were required. The first simplification was to make the missingness monotone. In the observed NSHD data, the majority of participants, 72%, had monotone missingness.

The next simplification was to group the variables by year of data collection, and create phases at which to drop data. The phases were created based on the number of observations for the variables in the observed NSHD data. Father's occupational SEP collected at birth was fully observed, and was phase 1. The range of observations for the variables collected at ages 2 and 4 was 2,275 – 2,452, a small enough range to classify both these time points as phase 2. The range of observations for ages 8 to 15 was 2,040 – 2,204, so the data collected at these time points was considered phase 3. Age 26 was considered phase 4, ages 36 and 43 were considered phase 5, and age 53 was considered phase 6. The average observed dropout for each phase was then calculated using all the variables from each phase (Table 5.4).

Table 5.4: Dropout between phases (grouped ages of data collection) in the simulation study

Phase (age)	Average number of subjects missing (cumulative)	Proportion of subjects missing (cumulative)	Proportion of remaining subjects who drop out at this phase
1 (birth)	0	0	0
2 (2 - 4)	429	0.152	0.152
3 (8 - 15)	690	0.245	0.110
4 (26)	732.5	0.260	0.020
5 (36 - 43)	1191.8	0.423	0.220
6 (53)	1408.5	0.500	0.133

The final two columns of Table 5.4 show the calculated proportion of participants who needed to be dropped from the simulated dataset at each phase to ensure the data represented the NSHD. The final column was calculated by considering the probability of an individual being missing at phase $k+1$ given that they were observed at phase k, to give the proportion of simulated data to delete from those with observed data at the previous phase. The final column was calculated from:

P(Y_{k+1} missing | Y_k observed) = $\dfrac{p_{k+1} - p_k}{1 - p_k}$

where Y_k is a variable at phase k, and p_k is the probability that the variable is missing at phase k.

Five independent variables, $u_2 \ldots u_6$, were created by generating 2,815 observations from a Uniform[0,1] distribution; 1 observation for each simulated participant. These variables were used to drop participants at each phase. The uniform variables were generated using a seed in order to make the results replicable, with the seed differing by 10,000 between the five uniform variables to ensure that different seeds were used for each. If u_{2j}<0.152, j = 1…2,815, then the variables from phase 2, as defined in Table 5.2, would be set to missing for participant j. As the missing data were monotone, the data for all further waves were also set to missing for those participants. Similarly, if u_{3j}<0.110, j = 1…2,815, then the variables from phase 3, and beyond, would be set to missing for participant j. Data were dropped using the same method for u_k, $k = 4\ldots6$. A test run was then carried out on 15 datasets of 2,815 participants to check the correct proportion of missingness occurred at each phase.

5.2.4 Simulating MAR missingness

When missing data are MAR, the missingness can depend on observed values but not on unobserved values. To make the dropout realistic, logistic regression models were initially carried out on the observed NSHD data with the outcome of 'missing at phase *k*', *k=2…6*. The initial logistic regression model contained all the available variables from previous phases; backward selection was then carried out to select the variables which predicted missingness. Although in MAR, the missingness can depend on observed data from future phases, as missingness was restricted to monotone missingness in this simulation study, this cannot occur here. This gave the equations

$f_k = \log\left(\dfrac{p_k}{1 - p_k}\right)$ for *k=2…6*.

The equations were then rearranged using $p_k = \dfrac{e^{f_k}}{1 + e^{f_k}}$, to obtain an expression for the proportion with missing data at phase *k*, p_k. The p_k were then compared to randomly generated values $v_{2j} \ldots v_{6j}$ (j=1…2,815) from a Uniform[0,1] distribution, divided by a constant c_k and dropped if $p_k<v_k/c_k$. It was necessary to divide by c_k to ensure the correct proportion of missingness at each phase. This is because the p_{mk} were calculated from the observed NSHD dataset, where a phase was defined as missing if any of the variables at that phase were missing, leading to higher levels of missingness than when a complete case analysis was carried out. As the missingness was monotone, the data were also dropped if the data for that simulated individual were missing at the previous phase. The uniform variables were again generated using a seed in order to ensure the results were replicable, with the seed differing by 10,000 between the five uniform variables.

To calculate the c_k (k=2…6), the mean p_k (k=2…6) was calculated using the observed NSHD data for men. When k=2, the initial value of c_k was calculated to produce $p_k<v_k/c_k$ in 15.3% of cases. The observed mean value of p_2 was 0.358. Therefore the initial value of c_2 was 0.358/0.153 = 2.37. The remaining c_k were calculated using the same method.

Each c_k was then adjusted via an iterative process, with c_k decreasing to drop a larger proportion of the data or increasing to drop a smaller proportion of the data, until the mean proportion dropped in a trial of 15 datasets at each phase was within 0.5% of the observed proportion missing in the observed NSHD data. The final values of c_k are presented in Table 5.5 below, along with the cumulative percentage of missing data at each phase from the trial of 15 datasets generated, to be compared to the observed values in Table 5.4.

Table 5.5: Cumulative percentage of missing data for each phase of the simulation study (MAR)

Phase k	Final constant c_k	cumulative percentage missing
2	2.36	0.153
3	4.00	0.240
4	3.60	0.257
5	2.65	0.419
6	4.10	0.501

5.2.5 Simulating MNAR missingness

For missing data to be MNAR, the missingness can depend on both observed and unobserved values; however it is not possible to use logistic regression models to predict missingness based on unobserved data. Therefore the equations from the MAR missingness were used, but for the following phase, so the probability of missingness at phase *k* could depend on variables from phase *k*. In practice, this meant that the equation for MAR missingness at phase *k+1* was the equation for MNAR missingness at phase *k*. This was not possible for the final phase of MNAR missingness, so the equation for the previous phase was updated, for example the reading score at age 15 was updated to the NART score at age 53. The possible range of scores for the reading test at age 15 and the NART were different, so the coefficient was altered to reflect this. The initial values of the constants c_k were calculated and the values then adjusted via an iterative procedure, as they were when generating MAR missingness (Table 5.6).

Table 5.6: Cumulative percentage of missing data for each phase of the simulation study (MNAR)

Phase *k*	Final constant c_k	cumulative percentage missing
2	3.63	0.150
3	2.81	0.246
4	4.65	0.263
5	3.45	0.427
6	3.70	0.504

5.3 Post-simulation methods

For each of the 1,428 simulations, once the complete data had been simulated, before any data were dropped, a linear regression analysis was carried out on the full dataset, investigating the effect of childhood SEP on cognitive function at age 53, here represented by the NART, after adjusting for the same variables as in the final model in section 4.4; educational qualifications, own occupational SEP at age 43, childhood cognitive function, and (childhood cognitive function)^2. This was done to obtain the true results, for comparison with the simulation results from the analyses on the datasets with deleted data.

The same analyses were then carried out within each of the MCAR, MAR and MNAR datasets using first a complete case analysis, then two Heckman selection analyses (with the same selection models as in section 4.6) and a multiple imputation analysis (with the same imputation model as in section 4.5). The Heckman models were stopped if they had not converged after 350 iterations to ensure the programme did not get stuck, and the convergence status was recorded. The p-value for rho, the correlation between the errors in the selection model and analysis model, was also recorded to check whether the simulated datasets were MCAR, according to the test within the Heckman command in Stata. This information was stored in addition to the beta coefficients and their standard errors for each analysis.

The mean of each regression coefficient and standard error was calculated in each of the analyses for the MCAR, MAR and MNAR datasets. The between simulation standard errors (Equation 5.1) were also calculated for the regression coefficients and standard errors.

Equation 5.1: Between-simulation standard error

$$SE(\hat{\beta}_i) = \sqrt{\frac{1}{B-1}\sum_{i=1}^{B}\left(\hat{\beta}_i - \bar{\hat{\beta}}\right)^2}$$

Three evaluation criteria were used to compare the results produced by each method of dealing with missing data: bias, coverage and accuracy.

<u>Bias</u>

Bias is calculated using $Bias = \bar{\hat{\beta}} - \beta$, where $\bar{\hat{\beta}} = \sum_{i=1}^{B}\frac{\hat{\beta}_i}{B}$, and $\hat{\beta}_i$ is the regression coefficient of interest within each of the 1,428 simulations. The bias of an estimator tells us how far the average of the estimator is from the true value of the parameter it is estimating, and a low bias is therefore desired. Large biases cause problems in estimation and hypothesis tests. A positive bias would result in the rejection of the null hypothesis more frequently than it should be, increasing the Type I error rate. Schafer and Graham (173) define a level of bias to be problematic if it is 'greater than about one half of the estimate's standard error', a slightly more lenient level than Collins et al (121) who considered a standardised bias of greater than 40% to be practically significant due to the 'noticeable adverse impact on efficiency, coverage, and error rates'.

<u>Coverage</u>

Coverage represents the percentage of confidence intervals that include the parameter value, which is the percentage of times the 100(1-alpha)% CI $(\hat{\beta}_i \pm Z_{1-\alpha/2} SE(\hat{\beta}_i))$ include beta, for *i*=1..1,428. Schafer and Graham (173) state that when the coverage rate is accurate, the probability of wrongly rejecting a true null hypothesis (Type I error) will be accurate, and therefore deem values near 95 to represent adequate coverage. The procedure is deemed to be 'troublesome' if the coverage is below 90% (121).

Accuracy

The measure of accuracy chosen was the mean square error (MSE), the sum of the squared differences between the estimates and their targets, divided by the number of simulations (Equation 5.2). This is equivalent to the bias squared plus the estimate's variance, combining bias and efficiency (121).

Equation 5.2: Mean Square Error

$$MSE = \left(\bar{\hat{\beta}} - \beta\right)^2 + \left(SE(\hat{\beta})\right)^2$$

5.4 Results

In the full data analyses father's occupational SEP was not a significant predictor of adult NART score, nor was (childhood cognitive function)^2. The same overall conclusion was reached regarding significance of variables for each of the missing data methods under each of the missing data mechanisms. However the p-values were smallest in the full data analysis for both father's occupational SEP and (childhood cognitive function)^2. Although father's occupational SEP was a significant predictor in 559 (39.1%) of the full data simulations, it was significant in a maximum of 23.7% of the simulations when missing data methodology was applied, with the highest proportion for the complete case analysis of the MNAR dataset. Within the MCAR and MAR datasets the complete case analyses also had the highest proportion of occasions where father's occupational SEP was significant.

Bias

When the missing data mechanism was MCAR, the complete case analysis had the lowest bias for 6 of the 7 variables in the model of interest (Table 5.7). This was also the case in 4 of the 7 variables for the MAR dataset, and 5 of the 7 variables for the MNAR dataset; in the other cases the multiple imputation analysis had the lowest bias. The bias was often very high for both the Heckman selection models. The bias was 'problematic', that is greater than half the size of the estimate's standard error, in many of the Heckman selection estimates, in particular for five of the seven covariates when the missing data mechanism was MNAR. The occasions where the bias was greater than half the value of the estimate's standard error are in bold in Table 5.7. For childhood cognitive function, all three methods of accounting for missing data resulted in 'problematic' levels of bias when the missing data mechanism was MNAR.

Coverage

For each covariate under all three of the missing data mechanisms, the coverage was highest for the multiple imputation analyses, with the exception of adult occupational SEP in the MAR dataset, where the complete case analyses had very slightly higher coverage (Table 5.7). As values near 95% represent adequate coverage (173), the coverage of the multiple imputation analyses was adequate for each of the variables in the MCAR dataset, and for 5 out of 7 variables in the MAR and MNAR datasets, with lower rates of coverage found for the childhood cognitive function and (childhood cognitive function)^2 variables. For childhood cognitive function, the Heckman selection coverage was extremely low for the MAR and MNAR datasets. The Heckman selection models resulted in the most cases where the level of coverage was 'troublesome'. These occasions, and those for complete case and multiple imputation, are in bold in Table 5.7.

MSE

For each of the missing data methods, under each of the missing data mechanisms, the multiple imputation analyses have the lowest MSE for each of the variables, with one exception; Heckman selection model 1 (with more stringent conditions) had the lowest MSE for father's occupational SEP during childhood with MCAR missingness (Table 5.7). The MSE was often highest for the Heckman selection models.

5.5 Discussion

Main findings

The multiple imputation analyses had the lowest MSE and highest coverage, regardless of the missing data mechanism. The Heckman selection models usually had the lowest coverage, twice dropping below 20%, once when the missing data mechanism was MAR, and once when it was MNAR. The complete case analyses had the lowest bias when the missing data mechanism was MCAR, and complete case and multiple imputation generally had similar levels of bias for the MAR and MNAR datasets, with much higher bias found in the Heckman selection results. The results for the two Heckman selection models were very similar, with the selection model chosen using the stricter conditions producing slightly better results.

Comparison with other studies

There are not many simulation studies which have investigated attrition bias in longitudinal studies. With 40% attrition, Kristman et al (174) found that with MNAR missingness multiple imputation ‘tended to improve the results’ over the complete case analyses, but that the precision and coverage were not at the required level for multiple imputation to be considered an appropriate method for this kind of missingness. Kristman et al (174) also found that multiple imputation tends to produce greater precision than the complete case analysis under MAR missingness, although the estimates of the odds ratios were no better.

Marshall et al (175) investigated missing covariate data in prognostic modelling studies, and found that the regression coefficients from the complete case analyses were less biased with better coverage rates than those using multiple imputation for each of the three missing data mechanisms, although they had larger standard errors due to the smaller samples being analysed. Marshall et al concluded that for low levels of missingness (<10%), it was unclear whether the benefits of multiple imputation outweighed the simplicity of a complete case analysis, although for higher levels of missingness the advantages of multiple imputation were clear. As the simulation study carried out in this chapter had over 50% missingness in the dependent variable, this would imply that a complete case analysis would not be advised in this situation.

Young and Johnson (176) investigated the differences between complete case, multiple imputation and Heckman selection analyses using the National Survey of Families and Households, an observed dataset rather than in a simulation study, and were unable to draw conclusions beyond that some different conclusions were drawn between the results of the multiple imputation analysis and the other analyses, with the difference in standard errors affecting the significance of variables between the complete case and multiple imputation results.

It may seem surprising that the complete case results were less biased than the multiple imputation results under the MAR and MNAR missing data mechanisms. However, complete case analyses have been shown to provide unbiased estimates in situations other than MCAR. White and Carlin (177) explain that that ‘when missingness is independent of the outcome given the covariates, CC has negligible bias and MI is biased away from the null.’ Another example is when the MAR missingness occurs in the independent variables only (178). Additionally, if missing data only occurred in a dependent variable that was measured once per participant, the complete case analysis would not be biased, as long as "all the variables associated with the outcome being missing can be included as covariates" (116). Although neither of these situations occurred in this simulation study, it is possible that the covariates included in the model provided a good enough prediction of the outcome being missing.

Explanation of findings

The Heckman selection analyses in Stata test whether the missingness is MCAR by testing whether rho, the correlation between the errors in the selection model and the analysis model of interest, was equal to zero. Despite the missing data being defined as MCAR and dropped 'completely at random' in the simulation study, only one of the 1,428 simulations concluded that the MCAR dataset was in fact MCAR. This is likely to be due to the large sample size, resulting in too much power; the same reason the skewness-kurtosis test and the Shapiro-Wilk tests were oversensitive when investigating the normality of the variables prior to simulation.

The Heckman selection model is often used in the situation of women's wages. In this situation a few key variables exist which are likely to strongly predict selection into employment, whereas when studying attrition in a longitudinal study, it is unlikely that selection could be accurately predicted with only a few variables. This is especially true when dropout has occurred over a period of 53 years, with different variables likely to predict dropout at different stages of the life course. It is likely that the Heckman selection model would produce better results when selection into something occurs, rather than dropout. Some caution is therefore needed when applying a Heckman selection model to a complex dataset, where dropout occurs over a long period.

Another potential issue with the Heckman selection model is that the results are sensitive to the selection model, and the selection model analysis only uses complete cases. Although variables from ages 43 and above were not included in the selection model, it is possible that the poor performance of the Heckman selection analyses is due to the inclusion of variables in the selection model with too much missing data.

Identifying the 'correct' results from Chapter 4 in light of these results

As it is not possible to distinguish between MAR and MNAR missingness, it has been suggested that the most conservative approach is to assume that the missingness is MNAR, as many studies have found that participants who were lost to follow up had different characteristics to those who remained in the study (174). However it is possible to reduce the effect of having a MNAR missing data mechanism by including variables that explain the missingness or are associated with the incomplete variables in the imputation model, in order to increase the likelihood of the missing data mechanism being MAR (111). As the Heckman selection models produced biased results, often with low coverage, whereas the multiple imputation results contained the lowest MSE and highest coverage of the three methods when the missing data mechanism was MAR or MNAR, the multiple imputation results would be accepted as showing the closest to the true relationship between life course SEP and cognitive function. This is despite the bias occasionally being greater than half of the estimate's standard error when the missing data mechanism was MNAR.

As the MID methodology was preferred theoretically to the MI methodology, the final results for the analyses carried out in chapter 4 are the MID results. This led to the following conclusions.

In the NSHD when father's occupational SEP was used as the measure of childhood SEP, childhood SEP remained a significant predictor for both men and women after adjusting for early adulthood and later adulthood SEP, but was fully attenuated for men after additionally adjusting for childhood cognitive function. When childhood material deprivation was used as the measure of childhood SEP it was fully attenuated by adjustment for early adult SEP for women, and fully adjusted by educational qualifications for men, although when occupational SEP was used as the early adulthood measure of SEP childhood SEP remained significant until later adult SEP was additionally adjusted for.

In the Whitehall II analyses childhood SEP remained significant after adjusting for educational qualifications and occupational SEP for men, but was fully attenuated by adjusting for educational qualifications for women.

These results were discussed in detail in section 4.10.

Limitations

Although the simulation study was based on a real dataset, it was, by necessity, only based on the observed parts of the dataset, with an unknown missing data mechanism. Therefore the means and standard deviations, as well as relationships between the variables such as correlations and equations to predict dropout were based only on the observed data, and therefore are not likely to accurately represent the true (observed and unobserved) data. When predicting dropout for the MNAR missing data mechanism in the simulation, the probability of missingness was allowed to depend on the NART score at age 53; however the existence of this relationship, as well as its magnitude, could not be verified from the observed data.

It was not possible to compare the results of multiple imputation and multiple imputation, then deletion in the simulation study due to the monotone missingness; there were no participants who had the dependent variable observed but were missing data for an independent variable; therefore the MID analysis reduced to a complete case analysis.

Although a Pearson's correlation coefficient is not necessarily appropriate for looking at the correlation between two binary variables, all the correlations were read by the programme as correlations of continuous variables, as continuous data were simulated for all the variables. Although the correlations of the binary variables in the trial simulations were slightly lower in the simulated data than the observed data, this methodology was deemed sufficient for this simulation study, and it is unlikely to influence how the missing data methods performed relative to each other.

Regenerating the continuous variables which were outside of the observed limits led to lower standard deviations for the continuous variables which slightly altered the distribution of the variables, however it was necessary to reach a balance. If the standard deviations of the continuous variables were made larger, a larger proportion of the simulated participants would need to be regenerated, as they were more likely to be outside the valid limits. Again, this should not affect how the missing data methods perform against each other.

As Marshall et al (175) point out, the generalizability of the results from any simulation study are limited, as they reflect the data from a single study, and investigate specific patterns of missing data.

Strengths

The main strength is that a simulation study was carried out, which is well suited to investigating the effect of attrition bias in longitudinal studies, as the attrition mechanism is known (174). The simulation study followed the guidelines provided by Burton et al (172), ensuring all important aspects were considered. The major advantage of all simulation studies is that the true results are known when using the full simulated datasets, before simulating missingness, allowing the results from each of the missing data methods to be compared to the true results. This also allows measures of bias, accuracy and coverage to be calculated.

The simulated data were based on a real dataset, which provided a ‘realistic framework for simulating missing data’ (175). The simulation study had a complex data structure, simulating data for a large number of variables. This was necessary to allow the imputation model to contain a wide range of variables, ensuring that all the important variables were included.

If a simulation study had not been carried out then it is possible the results of the Heckman selection model would have been considered the most appropriate, as it has the most lenient assumptions regarding missing data. However, as observed in the simulation study, this would not have been appropriate due to the low coverage and high bias found in the Heckman selection results when the missing data mechanism was MAR or MNAR, the two possible scenarios in the NSHD and Whitehall II datasets.

Conclusion

This simulation study found that the complete case analysis had the lowest bias when the missing data mechanism was MCAR, and similar levels of bias to the multiple imputation results when the missing data mechanism was MAR and MNAR. However the Heckman selection results had the highest level of bias and lowest levels of coverage for all three missing data mechanisms. Although multiple imputation was found to be the best method of the three considered for MNAR missingness, the bias was still beyond the acceptable limit for some of the variables, showing that none of the methods considered can fully solve the problems caused by missing data. It is therefore important to attempt to keep the dropout as low as possible, and to collect as much information as possible on the reasons for missingness, in order to increase the possibility of the missingness being MAR. The Heckman selection model using the stricter conditions provided better results, so that selection model was used in future analyses in this thesis.

Table 5.7: Comparison of missing data methodology using a simulation (Heckman 1: stricter conditions, Heckman 2: more lenient conditions)

	Father's occupational SEP in childhood			Education - up to GCE 'O'-Level			Education - GCE 'A'-Level			Education - Degree		
	Coverage (%)	Bias	MSE	Coverage (%)	Bias	MSE	Coverage (%)	Bias	MSE	Coverage (%)	Bias	MSE
MCAR												
CC	93.6	-0.014	0.158	95.2	0.009	0.267	94.8	-0.001	0.255	94.1	0.004	0.814
MI	95.9	0.007	0.138	97.8	0.180	0.142	96.8	0.072	0.181	97.3	0.065	0.506
Heckman 1	94.4	0.014	0.111	93.4	0.119	0.282	92.5	0.174	0.287	91.5	**0.258**	0.877
Heckman 2	94.0	0.013	0.158	93.5	0.118	0.278	92.5	0.172	0.287	91.1	**0.258**	0.887
MAR												
CC	94.9	-0.012	0.151	94.5	-0.008	0.304	94.5	-0.015	0.289	**88.8**	-0.055	1.179
MI	95.9	-0.017	0.130	97.8	0.165	0.141	96.4	0.042	0.200	94.9	0.010	0.623
Heckman 1	94.4	0.047	0.149	92.5	0.159	0.320	**85.4**	**0.316**	0.389	**78.6**	**0.546**	1.428
Heckman 2	95.0	0.050	0.148	91.9	0.195	0.334	**82.8**	**0.379**	0.432	**74.0**	**0.652**	1.563
MNAR												
CC	96.2	0.003	0.152	95.1	0.081	0.287	92.9	0.113	0.299	**86.7**	0.068	1.235
MI	96.8	0.015	0.131	96.8	**0.217**	0.161	96.6	0.122	0.205	94.7	0.108	0.620
Heckman 1	95.4	0.049	0.151	91.5	**0.214**	0.326	**84.0**	**0.363**	0.420	**77.8**	**0.505**	1.486
Heckman 2	94.7	0.067	0.153	**89.1**	**0.258**	0.349	**80.1**	**0.438**	0.479	**73.1**	**0.629**	1.650

	Age 43 own occupational SEP - non-manual			Standardised age 8 cognitive score			Standardised age 8 cognitive score squared		
	Coverage (%)	Bias	MSE	Coverage (%)	Bias	MSE	Coverage (%)	Bias	MSE
MCAR									
CC	93.1	0.005	0.168	94.3	0.002	0.031	94.5	-0.003	0.014
MI	94.5	0.041	0.157	95.7	0.006	0.028	96.6	-0.009	0.012
Heckman 1	91.5	0.149	0.192	**72.5**	**0.205**	0.081	94.7	-0.014	0.014
Heckman 2	91.2	0.155	0.197	**75.2**	**0.190**	0.075	94.0	-0.013	0.014
MAR									
CC	95.0	0.019	0.152	91.0	0.050	0.039	**88.2**	**-0.063**	0.021
MI	94.8	0.057	0.146	92.9	0.033	0.034	92.6	-0.008	0.016
Heckman 1	**87.4**	**0.254**	0.214	**19.6**	**0.490**	0.280	**89.8**	0.030	0.018
Heckman 2	**83.0**	**0.318**	0.252	**15.3**	**0.521**	0.312	91.4	0.024	0.018
MNAR									
CC	**94.5**	0.068	0.144	**80.6**	**0.167**	0.061	**79.4**	**-0.119**	0.031
MI	94.7	0.087	0.139	**85.5**	**0.140**	0.050	92.4	-0.046	0.017
Heckman 1	**88.7**	**0.233**	0.196	**14.4**	**0.524**	0.315	**88.7**	-0.062	0.020
Heckman 2	**83.9**	**0.306**	0.234	**9.0**	**0.577**	0.371	**89.1**	-0.064	0.021

Chapter 6: The life course effect of SEP on adult crystallized cognitive function

6.1 Introduction

Life course epidemiology has been defined as the study of long-term effects of physical or social exposures during gestation, childhood, adolescence, young adulthood and later adult life on later health or disease risk (86). Life course hypotheses fall into two general categories: accumulation and critical period. The accumulation hypothesis assumes that the cumulative time with a certain exposure is associated with the outcome, regardless of when the exposure occurred. The critical period hypothesis states that it is only the exposure at a certain point in time which influences the outcome, and assumes that once that critical period has passed, it is no longer possible to alter the implications for later health. The idea of a sensitive period is similar to that of a critical period, however it allows for smaller effects to occur at times outside of the critical period.

In the social sciences, there has long been an interest in the effects of social mobility on health (87). Social mobility models are less well defined empirically than accumulation and critical period models, and unlike these, were not defined by Kuh et al in their glossary of life course epidemiology (86). In the literature, social mobility is often investigated by using binary SEP measures at two time points, and considering whether the participant is in the lower category at both time points, the higher category at both time points, upwardly mobile or downwardly mobile (55). This is similar to an accumulation analysis, but splits those who were in the higher category at one time point into those who were upwardly mobile and those who were downwardly mobile. It is, however, usually observed in such analyses, that those with mobility show health intermediate to the always high and always low groups, consistent with an accumulation model. Hence, it is not possible to distinguish accumulation from mobility. Alternatively, social mobility has been considered to imply that an interaction exists between the exposures at (a minimum of) two different time points i.e. that the impact of downward mobility, for example, is greater than just the sum of the effects from the 2 time points. These can also be thought of as sensitive period models, with later effect modification (89).

This chapter investigates the effect of life course SEP on crystallized cognitive function in older age. Two methods are compared in this chapter; first a standard method of model selection, using linear regression models with backwards selection, to identify which of the life course hypotheses are supported by the data through examining which variables remain in the model, and the coefficients of these variables. The second method uses the life course methodology developed by Mishra et al (89), which tests the life course models against a saturated model. The hypotheses supported by each method are compared, and the relative benefits of each method discussed. This chapter reports the complete case analyses and the next chapter considers missing data and the issue of the different lengths of time spent in each stage of the life course.

6.2 Method 1: Linear regression, using backwards selection

Linear regression analyses were carried out using backwards selection. This started from the saturated model, containing each of the three main SEP effects, the three two-way interactions, and the one three-way interaction between the three SEP measures. For women, as two SEP variables were considered for each time point, there were eight models; for men there were four models, as head of household SEP was the same as own SEP at age 43 (as in section 4.4). The variables were dichotomised, as this was necessary for the life course analyses using the methodology of Mishra et al (89) (Method 2), and the aim was to compare the results of the two methods. SEP at the same time as the outcome variable was not used to ensure that all the SEP variables were measured prior to the measure of cognitive function.

6.2.1 NSHD specific methodology

Father's occupational SEP, own occupational SEP at age 26, and own and head of household's occupational SEP at age 43 were dichotomised with state 1 defined as Registrar General levels I, II and IIINM, the non-manual occupations, and state 0 defined as Registrar General levels IIIM, IV and V, the manual occupations. Childhood material deprivation was dichotomised, with participants who had access to all three of the amenities (running hot water, their own kitchen and their own bathroom) in the more advantaged group (state 1). Educational qualifications were dichotomised, so GCE 'O' level (or equivalent) or below were considered to be the lower SEP group, and those who attained qualifications at GCE 'A' level, Burnham B level or above, were considered to be in the higher SEP group (state 1). The outcome variable was the NART score at age 53. Only the SEP variables were included in the model when backward selection was carried out.

6.2.2 Whitehall II specific methodology

In Whitehall II, father's occupational SEP was dichotomised with state 1 defined as Registrar General levels I, II and IIINM and state 0 defined as Registrar General levels IIIM, IV and V. Childhood material deprivation was dichotomised; participants who experienced zero to two of the four situations used to create the childhood material deprivation variable (parents were unemployed during their childhood, family had financial problems during their childhood, family did not have an outside toilet, and family did not have a car) were in state 1, the higher SEP category. Educational qualifications were dichotomised so having a university degree or above was considered higher SEP (state 1). This is a different cutoff to the NSHD due to the different characteristics of the cohorts. Adult SEP was measured using the variable last known employment grade at phase 7. This was dichotomised by considering Unified Grades 1-6 positions as higher SEP (state 1), and the remainder as lower SEP (state 0). The outcome of interest was the Mill Hill test score at phase 9. All the analyses for Whitehall II were adjusted for age at phase 9, and the number of times the cognitive tests had previously been taken.

6.3 Method 2: Life course methodology

Using the methodology described in Mishra et al (89), the life course influence of SEP on crystallized cognitive function in older age was examined. The methodology is outlined below. Initially only the life course SEP variables were included in the models, without additional covariates. The SEP variable at each time point was a binary variable, otherwise the method would become too complex, and there would not be sufficient data within each trajectory. For example, with 3 categories at each time point there would be 27 possible trajectories, and within the NSHD, half of the trajectories already contain fewer than 100 participants.

The intention of this section is to test the different life course hypotheses described above. When only one life course hypothesis is tested it is possible that it will be supported; however if a different life course hypothesis had been tested, it too may have been supported; therefore it is also important to compare the models. One method of doing this involves comparing the model relating to each of the hypotheses to a saturated model, and carrying out partial F-tests.

The explanatory life course SEP variables are binary variables S_i, with S_i equal to 0 when the participant is in the lower SEP group at time point *i*, and S_i equal to 1 when the participant is in the higher SEP group at time point *i*. S_iS_j represents the interaction term resulting from being in the higher SEP group at both time point *i* and time point *j*, and $S_iS_jS_k$ represents the three-way interaction term resulting from being in the higher SEP group at all three stages of the life course.

The saturated model (Equation 6.1) consists of main effects for SEP at each of the three time points, as well as the three two-way interactions and the single three-way interaction. This model assumes that each of the 8 possible life course SEP trajectories has a different mean for outcome Y. The other models are nested models of the saturated model, with explicit constraints on the parameters relating to different life course hypotheses.

Equation 6.1

$$E(Y) = \alpha + \beta_1 S_1 + \beta_2 S_2 + \beta_3 S_3 + \theta_{12} S_1 S_2 + \theta_{23} S_2 S_3 + \theta_{13} S_1 S_3 + \theta_{123} S_1 S_2 S_3$$

This is equivalent to the parameterisation in Equation 6.2, which expresses the saturated model in terms of social mobility.

Equation 6.2

$$E(Y) = \alpha + \delta_{12} D_{12} + \gamma_{12} U_{12} + \delta_{23} D_{23} + \gamma_{23} U_{23} + \psi_1 D_{12} U_{23} + \psi_2 U_{12} D_{23} + \eta S_1 S_2 S_3$$

where $D_{j,j+1}$ is a binary indicator for a downward change in social class ($S_j = 1$, $S_{j+1} = 0$) and $U_{j,j+1}$ is a binary indicator for an upward change in social class ($S_j = 0$, $S_{j+1} = 1$).

The accumulation model hypothesises that the longer spent in the higher SEP group, the better the outcome (which in this situation is equivalent to a higher cognitive score), regardless of which parts of the life course were spent in the higher SEP group. The possible values for the accumulated time spent in a higher SEP group are 0 (in the lower SEP group at all three time points) to 3 (in the higher SEP group at all three time points). The three S_j (j=1,2,3) therefore have the same coefficient in the model, resulting in Equation 6.3, which is equivalent to testing the constraints that $\beta_1 = \beta_2 = \beta_3$, and $\theta_{12} = \theta_{23} = \theta_{13} = \theta_{123} = 0$ in Equation 6.1.

Equation 6.3

$$E(Y) = \alpha + \beta(S_1 + S_2 + S_3)$$

In addition to the accumulation model mentioned in Mishra et al (89), two other accumulation hypotheses were tested here. An alternative accumulation model hypothesis is that the exposure at each of the time points has an effect on the outcome, but that the influences do not need to be of the same magnitude. This is equivalent to mutually adjusting for SEP at the other time points, and is equivalent to the analyses in chapter 4, although the variables in the analyses in this section have been dichotomised. This model is equivalent to testing the constraint $\theta_{12} = \theta_{23} = \theta_{13} = \theta_{123} = 0$ in Equation 6.1, and results in Equation 6.4.

Equation 6.4

$$E(Y) = \alpha + \beta_1 S_1 + \beta_2 S_2 + \beta_3 S_3$$

An adult accumulation model was also considered, which tested a similar hypothesis to the accumulation model, but considered only early adulthood and adult SEP as important, and assumed that childhood SEP had no impact after considering the accumulated effect of SEP in early adulthood and adulthood. This resulted in Equation 6.5, which is equivalent to testing the constraints that $\beta_2 = \beta_3$, and $\beta_1 = \theta_{12} = \theta_{23} = \theta_{13} = \theta_{123} = 0$ in Equation 6.1.

Equation 6.5

$$E(Y) = \alpha + \beta(S_2 + S_3)$$

The critical period model hypothesises that there is one time point at which the exposure affects the outcome of interest, and that the level of the exposure at other time points will have no impact on the outcome. This is modelled by Equation 6.6, with i representing the life course stage of interest ($i=1, 2, 3$), and is equivalent to testing the constraints $\beta_k = \theta_{12} = \theta_{23} = \theta_{13} = \theta_{123} = 0$ for all $k \neq i$.

Equation 6.6

$$E(Y) = \alpha + \beta_i S_i$$

The remaining models relate to social mobility. The first social mobility model investigated was inter-generational mobility; either upward or downward mobility between childhood SEP, when the parent's SEP was measured, and early adulthood SEP, when the participant's own SEP was measured. No assumption was made relating to the relative size of the upward and downward mobility effects.

An individual was defined as being inter-generationally upwardly mobile if they were in the lower SEP group in childhood and the higher SEP group in early adulthood, and inter-generationally downwardly mobile if they were in the higher SEP group in childhood and the lower SEP group in early adulthood. Using the notation in Equation 6.2, this implies the constraints $\delta_{23} = \gamma_{23} = \psi_1 = \psi_2 = \eta = 0$, resulting in Equation 6.7.

Equation 6.7

$$E(Y) = \alpha + \delta_{12} D_{12} + \gamma_{12} U_{12}$$

As $D_{12}= S_1(1- S_2)$ and $U_{12} = (1- S_1)S_2$, this is equivalent to Equation 6.8, with constraints $\beta_1 + \beta_2 = - \theta_{12}$ and $\beta_3 = \theta_{13} = \theta_{23} = \theta_{123}=0$ using the notation in Equation 6.1.

Equation 6.8

$$\begin{aligned} E(Y) &= \alpha + \delta_{12}(S_1(1-S_2)) + \gamma_{12}((1-S_1)S_2) \\ &= \alpha + \delta_{12}S_1 - \delta_{12}S_1S_2 + \gamma_{12}S_2 - \gamma_{12}S_1S_2 \\ &= \alpha + \delta_{12}S_1 + \gamma_{12}S_2 - (\delta_{12} + \gamma_{12})S_1S_2 \\ &= \alpha + \beta_1 S_1 + \beta_2 S_2 + \theta_{12} S_1 S_2 \end{aligned}$$

The next social mobility model investigates intra-generational social mobility; that is upward or downward mobility between early adulthood and later adulthood, with the participant's own SEP measured at both time points. No assumption was made relating to the relative size of the upward and downward mobility effects.

An individual was defined as being intra-generationally upwardly mobile if they were in the lower SEP group in early adulthood, and the higher SEP group in middle age. Using the notation in Equation 6.2, this implies the constraints $\delta_{12} = \gamma_{12} = \psi_1 = \psi_2 = \eta = 0$, resulting in Equation 6.9.

Equation 6.9

$$E(Y) = \alpha + \delta_{23} D_{23} + \gamma_{23} U_{23}$$

As $D_{23} = S_2(1- S_3)$ and $U_{23} = (1- S_2)S_3$, this is equivalent to Equation 6.10, with constraints $\beta_2 + \beta_3 = - \theta_{23}$ and $\beta_1 = \theta_{13} = \theta_{12} = \theta_{123}=0$ using the notation in Equation 6.1.

Equation 6.10

$$\begin{aligned} E(Y) &= \alpha + \delta_{23}(S_2(1-S_3)) + \gamma_{23}((1-S_2)S_3) \\ &= \alpha + \delta_{23}S_2 - \delta_{23}S_2S_3 + \gamma_{23}S_3 - \gamma_{23}S_2S_3 \\ &= \alpha + \delta_{23}S_2 + \gamma_{23}S_3 - (\delta_{23} + \gamma_{23})S_2S_3 \\ &= \alpha + \beta_2S_2 + \beta_3S_3 + \theta_{23}S_2S_3 \end{aligned}$$

The third social mobility model investigates any mobility. It assumes that any upward mobility has an equal effect, regardless of whether it is inter-generational or intra-generational, and similarly assumes that all downward mobility has an equal effect. No assumption was made relating to the relative size of the upward and downward mobility effects.

An individual was defined as upwardly mobile if they were either inter-generationally or intra-generationally upwardly mobile. Similarly, someone was downwardly mobile if they were either inter-generationally or intra-generationally mobile. As SEP is a binary variable, it is not possible to be both inter-and intra-generationally upwardly mobile or inter- and intra-generationally downwardly mobile; however unlike in the previous mobility models, it is possible to be both upwardly and downwardly mobile. According to this hypothesis, the expected cognitive score would be the same regardless of whether they were first upwardly or downwardly mobile. Using the notation from Equation 6.2, this corresponds to the constraints $\delta_{12} = \delta_{23}$, $\gamma_{12} = \gamma_{23}$ and $\psi_1 = \psi_2 = \eta = 0$, resulting in Equation 6.11. This is equivalent to Equation 6.12, using the notation from Equation 6.1, with the constraints $\beta_2 = \beta_1 + \beta_3 = - \theta_{12} = - \theta_{23}$ and $\theta_{13} = \theta_{123} = 0$.

Equation 6.11

$$E(Y) = \alpha + \delta(D_{12} + D_{23}) + \gamma(U_{12} + U_{23})$$

Equation 6.12

$$\begin{aligned} E(Y) &= \alpha + \delta(S_1(1-S_2) + S_2(1-S_3)) + \gamma((1-S_1)S_2 + (1-S_2)S_3) \\ &= \alpha + \delta S_1 + \delta S_2 - \delta S_1 S_2 - \delta S_2 S_3 + \gamma S_2 + \gamma S_3 - \gamma S_1 S_2 - \gamma S_2 S_3 \\ &= \alpha + \delta S_1 + (\delta + \gamma) S_2 + \gamma S_3 - (\delta + \gamma) S_1 S_2 - (\delta + \gamma) S_2 S_3 \\ &= \alpha + \beta_1 S_1 + \beta_2 S_2 + \beta_3 S_3 + \theta_{12} S_1 S_2 + \theta_{23} S_2 S_3 \end{aligned}$$

An additional social mobility model was considered beyond those described in Mishra et al (89). This extends the third social mobility model by additionally allowing the outcome to differ between those who remain in the lower social class for all three time points, and those who remain in the higher social class for all three time points. This adds the three-way interaction to the model, but maintains the restriction that the coefficients for upward mobility remain the same, regardless of whether it was inter- or intra-generational, and similarly for downward mobility. Using the notation from Equation 6.2, this corresponds to the constraints $\delta_{12} = \delta_{23}$, $\gamma_{12} = \gamma_{23}$ and $\psi_1 = \psi_2 = 0$, resulting in Equation 6.13. This is equivalent to Equation 6.14, using the notation from Equation 6.1, with the constraints $\beta_2 = \beta_1 + \beta_3 = -\theta_{12} = -\theta_{23}$ and $\theta_{13} = 0$.

Equation 6.13

$$E(Y) = \alpha + \delta(D_{12} + D_{23}) + \gamma(U_{12} + U_{23}) + \eta S_1 S_2 S_3$$

Equation 6.14

$$\begin{aligned} E(Y) &= \alpha + \delta(S_1(1-S_2) + S_2(1-S_3)) + \gamma((1-S_1)S_2 + (1-S_2)S_3) + \eta S_1 S_2 S_3 \\ &= \alpha + \delta S_1 + \delta S_2 - \delta S_1 S_2 - \delta S_2 S_3 + \gamma S_2 + \gamma S_3 - \gamma S_1 S_2 - \gamma S_2 S_3 + \eta S_1 S_2 S_3 \\ &= \alpha + \delta S_1 + (\delta + \gamma) S_2 + \gamma S_3 - (\delta + \gamma) S_1 S_2 - (\delta + \gamma) S_2 S_3 + \eta S_1 S_2 S_3 \\ &= \alpha + \beta_1 S_1 + \beta_2 S_2 + \beta_3 S_3 + \theta_{12} S_1 S_2 + \theta_{23} S_2 S_3 + \theta_{123} S_1 S_2 S_3 \end{aligned}$$

Each of the above models was then tested against the saturated model using a partial F-test. The null hypothesis was that there was no difference between the fit of the hypothesised life course model of interest and the saturated model, and the alternative hypothesis was that there was a difference between the fit of the model of interest and the saturated model. If $p<0.05$, then there was sufficient evidence to reject the null hypothesis and conclude that there was a difference between the fit of the model of interest and the saturated model. This would mean that the model of interest was not as good a fit as the saturated model. However if $p \geq 0.05$, then there was not sufficient evidence to reject the null hypothesis, and the null hypothesis was therefore accepted. This lead to the conclusion that, as there was no difference between the fit of the model of interest and the saturated model, the simpler model would suffice.

Predicted values of the mean cognitive score for each of the eight possible trajectories under each of the life course hypotheses were calculated, to compare the accuracy of predictions for the different models.

6.3.1 Life course analyses adjusting for confounders

It was also possible to adjust for other variables in the life course methodology. Therefore cognitive function at age 8 and (cognitive function at age 8)^2 were adjusted for in the NSHD analyses, and age at phase 9 and the number of times the participants had taken the cognitive tests in the Whitehall II analyses.

6.4 Descriptive results

6.4.1 NSHD

Complete data on each of the six SEP measures and the outcome variable, as well as childhood cognitive function (as measured in the previous chapter), were available for 893 men (32%) and 955 women (37%). This sample was used to allow comparisons between the adjusted and unadjusted results.

When father's occupational SEP, educational qualifications and own adulthood occupational SEP were used as the SEP measures (Table 6.1), the SEP trajectories with largest frequency differed for men and women. For men, the largest group was those who remained in the higher SEP category at all three time points (27%), whereas for women the largest group was those who were in the lower SEP category in childhood and early adulthood, but were in the higher SEP category for own occupational SEP at age 43 (29%). Three of the trajectories contain fewer than 50 men and 50 women. For men, the group with the lowest frequency was those who were in the higher childhood SEP and educational qualifications groups, then in the lower SEP group for own occupational SEP at age 43 (2%). However for women, the group with the lowest frequency were those who were in the lower SEP category in childhood and adulthood, but were in the higher SEP category for educational qualifications (3 participants, 0%).

Table 6.1: SEP trajectory $S_1S_2S_3$ frequencies, using father's occupational SEP (S_1), educational qualifications (S_2) and own adult occupational SEP (S_3)

Trajectory $S_1S_2S_3$	Men N (%)	Women N (%)
000	204 (23)	184 (19)
001	130 (15)	280 (29)
010	47 (5)	3 (0)
011	117 (13)	80 (8)
100	47 (5)	48 (5)
101	86 (10)	165 (17)
110	19 (2)	9 (1)
111	243 (27)	186 (19)
Total	893 (100)	955 (100)

As the stratifying variable for the data collection in the NSHD was father's occupational SEP at birth, different numbers of participants started in the lower and higher SEP groups. Of those men who started in the higher SEP group (1) 62% stayed in the higher group for both of the other time points, compared to 46% of women, and 41% of those men who started in the lower SEP group (0) stayed in the lower SEP group, compared to 34% of women.

6.4.2 Whitehall II

When father's occupational SEP was used as the childhood SEP measure (Table 6.2), the trajectory with the highest frequency differed for men and women. For men the largest group was those who remained in the higher SEP category at all three time points (21%), as it was in the NSHD. However for women the largest group was those who were in the lower SEP category at all three time points (38%). Half of the trajectories for women had a frequency of less than 50. The group with the lowest frequency was the same for men and women; those who were in the lower childhood SEP and adult SEP categories, but were in the higher educational qualifications category.

There were differences between the genders in social mobility; 44% of women were downwardly mobile at some point during the life course, and 13% were upwardly mobile (8% were both upwardly and downwardly mobile). However, 42% of men were downwardly mobile and 37% were upwardly mobile at some point during the life course (17% were both upwardly and downwardly mobile).

Complete data on each of the four SEP measures and the Mill Hill test score at phase 9, as well as age at phase 9 and the number of times the cognitive tests had been taken were available for 2,440 men (35%) and 826 women (24%).

Table 6.2: Whitehall II SEP trajectories, using father's occupational SEP (S_1), educational qualifications (S_2) and occupational SEP (S_3)

Trajectory $S_1S_2S_3$	Men N (%)	Women N (%)
000	416 (17)	313 (38)
001	258 (11)	22 (3)
010	79 (3)	19 (2)
011	223 (9)	22 (3)
100	459 (19)	218 (26)
101	339 (14)	46 (6)
110	151 (6)	77 (9)
111	515 (21)	109 (13)
Total	2,440 (100)	826 (100)

6.5 Method 1: Backwards selection results

6.5.1 NSHD

Men

The results from the backwards selection models for men are presented in Table 6.4 and Appendix 15. For men, the R^2 values were similar for each of the saturated models (Model 1) using different measures of SEP, although they were slightly higher when educational qualifications attained was used as the measure of early adulthood SEP rather than own occupational SEP at age 26 (Appendix 15). All three main effects remained in the model when father's occupational SEP was used as the measure of childhood SEP (Table 6.4). The coefficients for educational qualifications attained and own occupational SEP at age 26 were very similar. These results indicate that the mutually adjusted accumulation model is likely to provide a good fit to the data, as all three of the main effects remained significant predictors of the NART score at age 53. However the adult accumulation model may also provide a good fit to the data, due to the similar, larger, coefficients for educational qualifications attained/own occupational SEP at age 26 and own occupational SEP at age 43. The adult accumulation model has fewer parameters to fit than the mutually adjusted accumulation model, which may improve the model fit when comparing the models using the BIC, which penalises models with more parameters.

When childhood household amenities was used as the measure of childhood SEP, with educational qualifications attained in the model as the early adulthood measure of SEP, only the main effects of educational qualifications attained and own occupational SEP at age 43 remained in the model. The coefficients were similar, indicating that the adult accumulation model was likely to provide a good fit to the data.

However, when childhood household amenities and own occupational SEP at age 26 were included in the model, the three-way interaction between childhood household amenities, own occupational SEP at age 26, and own occupational SEP at age 43 was significant. To aid the interpretation of the three-way interaction, a table was produced (Table 6.3), showing the predicted NART score for each combination of the three SEP variables. As the table shows, childhood SEP only had an effect when early adulthood and adult SEP were in different categories. It is not clear from the results which of the life course hypotheses would be supported, although the significance of the interaction terms indicates that a social mobility hypothesis may be supported.

Table 6.3: Displaying the three-way interaction for men, NSHD (S_{1b}: childhood household amenities, S_{2b}: own occupational SEP at age 26, S_3: own occupational SEP at age 43)

		Predicted NART score	
S_{2b}	S_3	S_{1b}=0	S_{1b}=1
0	0	27.98	28.51
1	0	30.40	37.79
0	1	31.71	34.48
1	1	38.16	39.30

Women

The results from the backwards selection models for women are presented in Table 6.5 and Appendix 15. The range of R^2 values for the saturated models was much higher for women than men, ranging from 0.184 (childhood household amenities, own occupational SEP at 26, own occupational SEP at 43) to 0.282 (father's occupational SEP, educational qualifications attained, own occupational SEP at 43).

When father's occupational SEP was used as the measure of childhood SEP, and own occupational SEP at age 26 was used as the measure of early adulthood SEP, all three main effects remained as significant predictors of the NART score at age 53, with similar sized coefficients for the three time points. This indicates that the accumulation hypothesis would be appropriate.

When father's occupational SEP and educational qualifications were included in the model, each of the three main effects remained significant, as well as the interaction between father's occupational SEP and educational qualifications attained. The coefficient for educational qualifications was much larger than the coefficients for father's occupational SEP or either of the adult SEP coefficients, indicating the critical period in early adulthood hypothesis may be appropriate. However the interaction between childhood SEP and early adulthood SEP remained significant, indicating that the inter-generational social mobility hypothesis may be appropriate, or perhaps the mutually adjusted accumulation hypothesis.

When childhood household amenities was used as the measure of childhood SEP only the main effects of early adulthood and adult SEP remained significant. When own occupational SEP was used as the measure of early adulthood SEP, the coefficients for early adulthood and adult SEP were similar, implying that the adult accumulation hypothesis may fit the data best. However when educational qualifications was used as the measure of early adulthood SEP, the coefficient for educational qualifications was much larger than the coefficient for adult SEP. This indicates that the mutually adjusted accumulation hypothesis or the critical period hypothesis for early adulthood would provide the best fit to the data.

Table 6.4: NSHD backwards selection for men from the saturated model (S_1: father's occupational SEP, S_2: educational qualifications, S_3: own occupational SEP at age 43), with outcome NART

Men	Model 1		Model 2		Model 3		Model 4		Model 5	
N=893	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_1	2.32 (2.06)	0.259	2.86 (1.78)	0.108	3.17 (1.16)	0.006	2.79 (0.67)	<0.001	2.91 (0.67)	<0.001
S_2	3.48 (1.69)	0.040	3.78 (1.56)	0.016	3.76 (1.55)	0.016	3.69 (1.53)	0.016	5.18 (0.74)	<0.001
S_3	4.57 (1.10)	<0.001	4.73 (1.05)	<0.001	4.82 (0.98)	<0.001	4.87 (0.96)	<0.001	5.56 (0.81)	<0.001
S_1S_2	1.86 (2.93)	0.526	-0.88 (1.38)	0.525	-0.71 (1.37)	0.603				
S_2S_3	2.92 (1.98)	0.140	2.43 (1.73)	0.161	2.41 (1.74)	0.165	2.24 (1.71)	0.191		
S_1S_3	1.43 (2.45)	0.560	0.51 (1.89)	0.785						
$S_1S_2S_3$	-3.33 (3.31)	0.315								
Constant	27.71 (0.73)	<0.001	27.65 (0.72)	<0.001	27.61 (0.70)	<0.001	27.65 (0.69)	<0.001	27.34 (0.63)	<0.001
R^2	0.2694		0.2688		0.2687		0.2685		0.2660	
BIC	6347.56		6341.56		6334.86		6328.35		6324.56	

Table 6.5: NSHD backwards selection for women from the saturated model (S_1: father's occupational SEP, S_2: educational qualifications, S_3: own occupational SEP at age 43), with outcome NART

Women	Model 1		Model 2		Model 3		Model 4	
N=955	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_1	4.24 (1.60)	0.008	4.39 (1.53)	0.004	4.44 (0.81)	<0.001	4.43 (0.81)	<0.001
S_2	4.61 (1.24)	<0.001	6.00 (1.86)	0.001	5.97 (1.82)	0.001	9.10 (0.88)	<0.001
S_3	4.55 (0.93)	<0.001	4.59 (0.91)	<0.001	4.60 (0.82)	<0.001	4.75 (0.78)	<0.001
S_1S_2	-0.55 (3.33)	0.869	-3.10 (1.20)	0.010	-3.08 (1.16)	0.008	-3.10 (1.17)	0.008
S_2S_3	4.75 (1.55)	0.002	3.29 (1.91)	0.086	3.32 (1.90)	0.081		
S_1S_3	0.28 (1.85)	0.878	0.07 (1.73)	0.967				
$S_1S_2S_3$	-2.73 (3.57)	0.444						
Constant	26.62 (0.75)	<0.001	26.59 (0.74)	<0.001	26.59 (0.69)	<0.001	26.49 (0.68)	<0.001
R^2	0.2819		0.2817		0.2817		0.2807	
BIC	6756.98		6750.33		6743.47		6737.97	

6.5.2 Whitehall II

Men

The results from the backwards selection models are presented in Table 6.6 for father's occupational SEP, and Appendix 16 for childhood material deprivation. The R^2 values were slightly higher in the models containing father's occupational SEP than in the models with childhood material deprivation. When father's occupational SEP was used as the measure of childhood SEP, each of the three main effects plus an interaction between father's occupational SEP and adult occupational SEP remained in the model. The positive interaction term between childhood SEP and adult SEP implied that there was an additional benefit of being in the higher group at both of these times beyond the sum of the two individual benefits. For individuals in the same early adulthood SEP group there was a benefit of 2.49 (0.21+1.64+0.64) points over those who were in the lower groups for both childhood and adult SEP, and a benefit of 0.85 points over those who were in the lower childhood SEP group and the higher adult SEP group. The coefficient for adult SEP was already the largest of the three main effects, even excluding the interaction term, so adult SEP is likely to be important, perhaps through the mutually adjusted accumulation model. None of the life course models tested include the childhood SEP and adulthood SEP interaction directly, although it is included in the any mobility model with the three-way interaction, interacting with early adulthood SEP.

When childhood material deprivation was used, each of the three main effects remained in the model, as well as the interaction between childhood material deprivation and educational qualifications. The interaction term must be considered with relation to the individual terms. In this case the positive interaction term between childhood SEP and educational qualifications implies that there was an additional benefit of being in the higher group at both of these times, beyond the sum of the two individual benefits. For individuals in the same adult SEP group there was a benefit of 1.72 (0.20+0.87+0.65) points over those who were in the lower group for childhood and early adulthood SEP, and a benefit of 0.85 points over those who were in the lower childhood SEP group and the higher educational qualifications group. These results imply that the inter-generational social mobility hypothesis may be the most appropriate. However, the coefficients were low for both the main effects involved and the interaction term, compared to the size of the coefficient for adult SEP, indicating that the mutually adjusted accumulation model may be appropriate.

Women

The results from the backwards selection models for women are presented in Table 6.7 for father's occupational SEP, and Appendix 16 for childhood material deprivation. When father's occupational SEP was used as the measure of childhood SEP, all three main effects remained as significant predictors of the Mill Hill test score at phase 9, with differing coefficients. This indicated that the mutually adjusted accumulation hypothesis would be the most appropriate.

When childhood material deprivation was included in the model, the same SEP variables remained in the model as for men, although the coefficient for the interaction between childhood and early adulthood SEP was negative for women, and positive for men. As the magnitude of the significant interaction term was relatively large, the inter-generational mobility hypothesis may be appropriate, or an accumulation hypothesis.

As can be seen from the results, more than one life course hypothesis was suggested by the backwards selection methodology. The advantage of the methodology developed by Mishra et al (Method 2) is that it identifies which of the life course models are significantly different from the saturated model. These models can then be directly compared.

Table 6.6: Whitehall II backwards selection for men from the saturated model (S1: father's occupational SEP, S2: educational qualifications, S3: occupational grade at phase 1), with outcome Mill Hill test

Men (N=2,440)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_1	0.23 (0.22)	0.299	0.23 (0.22)	0.298	0.14 (0.21)	0.488	0.15 (0.21)	0.460	0.21 (0.20)	0.286
S_2	1.46 (0.41)	<0.001	1.46 (0.40)	<0.001	1.17 (0.31)	<0.001	1.06 (0.24)	<0.001	1.26 (0.15)	<0.001
S_3	1.87 (0.26)	<0.001	1.87 (0.26)	<0.001	1.75 (0.24)	<0.001	1.71 (0.22)	<0.001	1.65 (0.22)	<0.001
S_1S_2	-0.15 (0.51)	0.775	-0.15 (0.51)	0.775	0.31 (0.30)	0.309	0.31 (0.30)	0.309		
S_2S_3	-0.61 (0.50)	0.228	-0.61 (0.50)	0.228	-0.16 (0.31)	0.596				
S_1S_3	0.32 (0.35)	0.358	0.32 (0.35)	0.357	0.54 (0.29)	0.065	0.52 (0.29)	0.072	0.62 (0.27)	0.023
$S_1S_2S_3$	0.71 (0.64)	0.265	0.71 (0.63)	0.265						
Age (phase 9)	0.00 (0.01)	0.990								
No. of times taken cog. tests	0.27 (0.09)	0.002	0.27 (0.09)	0.002	0.27 (0.09)	0.002	0.27 (0.09)	0.002	0.27 (0.09)	0.002
Constant	23.00 (0.84)	<0.001	23.00 (0.34)	<0.001	23.04 (0.34)	<0.001	23.06 (0.34)	<0.001	23.03 (0.34)	<0.001
R^2	0.1636		0.1636		0.1632		0.1631		0.1627	
BIC	12810.70		12802.90		12796.35		12788.83		12782.07	

Table 6.7: Whitehall II backwards selection for women from the saturated model (S1: father's occupational SEP, S2b: own occupational SEP at age 26, S3: own occupational SEP at age 43), with outcome Mill Hill test

Women (N=826)	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_1	1.28 (0.38)	0.001	1.27 (0.38)	0.001	1.14 (0.37)	0.002	1.06 (0.36)	0.003	0.87 (0.32)	0.008	0.82 (0.32)	0.011
S_2	2.95 (1.02)	0.004	2.92 (1.02)	0.004	1.96 (0.87)	0.024	1.38 (0.51)	0.007	1.45 (0.50)	0.004	1.91 (0.40)	<0.001
S_3	3.91 (0.95)	<0.001	3.92 (0.96)	<0.001	3.08 (0.83)	<0.001	3.34 (0.77)	<0.001	2.69 (0.57)	<0.001	3.29 (0.41)	<0.001
S_1S_2	-2.04 (1.17)	0.081	-2.03 (1.17)	0.83	-0.77 (0.92)	0.405						
S_2S_3	-1.10 (1.65)	0.505	-1.10 (1.65)	0.504	1.41 (0.83)	0.089	1.42 (0.83)	0.086	1.22 (0.81)	0.133		
S_1S_3	-1.97 (1.18)	0.096	-1.96 (1.18)	0.097	-0.68 (0.93)	0.466	-1.02 (0.83)	0.221				
$S_1S_2S_3$	3.34 (1.90)	0.079	3.34 (1.90)	0.080								
Age (phase 9)	-0.07 (0.03)	0.012	-0.07 (0.03)	0.009	-0.07 (0.03)	0.009	-0.07 (0.03)	0.009	-0.07 (0.03)	0.008	-0.07 (0.03)	0.011
No. of times taken cog. tests	0.19 (0.22)	0.392										
Constant	25.85 (1.98)	<0.001	26.64 (1.75)	<0.001	26.70 (1.75)	<0.001	26.74 (1.75)	<0.001	26.91 (1.75)	<0.001	26.61 (1.74)	<0.001
R^2	0.2257		0.2250		0.2220		0.2214		0.2200		0.2178	
BIC	4813.81		4807.84		4804.23		4798.22		4793.01		4788.58	

6.6 Method 2: Life course methodology results

6.6.1 NSHD

6.6.1.1 Predicted mean score, by life course hypothesis

There was a slightly wider range of cognitive scores between the different SEP trajectories for women (26.6 to 41.8) than men (27.7 to 41.0). For men the observed mean scores fell into four categories along the lines of the accumulation groups, whereas for women there was not such a clear pattern.

The results of these analyses are in Table 6.8, with the predicted values which were within 1 point of the observed mean NART scores in bold. For women, the accumulation model and mutually adjusted accumulation model were within 1 point of the observed mean NART scores the most frequently, for five of the trajectories. At the other extreme, the adulthood critical period model, intra-generational mobility model and any mobility models were each within 1 point from the observed mean NART score for only one of the trajectories. For men, the mutually adjusted accumulation model predicted the NART score to within 1 point of the observed NART score for six of the eight trajectories. However, four of the life course models predicted the mean NART score to within 1 point of the observed NART score for only one of the trajectories (adulthood critical period, inter-generational social mobility, intra-generational social mobility and the any social mobility model).

Table 6.8: Distribution of SEP trajectories (using father's occupational SEP (S_1), educational qualifications (S_2) and own occupational SEP at age 43 (S_3)), and observed and predicted mean (se) NART scores according to the hypotheses[1]

						Mean (se) predicted NART score					
SEP						Accumulation models			Critical period		
Sex	S_1	S_2	S_3	N (%)	Observed mean NART score (se)	Accumulation	Adult accumulation	Mutually adjusted accumulation	Childhood	Early adulthood	Adulthood
Women	0	0	0	184 (19.3)	26.6 (0.7)	**26.3 (0.6)**	25.3 (0.8)	**26.6 (0.7)**	30.9 (0.4)	30.3 (0.4)	**27.4 (0.7)**
(N=955)	1	0	0	48 (5.0)	30.9 (1.4)	**31.6 (0.3)**	25.3 (0.8)	**30.1 (0.8)**	37.3 (0.5)	**30.3 (0.4)**	27.4 (0.7)
	0	1	0	3 (0.3)	31.2 (1.2)	**31.6 (0.3)**	**30.5 (0.4)**	34.4 (0.9)	**30.9 (0.4)**	40.8 (0.4)	27.4 (0.7)
	0	0	1	280 (19.3)	31.2 (0.5)	**31.6 (0.3)**	**30.5 (0.4)**	**31.4 (0.5)**	**30.9 (0.4)**	**30.3 (0.4)**	34.7 (0.4)
	1	1	0	9 (0.9)	34.9 (2.9)	37.0 (0.3)	30.5 (0.4)	37.9 (0.9)	37.3 (0.5)	40.8 (0.4)	27.4 (0.7)
	1	0	1	165 (17.3)	35.7 (0.8)	37.0 (0.3)	30.5 (0.4)	**34.9 (0.6)**	37.3 (0.5)	30.3 (0.4)	34.7 (0.4)
	0	1	1	80 (8.4)	40.5 (0.8)	37.0 (0.3)	35.6 (0.4)	39.2 (0.6)	30.9 (0.4)	**40.8 (0.4)**	34.7 (0.4)
	1	1	1	186 (19.5)	41.8 (0.4)	**42.3 (0.5)**	35.6 (0.4)	**42.7 (0.5)**	37.3 (0.5)	40.8 (0.4)	34.7 (0.4)
Men	0	0	0	204 (22.8)	27.7 (0.7)	**27.6 (0.6)**	**28.0 (0.6)**	**27.3 (0.6)**	31.8 (0.5)	30.2 (0.5)	28.7 (0.6)
(N=893)	1	0	0	47 (5.3)	30.0 (1.9)	32.3 (0.4)	28.0 (0.6)	**30.3 (0.9)**	38.0 (0.5)	**30.2 (0.5)**	28.7 (0.6)
	0	1	0	47 (5.3)	31.2 (1.5)	32.3 (0.4)	33.3 (0.3)	32.5 (0.9)	**31.8 (0.5)**	38.1 (0.5)	28.7 (0.6)
	0	0	1	130 (14.6)	32.3 (0.8)	**32.3 (0.4)**	33.3 (0.3)	**32.9 (0.7)**	**31.8 (0.5)**	30.2 (0.5)	36.9 (0.4)
	1	1	0	19 (2.1)	35.4 (1.4)	36.9 (0.3)	33.3 (0.3)	**35.4 (0.9)**	38.0 (0.5)	38.1 (0.5)	28.7 (0.6)
	1	0	1	86 (9.6)	36.0 (1.1)	**36.9 (0.3)**	33.3 (0.3)	**35.8 (0.7)**	38.0 (0.5)	30.2 (0.5)	**36.9 (0.4)**
	0	1	1	117 (13.1)	38.7 (0.6)	36.9 (0.3)	**38.7 (0.4)**	**38.1 (0.6)**	31.8 (0.5)	**38.1 (0.5)**	36.9 (0.4)
	1	1	1	243 (27.2)	41.0 (0.5)	**41.6 (0.5)**	38.7 (0.4)	**41.0 (0.5)**	38.0 (0.5)	38.1 (0.5)	36.9 (0.4)

[1] Predicted values within 1 point of the observed mean NART scores are in bold

Sex	SEP S_1	S_2	S_3	N (%)	Observed mean NART score (se)	Social mobility: Inter-generational mobility	Intra-generational mobility	Any mobility	Any mobility with 3-way interaction
Women									
(N=955)	0	0	0	184 (19.3)	26.6 (0.7)	31.1 (0.4)	32.9 (0.6)	30.8 (0.6)	**26.9 (0.7)**
	1	0	0	48 (5.0)	30.9 (1.4)	34.3 (0.7)	32.9 (0.6)	32.6 (0.9)	**30.1 (0.9)**
	0	1	0	3 (0.3)	31.2 (1.2)	40.2 (0.8)	33.3 (1.7)	34.9 (0.7)	36.1 (0.7)
	0	0	1	280 (19.3)	31.2 (0.5)	**31.1 (0.4)**	**32.1 (0.5)**	33.2 (0.5)	32.9 (0.5)
	1	1	0	9 (0.9)	34.9 (2.9)	31.1 (0.4)	33.3 (1.7)	32.6 (0.9)	30.1 (0.9)
	1	0	1	165 (17.3)	35.7 (0.8)	34.3 (0.7)	32.1 (0.5)	**34.9 (0.7)**	**36.1 (0.7)**
	0	1	1	80 (8.4)	40.5 (0.8)	**40.2 (0.8)**	32.9 (0.6)	33.2 (0.5)	32.9 (0.5)
	1	1	1	186 (19.5)	41.8 (0.4)	31.1 (0.4)	32.9 (0.6)	30.8 (0.6)	**41.8 (0.4)**
Men									
(N=893)	0	0	0	204 (22.8)	27.7 (0.7)	32.2 (0.5)	33.7 (0.5)	32.2 (0.6)	**28.2 (0.7)**
	1	0	0	47 (5.3)	30.0 (1.9)	33.8 (1.0)	33.7 (0.5)	**30.5 (1.0)**	27.7 (1.0)
	0	1	0	47 (5.3)	31.2 (1.5)	36.4 (0.7)	**31.6 (1.4)**	33.6 (0.9)	34.5 (0.9)
	0	0	1	130 (14.6)	32.3 (0.8)	**32.2 (0.5)**	33.3 (0.7)	35.3 (0.5)	35.0 (0.5)
	1	1	0	19 (2.1)	35.4 (1.4)	32.2 (0.5)	31.6 (1.4)	30.5 (1.0)	27.7 (1.0)
	1	0	1	86 (9.6)	36.0 (1.1)	33.8 (1.0)	33.3 (0.7)	33.6 (0.9)	34.5 (0.9)
	0	1	1	117 (13.1)	38.7 (0.6)	36.4 (0.7)	33.7 (0.5)	35.3 (0.5)	35.0 (0.5)
	1	1	1	243 (27.2)	41.0 (0.5)	32.2 (0.5)	33.7 (0.5)	32.2 (0.6)	**41.0 (0.5)**

6.6.1.2 Life course hypotheses

Men

Only accumulation hypotheses were not significantly different from the saturated model for all the SEP combinations (Table 6.9 and Table 6.10). For all the SEP combinations the mutually adjusted accumulation model was not significantly different from the saturated model. For two of the SEP combinations, other accumulation models were also not significantly different from the saturated model. When father's occupational SEP and own occupational SEP at 26 were included in the model, the accumulation hypothesis was also not significantly different from the saturated model, and had the lower BIC. There was an increase of 4.22 (95% CI: 3.66, 4.77) points in the NART score for each stage of the life course spent in the higher SEP category. When childhood household amenities and educational qualifications were included in the model the adult accumulation model was not significantly different from the saturated model, and also had a lower BIC than the mutually adjusted accumulation model. Each additional time point spent in the higher SEP for the two adult SEP measures increased the NART score by 5.88 points (95% CI: 5.10, 6.66).

Women

The same final models were selected for own (Table 6.9 and Table 6.10) and head of household (Table 6.11 and Table 6.12) occupational SEP at age 43; therefore only the results using own occupational SEP are discussed here. As for men, only accumulation models were not significantly different from the saturated model. In analyses using father's occupational SEP and educational qualifications (Table 6.9), all of the life course models were significantly different from the saturated model, showing that the relationship between life course SEP and crystallized cognitive function was more complex than any of the life course models allowed for.

When childhood household amenities and educational qualifications were used, only the mutually adjusted accumulation model was not significantly different from the saturated model. The coefficient for childhood SEP (0.60 (95% CI: -0.66, 1.85)) was much lower than the coefficients for early adulthood (8.77 (95% I: 7.57, 9.98)) and adult SEP (4.98 (95% CI: 3.42, 6.53)). For childhood household amenities and occupational SEP, both the adult accumulation and mutually adjusted accumulation models were not significantly different from the saturated model. The adult accumulation model had the lower BIC. Each additional time point of the two adult SEP measures spent in the higher SEP category increased the NART score by 5.15 points (95% CI: 4.22, 6.07).

Using father's occupational SEP and occupational SEP at age 26, both the accumulation and mutually adjusted accumulation model were not significantly different from the saturated model. The accumulation model had the lower BIC, with an increase in NART score of 4.66 (95% CI: 3.98, 5.34) points for each stage of the life course spent in the higher SEP category.

Table 6.9: NSHD tests of models for alternative life course hypotheses using father's occupational SEP at age 4, educational qualifications by age 26 and own occupational SEP at age 43, with outcome NART

	Women			Men		
Hypothesis (Equation number)	df	F statistic	P-value*	df	F statistic	P-value*
No effect	7,947	72.36	<0.0001	7,885	43.22	<0.0001
Accumulation models						
Accumulation (6.3)	6,947	4.78	0.0001	6,885	2.20	0.0408
Adult accumulation (6.5)	6,947	8.09	<0.0001	6,885	3.74	0.0011
Mutually adjusted accumulation (6.4)	4,947	3.26	0.0115	4,885	0.60	**0.6615**
Critical period models						
Childhood (6.6)	6,947	46.57	<0.0001	6,885	31.78	<0.0001
Early adulthood (6.6)	6,947	28.92	<0.0001	6,885	15.90	<0.0001
Adulthood (6.6)	6,947	47.19	<0.0001	6,885	19.48	<0.0001
Social mobility models						
Inter generational (6.8)	5,947	99.71	<0.0001	5,885	58.84	<0.0001
Intra generational (6.10)	5,947	75.48	<0.0001	5,885	53.37	<0.0001
Any mobility (6.12)	5,947	89.82	<0.0001	5,885	57.60	<0.0001
Any mobility with 3-way interaction (6.14)	4,947	31.84	<0.0001	4,885	17.90	<0.0001

* The p-values test whether the life course model is significantly different from the saturated model

Table 6.10: NSHD tests of models for alternative life course hypotheses using father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43, with outcome NART

	Women				Men			
Hypothesis (Equation number)	df	F statistic	P-value*	BIC	df	F statistic	P-value*	BIC
No effect	7,947	33.25	<0.0001		7,885	36.27	<0.0001	
Accumulation models								
Accumulation (6.3)	6,947	0.88	**0.5075**	6795.66	6,885	1.33	**0.2405**	6337.54
Adult accumulation (6.5)	6,947	11.48	<0.0001		6,885	3.39	0.0026	
Mutually adjusted accumulation (6.4)	4,947	1.06	**0.3761**	6809.18	4,885	0.40	**0.8078**	6343.33
Critical period models								
Childhood (6.6)	6,947	17.87	<0.0001		6,885	24.16	<0.0001	
Early adulthood (6.6)	6,947	23.83	<0.0001		6,885	7.87	<0.0001	
Adulthood (6.6)	6,947	18.40	<0.0001		6,885	16.01	<0.0001	
Social mobility models								
Inter generational (6.8)	5,947	44.62	<0.0001		5,885	48.52	<0.0001	
Intra generational (6.10)	5,947	38.08	<0.0001		5,885	46.80	<0.0001	
Any mobility (6.12)	5,947	39.31	<0.0001		5,885	49.09	<0.0001	
Any mobility with 3-way interaction (6.14)	4,947	15.01	<0.0001		4,885	13.65	<0.0001	

* The p-values test whether the life course model is significantly different from the saturated model

Table 6.11: NSHD tests of models for alternative life course hypotheses using father's occupational SEP, educational qualifications and head of household occupational SEP at age 43, with outcome NART

	Women		
Hypothesis (Equation number)	df	F statistic	P-value*
No effect	7,947	63.91	<0.0001
Accumulation models			
Accumulation (6.3)	6,947	6.02	<0.0001
Adult accumulation (6.5)	6,947	9.87	<0.0001
Mutually adjusted accumulation (6.4)	4,947	2.63	0.0249
Critical period models			
Childhood (6.6)	6,947	45.40	<0.0001
Early adulthood (6.6)	6,947	13.24	<0.0001
Adulthood (6.6)	6,947	49.96	<0.0001
Social mobility models			
Inter generational (6.8)	5,947	82.00	<0.0001
Intra generational (6.10)	5,947	74.79	<0.0001
Any mobility (6.12)	5,947	86.43	<0.0001
Any mobility with 3-way interaction (6.14)	4,947	36.63	<0.0001

* The p-values test whether the life course model is significantly different from the saturated model

Table 6.12: NSHD tests of models for alternative life course hypotheses using father's occupational SEP, own occupational SEP at age 26 and head of household occupational SEP at age 43, with outcome NART

	Women			
Hypothesis (Equation number)	df	F statistic	P-value*	BIC
No effect	7,947	31.03	<0.0001	
Accumulation models				
Accumulation (6.3)	6,947	0.13	**0.9932**	6786.70
Adult accumulation (6.5)	6,947	8.64	<0.0001	
Mutually adjusted accumulation (6.4)	4,947	0.16	**0.9606**	6800.19
Critical period models				
Childhood (6.6)	6,947	18.00	<0.0001	
Early adulthood (6.6)	6,947	19.14	<0.0001	
Adulthood (6.6)	6,947	20.06	<0.0001	
Social mobility models				
Inter generational (6.8)	5,947	42.47	<0.0001	
Intra generational (6.10)	5,947	36.84	<0.0001	
Any mobility (6.12)	5,947	37.82	<0.0001	
Any mobility with 3-way interaction (6.14)	4,947	19.24	<0.0001	

* The p-values test whether the life course model is significantly different from the saturated model

6.6.2 Whitehall II

6.6.2.1 Predicted mean score, by life course hypothesis

The results of these analyses are in Table 6.13, with the predicted values that were within 1 point of the observed mean Mill Hill test score in bold. There was a wider range of cognitive scores between the SEP trajectories for women (22.1 to 28.7) than men (23.9 to 27.8). For women the mean observed Mill Hill test scores fall into three categories along the lines of the adult accumulation model. For men almost all of the trajectories fall into three categories along the lines of the adult accumulation model. However the mean observed score was lower than would be expected for those participants who were in the lower SEP category during childhood, but the higher SEP category during early adulthood and adulthood.

For women the adult accumulation model was within 1 point of the observed mean Mill Hill test score for all eight of the trajectories. At the other extreme, the early adulthood critical period and inter-generational mobility models were each within 1 point from the observed mean NART score for only two of the trajectories. For men all three of the accumulation models predicted the Mill Hill test score to within 1 point of the observed Mill Hill test score for all eight of the trajectories. All of the life course models were within 1 point of the observed mean Mill Hill test score for at least three of the eight trajectories.

Table 6.13: Distribution of SEP trajectories, and observed and predicted mean (se) Mill Hill test scores according to the hypotheses, using father's occupational SEP (S_1), educational qualifications (S_2) and occupational SEP (S_3), with outcome Mill Hill test[2]

	SEP					Mean (se) predicted Mill Hill score					
						Accumulation models			Critical period		
Sex	S_1	S_2	S_3	N (%)	Observed mean Mill Hill test score (se)	Accumulation	Adult Accumulation	Mutually Adjusted accumulation	Childhood	Early adulthood	Adulthood
Women	0	0	0	313 (37.9)	22.1 (0.2)	**21.9 (0.2)**	**22.6 (0.2)**	**22.2 (0.2)**	**22.9 (0.2)**	**23.0 (0.2)**	**22.9 (0.2)**
(N=826)	1	0	0	218 (26.4)	23.4 (0.3)	**23.9 (0.2)**	**22.6 (0.2)**	**23.1 (0.3)**	25.1 (0.2)	**23.0 (0.2)**	**22.9 (0.2)**
	0	1	0	19 (2.3)	25.3 (1.3)	23.9 (0.2)	**25.5 (0.2)**	24.3 (0.4)	22.9 (0.2)	26.9 (0.3)	22.9 (0.2)
	0	0	1	22 (2.7)	26.3 (0.4)	23.9 (0.2)	**25.5 (0.2)**	**25.6 (0.4)**	22.9 (0.2)	23.0 (0.2)	27.6 (0.3)
	1	1	0	77 (9.3)	24.5 (0.7)	26.0 (0.2)	**25.5 (0.2)**	**25.1 (0.4)**	**25.1 (0.2)**	26.9 (0.3)	22.9 (0.2)
	1	0	1	46 (5.6)	25.5 (0.7)	**26.0 (0.2)**	**25.5 (0.2)**	26.5 (0.4)	**25.1 (0.2)**	23.0 (0.2)	27.6 (0.3)
	0	1	1	22 (2.7)	28.1 (0.6)	26.0 (0.2)	**28.4 (0.3)**	**27.7 (0.4)**	22.9 (0.2)	26.9 (0.3)	**27.6 (0.3)**
	1	1	1	109 (13.2)	28.7 (0.2)	**28.0 (0.3)**	**28.4 (0.3)**	**28.5 (0.3)**	25.1 (0.2)	26.9 (0.3)	27.6 (0.3)
Men	0	0	0	416 (17.1)	23.9 (0.2)	**23.7 (0.1)**	**24.1 (0.1)**	**23.8 (0.1)**	25.2 (0.1)	**24.9 (0.1)**	**24.4 (0.1)**
(N=2,440)	1	0	0	459 (18.8)	24.2 (0.2)	**25.0 (0.1)**	**24.1 (0.1)**	**24.3 (0.1)**	26.1 (0.1)	**24.9 (0.1)**	**24.4 (0.1)**
	0	1	0	79 (3.2)	25.4 (0.4)	**25.0 (0.1)**	**25.8 (0.1)**	**25.1 (0.2)**	**25.2 (0.1)**	26.9 (0.3)	24.4 (0.1)
	0	0	1	258 (10.6)	25.8 (0.2)	**25.0 (0.1)**	**25.8 (0.1)**	**25.8 (0.1)**	**25.2 (0.1)**	**24.9 (0.1)**	26.9 (0.1)
	1	1	0	151 (6.2)	25.5 (0.4)	**26.3 (0.1)**	**25.8 (0.1)**	**25.6 (0.2)**	**26.1 (0.1)**	26.9 (0.3)	24.4 (0.1)
	1	0	1	339 (13.9)	26.4 (0.1)	**26.3 (0.1)**	**25.8 (0.1)**	**26.4 (0.1)**	**26.1 (0.1)**	24.9 (0.1)	**26.9 (0.1)**
	0	1	1	223 (9.1)	26.7 (0.2)	**26.3 (0.1)**	**27.5 (0.1)**	**27.1 (0.1)**	25.2 (0.1)	**26.9 (0.3)**	**26.9 (0.1)**
	1	1	1	515 (21.1)	27.8 (0.1)	**27.7 (0.1)**	**27.5 (0.1)**	**27.7 (0.1)**	26.1 (0.1)	**26.9 (0.3)**	**26.9 (0.1)**

[2] Predicted values within 1 point of the observed mean Mill Hill test scores are in bold

Table 6.13 continued

	SEP					Social mobility			
Sex	S_1	S_2	S_3	N (%)	Observed mean Mill Hill test score (se)	Inter-generational Mobility	Intra-generational mobility	Any mobility	Any mobility with 3-way interaction
Women	0	0	0	313 (37.9)	22.1 (0.2)	24.0 (0.2)	23.8 (0.2)	23.9 (0.2)	**22.3 (0.2)**
(N=826)	1	0	0	218 (26.4)	23.4 (0.3)	**23.8 (0.3)**	**23.8 (0.2)**	**23.6 (0.3)**	**23.4 (0.2)**
	0	1	0	19 (2.3)	25.3 (1.3)	26.8 (0.8)	**24.7 (0.5)**	26.4 (0.5)	26.6 (0.4)
	0	0	1	22 (2.7)	26.3 (0.4)	24.0 (0.2)	**25.8 (0.6)**	**26.0 (0.5)**	**25.5 (0.5)**
	1	1	0	77 (9.3)	24.5 (0.7)	**24.0 (0.2)**	**24.7 (0.5)**	**23.6 (0.3)**	23.4 (0.2)
	1	0	1	46 (5.6)	25.5 (0.7)	23.8 (0.3)	**25.8 (0.6)**	**26.4 (0.5)**	26.6 (0.4)
	0	1	1	22 (2.7)	28.1 (0.6)	26.8 (0.8)	23.8 (0.2)	26.0 (0.5)	25.5 (0.5)
	1	1	1	109 (13.2)	28.7 (0.2)	24.0 (0.2)	23.8 (0.2)	23.9 (0.2)	**28.7 (0.4)**
Men	0	0	0	416 (17.1)	23.9 (0.2)	26.0 (0.1)	25.6 (0.1)	25.8 (0.1)	**24.1 (0.1)**
(N=2,440)	1	0	0	459 (18.8)	24.2 (0.2)	25.1 (0.1)	25.6 (0.1)	**24.9 (0.1)**	**24.4 (0.1)**
	0	1	0	79 (3.2)	25.4 (0.4)	26.4 (0.2)	**25.5 (0.2)**	26.7 (0.1)	26.4 (0.1)
	0	0	1	258 (10.6)	25.8 (0.2)	**26.0 (0.1)**	**26.1 (0.1)**	**25.7 (0.1)**	**26.1 (0.1)**
	1	1	0	151 (6.2)	25.5 (0.4)	**26.0 (0.1)**	**25.5 (0.2)**	**24.9 (0.1)**	24.4 (0.1)
	1	0	1	339 (13.9)	26.4 (0.1)	25.1 (0.1)	**26.1 (0.1)**	**26.7 (0.1)**	**26.4 (0.1)**
	0	1	1	223 (9.1)	26.7 (0.2)	**26.4 (0.2)**	25.6 (0.1)	25.7 (0.1)	**26.1 (0.1)**
	1	1	1	515 (21.1)	27.8 (0.1)	26.0 (0.1)	25.6 (0.1)	25.8 (0.1)	**27.8 (0.1)**

6.6.2.2 Life course hypotheses

As can be seen from Table 6.14 and Appendix 17: Table A17.5, when father's occupational SEP was used as the measure of childhood SEP, the mutually adjusted accumulation model was the only model which was not significantly different from the saturated model for both men and women. When childhood material deprivation was used as the measure of childhood SEP, the life course models were all significantly different from the saturated model for women, implying that none of the life course models were complex enough to describe the relationship between life course SEP and crystallized cognitive function in adulthood. For men, the mutually adjusted accumulation model was not significantly different from the saturated model.

In both models for men and the model with father's occupational SEP for women, the mutually adjusted accumulation model indicated a significant increase in Mill Hill test score with each time period spent in the higher SEP category. For all these models, the largest effect was for adult SEP, followed by early adulthood SEP, then childhood SEP; for example, when father's occupational SEP was used as the measure of childhood SEP for women the coefficient for adult occupational SEP was 3.29 (95% CI: 2.49, 4.09), for educational qualifications was 1.91 (95% CI: 1.12, 2.70), and the coefficient of childhood SEP was 0.82 (95% CI: 0.19, 1.46). A summary of the results can be seen in Table 6.15.

Table 6.14: Whitehall II tests of models for alternative life course using father's occupational SEP as the measure of childhood SEP, adjusted for age at phase 9 and the number of times the cognitive tests were taken, with outcome Mill Hill test

	Women			Men		
Hypothesis (Equation number)	df	F statistic	P-value*	df	F statistic	P-value*
No effect	7,816	28.36	<0.0001	7,2,430	64.89	<0.0001
Accumulation models						
Accumulation (6.3)	6,816	4.71	0.0001	6,2,430	10.17	<0.0001
Adult accumulation (6.5)	6,816	2.90	0.0083	6,2,430	5.31	<0.0001
Mutually adjusted accumulation (6.4)	4,816	1.91	**0.1073**	4,2,430	1.93	**0.1027**
Critical period models						
Childhood (6.6)	6,816	25.39	<0.0001	6,2,430	68.35	<0.0001
Early adulthood (6.6)	6,816	14.23	<0.0001	6,2,430	37.55	<0.0001
Adulthood (6.6)	6,816	7.55	<0.0001	6,2,430	17.78	<0.0001
Social mobility models						
Inter generational (6.8)	5,816	37.12	<0.0001	5,2,430	82.02	<0.0001
Intra generational (6.10)	5,816	37.95	<0.0001	5,2,430	88.23	<0.0001
Any mobility (6.12)	5,816	35.23	<0.0001	5,2,430	74.34	<0.0001
Any mobility with 3-way interaction (6.14)	4,816	4.20	0.0023	4,2,430	8.76	<0.0001

* The p-values test whether the life course model is significantly different from the saturated model

Table 6.15: Whitehall II life course model summary table, with outcome Mill Hill test

			Adjusted for age at phase 9 and number of times previously taken cognitive tests	
Childhood SEP	Early adulthood SEP	Adult SEP	Men	Women
Father's occupational SEP	Educational qualifications	Own occupational SEP at phase 7	Mutually adjusted accumulation	Mutually adjusted accumulation
Childhood material deprivation	Educational qualifications	Own occupational SEP at phase 7	Mutually adjusted accumulation	Saturated

6.6.3 NSHD life course analyses adjusting for childhood cognitive function

The NSHD analyses were re-run adjusting for childhood cognitive function and (childhood cognitive function)^2, in order to investigate whether the same life course hypotheses were supported after adjusting for the effect of childhood cognitive function on adult cognitive function. If childhood SEP was included in the life course model selected, this would indicate that it remained an important predictor of adult cognitive function through a mechanism other than childhood cognitive function. A summary of the results are presented in Table 6.16, alongside the results of the unadjusted analyses.

For men the adult accumulation hypothesis model had the lowest BIC of all the models which were not different from the saturated model, for all of the SEP combinations. This is different to the unadjusted analyses for three of the four SEP combinations, where childhood SEP was also involved in the hypothesis, either through the accumulation or mutually adjusted accumulation hypotheses.

For women the same results were found for own and head of household occupational SEP at age 43, except for childhood household amenities and educational qualifications, where the mutually adjusted accumulation model was supported for own occupational SEP (Appendix 17 Table A17.10), and adult accumulation for head of household occupational SEP (Appendix 17 Table A17.12). When own occupational SEP was used for early adulthood SEP, the same conclusions were drawn as in the unadjusted analyses.

When father's occupational SEP and educational qualifications were included in the model only the mutually adjusted accumulation model was not significantly different from the saturated model; in the unadjusted analyses none of the life course models were sufficiently complex to describe the relationship between life course SEP and crystallized cognitive function.

Table 6.16: NSHD life course model summary table, with outcome NART

			Unadjusted		Adjusted for childhood cognitive function and (childhood cognitive function)^2	
Childhood SEP	Early adulthood SEP	Adult SEP	Men	Women	Men	Women
Father's occupational SEP	Educational qualifications	Own occupational SEP at age 43	Mutually adjusted accumulation	Saturated	Adult accumulation	Mutually adjusted accumulation
Father's occupational SEP	Educational qualifications	Head of household occupational SEP at age 43	-	Saturated	-	Mutually adjusted accumulation
Childhood household amenities	Educational qualifications	Own occupational SEP at age 43	Adult accumulation	Mutually adjusted accumulation	Adult accumulation	Mutually adjusted accumulation
Childhood household amenities	Educational qualifications	Head of household occupational SEP at age 43	-	Mutually adjusted accumulation	-	Adult accumulation
Father's occupational SEP	Own occupational SEP at age 26	Own occupational SEP at age 43	Accumulation	Accumulation	Adult accumulation	Accumulation
Father's occupational SEP	Own occupational SEP at age 26	Head of household occupational SEP at age 43	-	Accumulation	-	Accumulation
Childhood household amenities	Own occupational SEP at age 26	Own occupational SEP at age 43	Mutually adjusted accumulation	Adult accumulation	Adult accumulation	Adult accumulation
Childhood household amenities	Own occupational SEP at age 26	Head of household occupational SEP at age 43	-	Adult accumulation	-	Adult accumulation

6.7 Discussion

Main findings

For both the NSHD and Whitehall II, accumulation models were supported in the majority of cases, unless none of the life course models were complex enough, and all of the life course models were significantly different to the saturated model. Most of the accumulation models identified as the best model included all three time points, although the adult accumulation model was selected for the NSHD when childhood household amenities was used, with educational qualifications for men, or occupational SEP at age 26 for women. When childhood cognitive function was adjusted for in the NSHD analyses, childhood SEP was no longer involved in any of the identified life course models for men, but the adjustment had less impact for women. As in the analyses in Chapter 4, the use of different SEP variables resulted in different conclusions. Further, different conclusions were drawn in NSHD compared with Whitehall II. If not all of the life course hypotheses had been considered, the conclusions may have been misleading; for example using backwards selection (Method 1), some of the interaction terms remained significant, indicating that social mobility hypotheses may be appropriate, however when the social mobility models were tested directly using the life course methodology (Method 2), they were significantly different from the saturated model.

Comparison with other studies

Both of the studies which investigated the effect of cumulative SEP on cognitive function (77;90) observed a dose-response relationship, and the three studies reviewed in section 1.2.8.3 found evidence to support the hypothesis that those participants who experienced social mobility between two time points had different cognitive scores from those participants who did not experience social mobility. However neither of the studies which investigated both the accumulation and social mobility hypotheses explored which of the two models was the most appropriate. In the current study accumulation models were supported, but all of the social mobility models were significantly different from the saturated model, implying that the social mobility models did not adequately describe the relationship between life course SEP and cognitive function.

Explanation of findings

Epidemiological

In the NSHD different life course hypotheses were best supported by the data depending on the specific SEP variables were used. Using each of the four SEP combinations for the earliest two time points resulted in four different conclusions being drawn for women, with three different results found for men; namely the three accumulation models.

For women, when father's occupational SEP was used as the childhood measure of SEP, one of the four models concluded that none of the life course hypotheses was supported by the data, but the other three models all suggested a form of accumulation, including childhood SEP. However when childhood household amenities was used as the measure of childhood SEP, only two of the four models included childhood SEP (mutually adjusted accumulation model), with the other two models supporting the adult accumulation hypothesis, implying that childhood household amenities were less strongly associated with adult cognitive function than other measures of SEP after adjusting for later life SEP. The accumulation hypothesis, where each variable has the same coefficient, was the hypothesis best supported by the data for both men and women when occupational SEP was used at all three time points (father's occupational SEP, own occupational SEP at age 26, own occupational SEP at age 43), whereas other combinations contained variables which were less similar to each other.

It would not necessarily be expected for the different SEP variables in various combinations to lead to the same conclusions for a number of reasons. The different SEP variables measure different aspects of SEP, which have different influences on the participants lives; for example father's occupational SEP is more likely to influence the household environment experienced during childhood, partially through parenting practices (37), whereas childhood household amenities may affect the participant's childhood through poor health or a lack of resources (179).

In addition, the different proportions of participants in the higher SEP category for each variable may also contribute to the different conclusions reached, especially for early adulthood SEP, where the most extreme proportions were observed for women, with 75.3% in the higher occupational SEP category, but only 29.1% in the higher educational qualifications category. As well as limiting the possible mobility trajectories, the proportions in each category influence the accumulation models, limiting the number of participants who could have an accumulation score of 3.

When childhood cognitive function was added to the NSHD models, the life course hypotheses best supported by the data changed in six of the twelve analyses. All three of the SEP combinations for men which included childhood SEP in the unadjusted life course analyses concluded that the adult accumulation model provided the best fit of the models not significantly different to the saturated model. This change was in the expected direction, with childhood SEP no longer necessary in the life course model after allowing for the effect of childhood cognitive function, as any variation in adult cognitive function that childhood SEP was explaining before adjustment has been explained by childhood cognitive function.

Differences between the NSHD and Whitehall II

The NSHD and Whitehall II study were sampled from very different populations, and hence it is of interest to compare the results between the studies. It is of interest to know whether similar findings hold in a selective occupational cohort with less variation in SEP as in a population sample. The most comparable of the unadjusted NSHD models are those using educational qualifications and own occupational SEP at age 43. It is important to note here that the Whitehall II childhood and educational qualifications data were collected retrospectively, whereas the NSHD data were collected prospectively. When father's occupational SEP was used as the measure of childhood SEP, for men both datasets concluded that the mutually adjusted accumulation model was the hypothesis best supported by the data. However for women the Whitehall II data supported the mutually adjusted accumulation model, whereas the NSHD data found all of the life course models to be significantly different from the saturated model, implying a more complex relationship.

When childhood household amenities (NSHD)/material deprivation (Whitehall II) was used as the measure of childhood SEP, the results were the opposite way round for women, with the mutually adjusted accumulation hypothesis supported in the NSHD, but all of the models being significantly different from the saturated model in the Whitehall II study. For men in the NSHD the adult accumulation model provided the best fit, whereas in the Whitehall II study the mutually adjusted accumulation hypothesis was supported. The measures of childhood material deprivation differed between the NSHD and Whitehall II; 50.8% of men and 50.4% of women were in the lower childhood material deprivation category in the NSHD, whereas in Whitehall II, 33.1% of men and 38.2% of women were in the lower SEP category. In Whitehall II the majority of women were in the *lower* adulthood occupational SEP category (75.7%), whereas in the NSHD the majority of women were employed in non-manual occupations at age 43, placing them in the *higher* SEP category (74.5%). For women, the proportion in the higher educational qualifications category was similar (NSHD: 29.1%, Whitehall II: 27.6%), whereas for men a larger proportion were in the higher category in the NSHD (47.7%, compared to 39.6% in Whitehall II). For men this difference is likely to be due to the younger age of the participants, as higher levels of education became more usual as time progressed.

These figures suggest a reason for the different results between men when childhood material deprivation was used as the measure of childhood SEP. The difference between two groups is likely to be less if the two groups were able to be split equally, whereas when the third with the worst circumstances is compared to the remainder, a difference is more likely to be observed if there is an association between the stratifying variable and the outcome. Therefore childhood SEP would have more influence in the Whitehall II dataset than the NSHD.

Methodological:

Using the backwards selection methodology (Method 1), it was not always simple to identify which of the life course hypotheses were indicated by the results, although accumulation models were often identified. In the NSHD when father's occupational SEP and educational qualifications were included in the model, it was not clear which of the hypotheses would be supported using Method 1; there was a significant interaction between childhood and early adulthood SEP, which indicated that an accumulation model may not be sufficient. Using the life course methodology (Method 2) each of the life course models was significantly different to the saturated model, indicating that none of the life course models were complex enough to be appropriate. In this situation, when the saturated model was selected using Method 2, there is a role for Method 1. One example of this is in Table 6.9, where the saturated model was selected, and Table 6.5 shows the same analyses using Method 1. From Method 1 the inter-generational social mobility and mutually adjusted accumulation models were identified, or potentially the early adulthood critical period. There was not one model clearly identified, which is why Method 2 failed to identify one life course model which was supported – instead there are aspects of many of the life course models (the high coefficient of the early adulthood SEP variable, the significant interaction between childhood SEP and early adulthood SEP, but when they are combined, they do not form one of the life course models tested by Method 2. If other parameterizations of the life course models had been tested, it is possible one of them would not have been significantly different from the saturated model, for example if an inter-generatioinal social mobility model had also allowed for an effect of adult SEP. By identifying that all of the life course models tested in Method 2 were significantly different to the saturated model, it was then possible to look back at the results from Method 1 and understand why this is the case.

The fact that Method 2 identifies which models were not significantly different from the saturated model, and then from those identifies the best fitting model by comparing the BIC, allows the life course hypotheses to be compared. Although it is possible to test accumulation and social mobility models as in Turrell et al (77) and Luo & Waite (90), and the BICs could then be compared, this would not be equivalent to Method 2 as it would not be known whether either of the models were sufficiently complex to describe the relationship. Therefore Method 2 is preferred for testing life course hypotheses, though as explained in the previous paragraph, there is still a role for Method 1.

The life course methodology described and implemented above was straightforward to implement in Stata, and was very flexible, allowing alternative hypotheses to be tested. It was also easy to adjust for other variables, in the same way as in standard linear regression models.

In the paper by Mishra et al describing the life course methodology (89) two of the models were not significantly different from the saturated model for men; one of the two models was identified as the model which provided the best fit without explaining how the decision between the two models had been made. As the two models had the same degrees of freedom it is likely the two p-values were compared. However due to the different number of parameters fitted in the different life course models, it is not always sufficient to compare the p-values for each model to identify which of the hypotheses was best supported by the data. The BIC (Equation 4.2), which penalises the model for each additional parameter fitted, was therefore used to compare the models.

This is why models which had lower p-values were often chosen as the model providing the best fit to the data.

Mobility models are dependent on the starting point; even with three time points it is important to remember that participants who are in the higher SEP category in childhood can only be upwardly mobile if they are first downwardly mobile. If mobility is split into inter-generational and intra-generational then this further restricts those who can be upwardly or downwardly mobile.

For the methodology described by Mishra et al it was necessary to dichotomise the SEP variables, which causes some information to be lost. One potential issue when dealing with life course analyses is collinearity, as the participant's SEP at each stage of the life course are likely to be correlated; however Mishra et al (89) concluded that collinearity was unlikely to be a problem unless the SEP indicators were measured closely in time.

Limitations

In order to use the methodology described by Mishra et al, it was necessary to choose three time points to represent life course SEP. The time points at which the SEP variables were chosen is more of an issue for the Whitehall II participants, as the participants are not all the same age; the adult SEP variable was the last recorded occupational SEP at phase 7, when the participants were aged 50-74. The ages at which the SEP variables used in the life course analyses were therefore less evenly distributed for some of the participants than others.

It was also necessary to choose an SEP variable for each of the three stages of the life course. In the NSHD two variables were chosen for each stage of the life course, and in the Whitehall II analyses two variables were compared for childhood SEP. As shown above, different results were found for different SEP combinations; therefore a wider range of SEP variables could be considered, especially for the Whitehall II participants, where other SEP variables may be more appropriate to represent older age, especially since a large proportion of the participants had retired.

As discussed in section 4.10, the childhood SEP variables were collected retrospectively in the Whitehall II dataset. The analyses above were complete case analyses, and it is known that the missingness is not MCAR. This is addressed in Chapter 7. The issues of overadjustment and collider bias which occur in some of the life course models were also mentioned in section 4.10. However as concluded above, overadjustment typically biases results towards the null (166), and a finding that childhood SEP was not included in the life course model selected would not imply that childhood SEP was not important, rather that it was important through its influence on later life SEP.

Strengths

Prior to this work, not many studies had considered life course SEP and cognitive function in adulthood. No studies have previously compared life course hypotheses when using cognitive function as the outcome measure, which may have led to misleading conclusions. Two methodologies for considering life course hypotheses were compared, with the advantages of the method of Mishra et al described. Additional life course hypotheses were tested beyond those defined in the original paper. As mentioned earlier (Chapter 4), all of the SEP variables were collected prospectively in the NSHD, which is rare when using SEP data spanning forty years.

Conclusions and implications

The accumulation hypotheses were generally supported when exploring the relationship between life course SEP and cognitive function in adulthood, although whether childhood SEP was included in the accumulation hypothesis or the stages of the life course were of equal importance varied depending on the SEP variables and dataset used. These analyses showed the advantages of the methodology developed by Mishra et al (89) over considering just one life course hypothesis, or using backwards selection to draw conclusions about which life course hypothesis is supported, especially when more than one hypothesis is supported by the data.

Chapter 7: Missing data and weighting for the life course methodology

7.1 Introduction

This chapter extends the work on the life course models described in the previous chapter, carrying out missing data analyses, and considering the effect of the different lengths of time spent in each stage of the life course. Although the Heckman selection models did not perform well in the simulation study, one Heckman selection model was included in this chapter to investigate how the results compared to the complete case and multiple imputation results.

7.2 Methods

7.2.1 Multiple imputation and Heckman selection

The life course analyses, adjusting for childhood cognitive function in the NSHD, were carried out on the multiply imputed data, which was generated for the analyses in Chapter 4, using multiple imputation, then deletion. The Heckman selection analyses were carried out using the selection model developed in Chapter 4 with the stricter conditions.

In the complete case analyses, the BIC was used to compare models which were not significantly different from the saturated model. The BIC can also be calculated for the Heckman selection analyses, but the log likelihood, which is involved in calculating the BIC, does not exist after multiple imputation, as multiple imputation does not involve the calculation of likelihood functions for the data (180). Therefore if more than one of the life course models was not significantly different from the saturated model, it was not possible to compare the models to identify which of the life course models provided the best fit to the data. However it was still possible to test the life course models against the saturated model using a specially written command in Stata to allow for this in multiple imputation analyses, based on work by Li et al (181).

7.2.2 Weighted analyses

One potential limitation of the life course methodology is that one period of the life course may have a stronger effect because it covers a longer period of time. One possible way of accounting for this is to weight the SEP variables to reflect the amount of time spent in each stage of the life course. The analyses in this section were carried out using complete case analyses.

In the NSHD, the participants were all aged 53 when the outcome measure, the NART score, was collected. The weightings used the ratio 16:18.5:18.5 for childhood, early adulthood and adulthood, respectively. This reflects childhood covering ages 0-16, with early adulthood and adulthood equally weighted (early adulthood: 16-34.5, adulthood: 34.5-53). The weights applied were these ratios divided by 53 to allow the sum of the weights to equal one.

However in Whitehall II the participants were not all the same age when the outcome measure, the Mill Hill test, was collected. As the participants were of different ages, ranging from 55 to 79 at phase 9, the proportions of their lives spent in each SEP stage varied. The weights therefore had the ratio 16:14: (age at phase 9 – 30), with each of these values divided by the participant's age at phase 9, in order for the sum of the weights to equal one. The childhood measure therefore represents age 0-16, the early adulthood measure represents 16-30, and adulthood represents age 30 – age at phase 9.

7.3 Results

7.3.1 Multiple imputation and Heckman selection

7.3.1.1 NSHD

In the tables below (Table 7.1 - Table 7.4, and Appendix 18: Table A18.1 – Table A18.4) models which were not significantly different from the saturated model are in bold, and of those, the model with the lowest BIC is in a box.

There were three differences between the models selected using complete case analyses and the Heckman selection analyses. The first difference can be seen in Table 7.2, where the three SEP measures were father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43. For women, the complete case analysis concluded that the accumulation hypothesis was best supported by the data, whereas in the Heckman selection analyses all of the models were significantly different from the saturated model. The same life course models were significantly different from the saturated model in the multiple imputation analysis as in the complete case analysis, although it was not possible to identify which model provided the best fit for the multiple imputation analysis.

The second difference can be seen in Table 7.3, where the SEP measures used were father's occupational SEP, educational qualifications and head of household occupational SEP at age 43. The complete case and multiple imputation analyses identified only the mutually adjusted accumulation model as not significantly different from the saturated model, whereas the Heckman selection model showed that both the accumulation and mutually adjusted accumulation models were not significantly different from the saturated model. In the Heckman selection analyses, the p-value was much higher in the mutually adjusted accumulation model (0.8452) than the accumulation model (0.0680). However, the BIC values were extremely similar, with a slightly lower value for the accumulation model (15834.74 vs. 15834.87), indicating that the accumulation model provided a better fit to the data. This is likely to be due to the lower number of parameters fitted in the accumulation model, as the BIC penalises the model for the number of parameters fitted.

The third difference in the life course model selected was for women when childhood material deprivation, educational qualifications and head of household occupational SEP were the SEP variables (Appendix 18: Table 18.3). The complete case analysis identified the adult accumulation model as the model which best fit the data, whereas the multiple imputation and Heckman selection model identified the mutually adjusted accumulation model.

There were no differences found in the overall conclusions drawn by the three missing data methods in only two of the eight SEP combinations (Appendix 18: Table 18.1 and Table 7.4). The differences were most often between the Heckman selection analyses and the other analyses.

Appendix 18: Table A18.3 contains the only situation in which the conclusion from the complete case analyses differed from both the other analyses, where the adult accumulation model was found not to be significantly different from the saturated model in the complete case analysis, but it was in the other analyses. Appendix 18: Table A18.2 contains the only situation where a conclusion from the multiple imputation analyses differed from the other two analyses, where the multiple imputation analyses found the any mobility with a three-way interaction model to be significantly different from the saturated model, for women.

7.3.1.2 Whitehall II

As in the NSHD, there were some discrepancies between the conclusions drawn by the complete case, multiple imputation and Heckman selection analyses. When father's occupational SEP was used as the measure of childhood SEP for women (Table 7.5) the complete case and multiple imputation analyses identified only the mutually adjusted accumulation model as not significantly different from the saturated model, whereas the Heckman selection analyses identified both the mutually adjusted accumulation model and the accumulation model, with the accumulation model having the lower BIC. When childhood material deprivation was used as the childhood SEP measure for men (Appendix 18: Table A18.5), the complete case and multiple imputation analyses identified the mutually adjusted accumulation model as the model which provided the best fit to the data, however the Heckman selection model showed all of the life course models to be significantly different from the saturated model.

For women, when childhood material deprivation was in the model (Appendix 18: Table A18.5) the complete case analyses found all the models to be significantly different from the saturated model, the multiple imputation analysis found the mutually adjusted accumulation model to be the only model that was not significantly different from the saturated model, whereas the Heckman selection analyses identified both the mutually adjusted accumulation model and the adult accumulation model, with the adult accumulation model having the lower BIC. This was the only difference in determining whether any of the life course models were significantly different from the saturated model between the complete case and multiple imputation analyses.

Table 7.1: NSHD: testing life course models using father's occupational SEP, educational qualifications and own occupational SEP at age 43, with outcome NART, under complete case, multiple imputation then deletion and Heckman selection

	Women			Men		
	p-value			p-value		
Hypothesis	CC	MID	Heckman	CC	MID	Heckman
No effect	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Accumulation models						
Accumulation	0.0092	0.0218	0.0416	0.0350	0.0020	**0.0882**
Adult accumulation	0.0026	0.0015	0.0048	**0.5305**	**0.3987**	**0.7500**
Mutually adjusted accumulation	**0.4519**	**0.8311**	**0.2487**	**0.6072**	**0.5583**	**0.6718**
Critical period models						
Childhood	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Early adulthood	0.0013	0.0011	0.0019	<0.0001	<0.0001	0.0047
Adulthood	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0001
Social mobility models						
Inter generational	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Intra generational	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility with 3-way interaction	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0020

Table 7.2: NSHD: testing life course models using father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43, with outcome NART, under complete case, multiple imputation then deletion and Heckman selection

	Women			Men		
	p-value			p-value		
Hypothesis	CC	MID	Heckman	CC	MID	Heckman
No effect	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Accumulation models						
Accumulation	**0.6516**	**0.8404**	0.0386	**0.1636**	**0.0695**	**0.2916**
Adult accumulation	0.0023	0.0011	0.0003	**0.3867**	**0.9392**	**0.3721**
Mutually adjusted accumulation	**0.3895**	**0.7138**	0.0100	**0.5019**	**0.9945**	**0.4886**
Critical period models						
Childhood	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Early adulthood	0.0001	<0.0001	<0.0001	<0.0001	0.0001	0.0014
Adulthood	<0.0001	<0.0001	<0.0001	<0.0001	0.0014	0.0909
Social mobility models						
Inter generational	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Intra generational	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility with 3-way interaction	0.0472	<0.0001	0.0061	<0.0001	0.0017	0.0023

Table 7.3: NSHD: testing life course models using father's occupational SEP, educational qualifications and head of household occupational SEP at age 43, with outcome NART, under complete case, multiple imputation then deletion and Heckman selection

	Women		
	p-value		
Hypothesis	CC	MID	Heckman
No effect	<0.0001	<0.0001	<0.0001
Accumulation models			
Accumulation	0.0071	0.0096	**0.0680**
Adult accumulation	0.0084	<0.0001	0.0123
Mutually adjusted accumulation	**0.5966**	**0.7181**	**0.8452**
Critical period models			
Childhood	<0.0001	<0.0001	<0.0001
Early adulthood	0.0004	0.0004	0.0085
Adulthood	<0.0001	<0.0001	<0.0001
Social mobility models			
Inter generational	<0.0001	<0.0001	<0.0001
Intra generational	<0.0001	<0.0001	<0.0001
Any mobility	<0.0001	<0.0001	<0.0001
Any mobility with 3-way interaction	<0.0001	<0.0001	<0.0001

Table 7.4: NSHD: testing life course models using father's occupational SEP, own occupational SEP at age 26 and head of household occupational SEP at age 43, with outcome NART, under complete case, multiple imputation then deletion and Heckman selection

	Women		
	p-value		
Hypothesis	CC	MID	Heckman
No effect	<0.0001	<0.0001	<0.0001
Accumulation models			
Accumulation	**0.7938**	**0.8978**	**0.9877**
Adult accumulation	0.0124	0.0023	0.0385
Mutually adjusted accumulation	**0.6512**	**0.7666**	**0.9242**
Critical period models			
Childhood	<0.0001	<0.0001	0.0020
Early adulthood	<0.0001	<0.0001	0.0002
Adulthood	0.0001	<0.0001	0.0004
Social mobility models			
Inter generational	<0.0001	<0.0001	<0.0001
Intra generational	<0.0001	<0.0001	<0.0001
Any mobility	<0.0001	<0.0001	<0.0001
Any mobility with 3-way interaction	0.0007	0.0001	0.0048

Table 7.5: Whitehall II: testing life course models using father's occupational SEP, educational qualifications and last recorded own occupational SEP at phase 7, with outcome Mill Hill test, under complete case, multiple imputation then deletion and Heckman selection

	Women			Men		
	p-value			p-value		
Hypothesis	CC	MID	Heckman	CC	MID	Heckman
No effect	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Accumulation models						
Accumulation	0.0001	<0.0001	**0.5023**	<0.0001	<0.0001	0.0003
Adult accumulation	0.0083	0.0008	0.0074	<0.0001	0.0001	0.0004
Mutually adjusted accumulation	**0.1073**	**0.6184**	**0.5846**	**0.1027**	**0.2661**	**0.1995**
Critical period models						
Childhood	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Early adulthood	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Adulthood	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Social mobility models						
Inter generational	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Intra generational	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility with 3-way interaction	0.0023	<0.0001	0.0029	<0.0001	<0.0001	<0.0001

7.3.2 Weighted for time spent in each stage of the life course

7.3.2.1 NSHD

The tables for these analyses are in Appendix 19. A much wider range of models were found not to be significantly different from the saturated model when the SEP variables were weighted for the length of time each section of the life course covered. For almost all the SEP combinations, the same model provided the best fit to the data as in the unweighted models. The only exception was for women when father's occupational SEP, educational qualifications and head of household SEP at age 43 were included in the model (Appendix 19: Table A19.5). With this combination of SEP variables, the unweighted model identified the mutually adjusted accumulation model as the model which was best supported by the data, whereas the weighted analyses supported the any mobility model.

7.3.2.2 Whitehall II

Only one of the four analyses using the weighted data (Appendix 19: Table A19.9 and Table 7.12) identified a different model which provided the best fit to the data as the unweighted analyses. When father's occupational SEP was used as the measure of childhood SEP (Table 7.12), the unweighted analyses concluded the mutually adjusted accumulation model was the only life course model which was not significantly different from the saturated model for women, whereas in the weighted analyses the accumulation and any mobility with the three way interaction models were also not significantly different from the saturated model, with the accumulation model providing the best fit of the three models.

Table 7.6: Whitehall II weighted life course models using father's occupational SEP, educational qualifications and last recorded own occupational SEP at phase 7, with outcome Mill Hill test

Hypothesis	Women				Men			
	df	F-statistic	P-value*	BIC	df	F-statistic	P-value*	BIC
No effect	7,816	28.51	<0.0001		7,2,430	66.37	<0.0001	
Accumulation models								
Accumulation	6,816	2.07	**0.0546**	5121.64	6,2,430	3.97	0.0006	
Adult accumulation	6,816	2.16	0.0445		6,2,430	5.54	<0.0001	
Mutually adjusted accumulation	4,816	1.40	**0.2313**	5132.75	4,2,430	1.98	**0.0944**	13853.73
Critical period models								
Childhood	6,816	25.89	<0.0001		6,2,430	70.00	<0.0001	
Early adulthood	6,816	14.33	<0.0001		6,2,430	38.91	<0.0001	
Adulthood	6,816	7.04	<0.0001		6,2,430	18.27	<0.0001	
Social mobility models								
Inter generational	5,816	14.76	<0.0001		5,2,430	46.63	<0.0001	
Intra generational	5,816	5.52	0.0001		5,2,430	11.58	<0.0001	
Any mobility	5,816	2.65	0.0220		5,2,430	6.79	<0.0001	
Any mobility with 3-way interaction	4,816	1.64	**0.1612**	5148.07	4,2,430	1.45	**0.2165**	13880.46

* The p-values test whether the life course model is significantly different from the saturated model

7.4 Discussion

Main findings

As in earlier analyses, when there were differences in the results between the missing data methods, the Heckman selection model usually reached a different conclusion to the complete case and multiple imputation results.

Both the NSHD and Whitehall II weighted results were very similar to the unweighted results; in each dataset there was only one model where a different life course model was selected in the unweighted and weighted analyses, despite the weightings in the Whitehall II analyses giving a much larger weight to adult SEP.

Explanation of findings

The differences in the life course model selected most often occurred between the complete case and Heckman selection models. However there were no consistent patterns in these differences; in three of the six situations where a difference was observed the Heckman selection model found an additional life course model not to be significantly different from the saturated model, and the additional model had a lower BIC. In the other three situations the Heckman selection model found the life course model selected by the complete case analysis to be significantly different from the saturated model. Neither the cohort nor gender affected which of the two situations occurred, nor the SEP variables in the model. In the multiple imputation analyses the models which were not significantly different to the saturated model were usually the same as in the complete case analyses.

When applying weights to the stages of the life course there were no rules to determining what the ratio of the weights should be. In the NSHD the weights were almost equal, so very little difference was expected between the unweighted and weighted results. However in the Whitehall II analyses the weights were less equal due to the older ages of the participants, and the weights differed between participants in the study, as the ratio depended on their age; therefore larger differences were expected between the unweighted and weighted results. However only one difference was found. Adult SEP had the largest weight for all of the participants, although the exact ratio varied by participant age. The weights applied were chosen to reflect the length of time the SEP measures were likely to be reflective of the participant's SEP, and the length of time spent in each section of the life course. However, other ratios exist which would be equally valid, especially when dividing the time spent in early adulthood and later adulthood SEP. Weighting the stages of the life course does not impact on the significance of the variables in the mutually adjusted accumulation model, where no constraints were attached to the coefficients of the three SEP variables; in the general case only the coefficients would differ, the significance would not.

In the Whitehall II study the only difference in the life course model identified as the model which provided the best fit from those not significantly different to the saturated model was for women when father's occupational SEP was used as the measure of childhood SEP. In the unweighted analysis the mutually adjusted accumulation model was selected, whereas the accumulation model was selected in the weighted analyses. Adult SEP had the largest coefficient in the unadjusted analyses, however as it was given the largest weighting in the weighted analyses, the adult SEP coefficient was no longer the largest in the weighted analyses. The standard error of each coefficient increased, which is likely to have led to the change in life course model selected, as the coefficients were no longer significantly different to each other due to the wider confidence intervals.

However in the NSHD, the selected model for women when using father's occupational SEP, educational qualifications and head of household occupational SEP changed from the mutually adjusted accumulation model in the unweighted analyses to the any mobility model in the weighted analyses. This model contains interaction terms, which magnify the effect of using weightings, as the weights are multiplied together. This shows that even small differences to the weightings can have an impact on which life course model was selected. When the unweighted any mobility model was run the downward mobility variable was not significant, whereas in the weighted analysis both the upward and downward mobility variables were significant.

In the unweighted analyses it is unclear whether the effect of adult SEP has been inflated, either due to the increased length of time spent in that stage of the life course, or because it represents the accumulation of earlier time periods. The weighted analyses allow for the length of time spent in each stage of the life course, but it is not possible to separate out the current effect of adult SEP and the extent to which it represents accumulation of SEP over the life course, beyond adjusting for earlier life SEP.

Limitations

Many of the same limitations in relation to the life course modelling approach apply to analyses in this chapter as in the previous chapter. The choice of weightings is inevitably based on assumptions and the choice made is likely to have affected the results. In the NSHD the weightings were 16:18.5:18.5, which is not far off 1:1:1, so differences would be expected to be minimal. However some differences were observed between the unweighted and weighted analyses.

This chapter has demonstrated that it was possible to apply the life course methodology developed by Mishra et al to multiply-imputed datasets and Heckman selection analyses, although it is not yet possible to compare the models which were not significantly different from the saturated model in the multiple imputation analyses. Maarten L. Buis recently described this as an area of active research (182).

Strengths

No previous work has investigated the effect of weighting in life course analyses; therefore this work represents an extension to the work carried out by Mishra et al (89).

Conclusions

The accumulation models were supported by the multiple imputation and Heckman selection analyses as well as the complete case analyses, although there were occasional differences in both the models which were different to the saturated model and the final life course model selected between the different missing data methods.

Overall the results for the weighted analyses were very similar to the unweighted analyses, yet in future work it is worth investigating any effect of weighting on life course analyses when the stages of the life course used were not equally spaced.

Chapter 8: The effect of childhood and life course SEP on cognitive decline

8.1 Introduction

The second part of this thesis extends the work of the first part, which investigated the effect of childhood and life course SEP on cognitive function at one time point. It investigates Aim 2 by examining the effect of childhood and life course SEP on cognitive trajectories. Life course SEP was first examined using measures of SEP from childhood, early-adulthood and midlife, modelled separately, before using the life course methodology developed by Mishra et al (89).

8.2 Methodology

The Whitehall II study was used to carry out these analyses, as there are 4 phases of memory scores available in the dataset (phases 3, 5, 7 and 9), spanning a period of 15 years. Memory was selected as the cognitive test of interest as it is a sensitive measure of fluid cognitive function, which is expected to decline with age. Descriptive results of memory scores at each phase were compared, to investigate the overall change in memory score across phases. The methodology used when collecting the phase 3 cognitive data differed to the methodology used in later phases. Additionally, the cognitive tests were only taken by some of the phase 3 participants, who differed from the phase 3 participants who did not take the phase 3 cognitive tests, as described in section 3.2.4.2. Hence the phase 3 memory data were excluded from this analysis. Age was centred at 50 years, and year of birth was centred at 1940. The analyses were carried out for men and women together to increase the power.

8.2.1 Introduction to multilevel models

Multilevel modelling is used to model data that has a hierarchical or clustered structure, such as pupils within a class. In longitudinal studies, an individual's responses over time are likely to be correlated, and repeated measures data can be considered hierarchical, with the measurements nested within individuals. Multilevel models allow for the hierarchical nature of the data by permitting residual components to be included at each level in the hierarchical structure. The variation both within- and between-individuals can be modelled and estimated explicitly. Additionally, the effects of exposures can be allowed to differ between individuals. Multilevel modelling was therefore used to investigate how life course SEP affected the trajectory of memory.

The simplest model has the form:

$y_{ij} = \beta_0 + e_{ij}$

where:

y_{ij} is the memory score at phase *i* for participant *j*,

β_0 is the overall mean memory score,

and e_{ij} is the level-1 residual, the difference from the mean memory score for participant *j* at phase *i*. e_{ij} is assumed to be normally distributed with mean 0 and variance σ_e^2.

A random intercept was introduced to the model, which allowed the intercept to vary for each participant. The form of a multilevel model with a random intercept is:

$$y_{ij} = \beta_{0j} + e_{ij}$$

where $\beta_{0j} = \beta_0 + u_{0j}$; y_{ij}, β_0 and e_{ij} are defined as above, β_{0j} is the mean memory score for each participant *j* over all phases *i*, and u_{0j} is the level-2 residual, the difference from the mean memory score for participant *j*.

Both time-invariant and time varying covariates can be added to a multilevel model. For example, if the time varying covariate *age* was added to the above model, the multilevel model would have the form:

$$y_{ij} = \beta_{0j} + \beta_1 age_{ij} + e_{ij}$$

This model would have the same slope for each participant *j*. If a covariate *x* is constant within participants, for example gender, then it would be expressed as x_j rather than x_{ij}. To allow for a random slope in the model, the coefficient of the covariate can differ for each participant, as in the following model:

$$y_{ij} = \beta_{0j} + \beta_{1j} age_{ij} + e_{ij}$$

The random slope $\beta_{1j} = \beta_1 + u_{1j}$, where β_{1j} is the effect of increasing the covariate *age* by one unit for person *j*, β_1 is the effect of increasing the covariate *age* by one unit across all participants, and u_{1j} is thus the difference in the effect of increasing the covariate *age* by one unit for participant *j* from the effect across all participants. Both the u_{0j} and u_{1j} are assumed to be normally distributed with mean 0, var(u_{0j})=σ_{u0}^2, var(u_{1j})=σ_{u1}^2 and cov(u_{0j},u_{1j}) = σ_{u01}.

8.2.2 Applying multilevel modelling

The variables were added to the models in blocks. As the primary interest is the memory trajectory with respect to ageing, **Block 1** contained age. The first model was a random effects model containing age as a fixed effect, allowing for a random intercept, as the mean baseline memory scores were not the same for each individual (Model 1). To test whether age influenced memory in a linear manner, an age-squared term was added to the model, also as a fixed effect (Model 2). To investigate whether age had the same effect on all individuals, age, in this model, was considered as a random effect, allowing each individual to have their own slope (i.e. rate of change with age) (Model 3). Age-squared was then also considered as a random effect, to investigate whether the shape of the curve varied between individuals (Model 4).

Practice effects, gender, period and cohort effects were considered in **Block 2**. It was necessary to allow for the practice effects resulting from the different number of times the participants had taken the cognitive tests, as some participants took the cognitive tests at phase 3 and/or the repeat tests after phases 3 and 7 (section 3.2). Therefore, the number of times the cognitive tests had previously been taken, was added as a time-varying fixed effect, to allow for practice effects. Initially, the number of times the cognitive tests had been taken was included as a linear variable (Model 5), then as a categorical variable (Model 6). Model 7 then contained the linear variable practice effects as a random effect, and the models in Block 2 were compared using the BIC. The next model (Model 8) added gender as a fixed effect. An interaction between age and gender was considered in Model 9 to investigate whether the rate of change in memory scores differed for men and women.

It was also necessary to consider cohort and period effects (183), and separate them from the effect of age. The advantage of having repeat measures of cognitive function is that it enables longitudinal change (the effect of chronological age) to be distinguished from cross-sectional change (the cohort effects) (184). Age at the time of data collection is the time in years between the participant being born and the year that the data were collected. ‘Period’ is a proxy for a set of ‘contemporaneous influences’, and ‘cohort’ is a proxy for ‘influences in the past’. Period effects could include the development of ‘brain training’ games which have become popular only in recent years, whereas a cohort effect could reflect differences by birth year in the minimum age of leaving full-time education. The minimum age of leaving full-time education was raised from 14 to 15 in the 1944 Education Act, which was implemented in April 1947. Therefore, participants who were born before 1932 (4%) were required by law to complete one less year of education.

Since age equals year of data collection (period) minus year of birth (cohort), an identification problem arises when modelling all three effects. It was decided to investigate cohort effects rather than period effects in order to investigate whether the rate of cognitive decline depended on year of birth. Therefore Model 10 included year of birth to test for cohort effects, and Model 11 included an interaction between year of birth and age to investigate whether cohort influenced the rate of decline.

Block 3 considered the SEP variables: childhood material deprivation, as defined in section 3.2.3.1 (Models 12 – 14), father's occupational SEP (Models 15 – 17), educational qualifications (Models 18 – 20) and grouped phase 1 occupational SEP (Models 21 – 23). Three models were run containing each of the SEP variables in turn; first a model adding only the SEP variable to the model chosen at the end of Block 2. The next model added an interaction between the SEP variable and gender, and the third model added an interaction between the SEP variable and age to the model containing the SEP variable.

The last stage of the complete case analyses was to fit models which contained more than one SEP variable, to investigate whether there was an effect of childhood SEP on memory or memory decline with age, after adjusting for education and occupation.

Finally, multiple imputation analysis was carried out for all the analyses, and results compared with those from the complete case analyses. The imputation was carried out with the data in 'wide' format, with one data record per individual (185). After the imputations had been carried out, the data were reshaped to the 'long' format, with a record for each memory score. The imputation model was developed as in section 4.7. The `Heckman` command in Stata cannot be implemented in multilevel models (see section 9.5).

8.3 Results

8.3.1 Descriptive results

The shapes of the density plots for memory were normally distributed, and similar for each of phases 5, 7 and 9 (see section 3.2.4.2). The mean memory score remained similar at phases 5 and 7, before dropping at phase 9 (Table 8.1). The mean score at phase 5 differed significantly for those who had taken the memory test at phase 3 and those who had not, with a mean score over 0.5 points higher for those who had taken the memory test at phase 3 (6.63 vs. 7.17).

Table 8.1: Descriptive statistics of the memory scores at each phase

Phase	N	Mean	Standard Deviation
5	6017	6.86	2.45
7	6349	6.79	2.43
9	6060	6.21	2.29

Four participants (0.07%) achieved the maximum possible score at phase 9, and 28 participants (0.47%) achieved the minimum score at phase 5. Apart from these cases, the maximum and minimum scores were not attained, so there was no concern over ceiling and floor effects.

8.3.2 Model development

In the complete case multilevel modelling analyses there were 10,540 memory observations from 4,106 individuals with all the required covariates.

<u>Block 1:</u>

Model 1: $y_{ij} = \beta_0 + \beta_1 age_{ij} + u_{0j} + e_{ij}$

Model 2: $y_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 age^2_{ij} + u_{0j} + e_{ij}$

Model 3: $y_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 age^2_{ij} + u_{0j} + u_{1j} age_{ij} + e_{ij}$

Model 4: $y_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 age^2_{ij} + u_{0j} + u_{1j} age_{ij} + u_{2j} age^2_{ij} + e_{ij}$

Both age and age-squared were associated with memory scores (Model 1 and Model 2). The model (Model 3) would not converge when the random effect age was allowed to have an unstructured covariance; therefore the model was refitted using an independent variance matrix. The BIC of Model 3 was higher than in Model 2 (45913.29 vs. 45906.36), indicating worse model fit. However as the effect of age was expected to differ between individuals, the random effect of age was retained in the model to see if the estimate of the between individual variance changed with the addition of further variables to the model. Model 4 attempted to model age-squared as a random effect, but the standard errors could not be computed, implying that the model was too complex for the data. Therefore Model 3 was carried forward as the basis for Block 2, consisting of a negative quadratic relationship between age and memory, with a random effect of age.

<u>Block 2:</u>

Model 5: $y_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 age^2_{ij} + \beta_3 practice\ effects_{ij} + u_{0j} + u_{1j} age_{ij} + e_{ij}$

Model 6: $y_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 age^2_{ij} + \beta_3 practice\ effects\_1_{ij} + \beta_4 practice\ effects\_2_{ij} + \beta_5 practice\ effects\_3_{ij} + \beta_6 practice\ effects\_4_{ij} + u_{0j} + u_{1j} age_{ij} + e_{ij}$

Model 7: $y_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 age^2_{ij} + \beta_3 practice\ effects_{ij} + u_{0j} + u_{1j} age_{ij} + u_{2j} practice\ effects_{ij} + e_{ij}$

The number of times the cognitive tests had previously been taken (practice effects in the models above) was significant when considered both as a linear (Model 5) and categorical (Model 6) variable. Model 6 had the lower BIC (Model 5: 45869.08 vs. Model 6: 45839.13), indicating that the relationship was not linear. Practice effects in linear form were also modelled as a random effect, but, the standard errors could not be computed. Therefore Model 6 was carried forward as the basis for Block 3, with a negative quadratic relationship for age, and positive coefficients for the number of times the cognitive tests had been taken.

Model 8: $y_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 age^2_{ij} + \beta_3 practice\ effects\_1_{ij} + \beta_4 practice\ effects\_2_{ij} + \beta_5 practice\ effects\_3_{ij} + \beta_6 practice\ effects\_4_{ij} + \beta_7 female_j + u_{0j} + u_{1j} age_{ij} + e_{ij}$

Model 9: $y_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 age^2_{ij} + \beta_3 practice\ effects\_1_{ij} + \beta_4 practice\ effects\_2_{ij} + \beta_5 practice\ effects\_3_{ij} + \beta_6 practice\ effects\_4_{ij} + \beta_7 female_j + \beta_8 female_j * age_{ij} + u_{0j} + u_{1j} age_{ij} + e_{ij}$

Model 10: $y_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 age^2_{ij} + \beta_3 practice\ effects\_1_{ij} + \beta_4 practice\ effects\_2_{ij} + \beta_5 practice\ effects\_3_{ij} + \beta_6 practice\ effects\_4_{ij} + \beta_7 female_j + \beta_8 year\ of\ birth_j + u_{0j} + u_{1j} age_{ij} + e_{ij}$

Model 11: $y_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 age^2_{ij} + \beta_3 practice\ effects\_1_{ij} + \beta_4 practice\ effects\_2_{ij} + \beta_5 practice\ effects\_3_{ij} + \beta_6 practice\ effects\_4_{ij} + \beta_7 female_j + \beta_8 year\ of\ birth_j + \beta_9 year\ of\ birth_j * age_{ij} + u_{0j} + u_{1j} age_{ij} + e_{ij}$

Model 8 shows that females had an intercept 0.23 (95% CI: 0.10, 0.36) points higher than males. The interaction between age and gender was not significant (Model 9), indicating that the rate of change in memory score was not significantly different for men and women. Year of birth (Model 10) was a significant predictor of memory score, with increasing memory scores associated with being born more recently. A significant interaction between year of birth and age was observed (Model 11), with a faster rate of cognitive decline for participants born more recently (Figure 8.1).

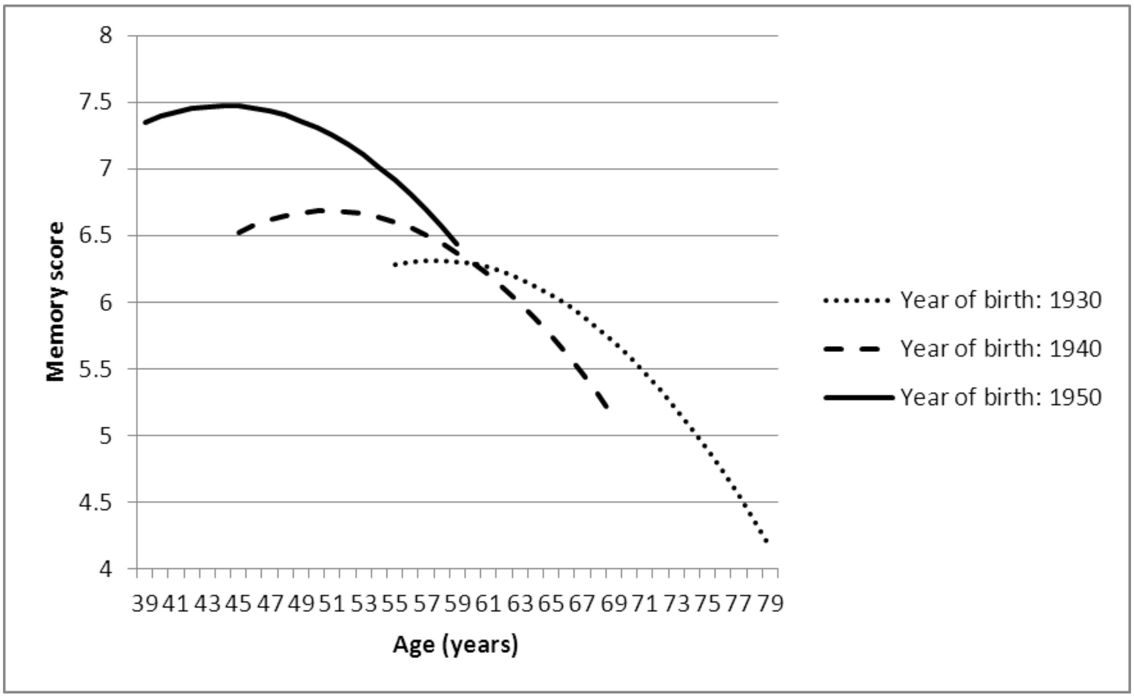


Figure 8.1: Predicted memory score trajectory by year of birth for men taking the cognitive tests for the first time at phase 5

<u>Block 3:</u>

Model 12: Model 11 + $\beta_{10} child\ mat.\ dep._j$

Model 13: Model 11 + $\beta_{10} child\ mat.\ dep._j + \beta_{11} female_j * child\ mat.\ dep._j$

Model 14: Model 11 + β_{10}*child mat. dep.*$_{j}$ + β_{11}*age*$_{ij}$**child mat. dep.*$_{j}$

Childhood material deprivation was a significant predictor of memory score (Model 12, Table 8.2), with an increase of 0.18 (95% CI: 0.09, 0.27) points per unit in childhood material deprivation score, with increasing scores representing material advantage. The interaction between childhood material deprivation and gender was not significant (Model 13, p=0.281). However the interaction between childhood material deprivation and age was significant (Model 14, p=0.040), with a slightly faster rate of decline for participants with more advantaged childhood material conditions (Figure 8.2), closing the gap with increasing age. To confirm that there was no significant difference between the memory scores by childhood SEP in older ages, where there is less data, an ANOVA was carried out to compare the memory scores when the childhood SEP scores were divided into tertiles. For both those over aged 70 and over aged 75 there was no significant difference in memory score by childhood SEP (p=0.136 and p=0.968 respectively). When Model 14 was run restricted to those under age 75, so as not to be influenced by those at the extreme of the age scale, the childhood SEP by age interaction remained significant.

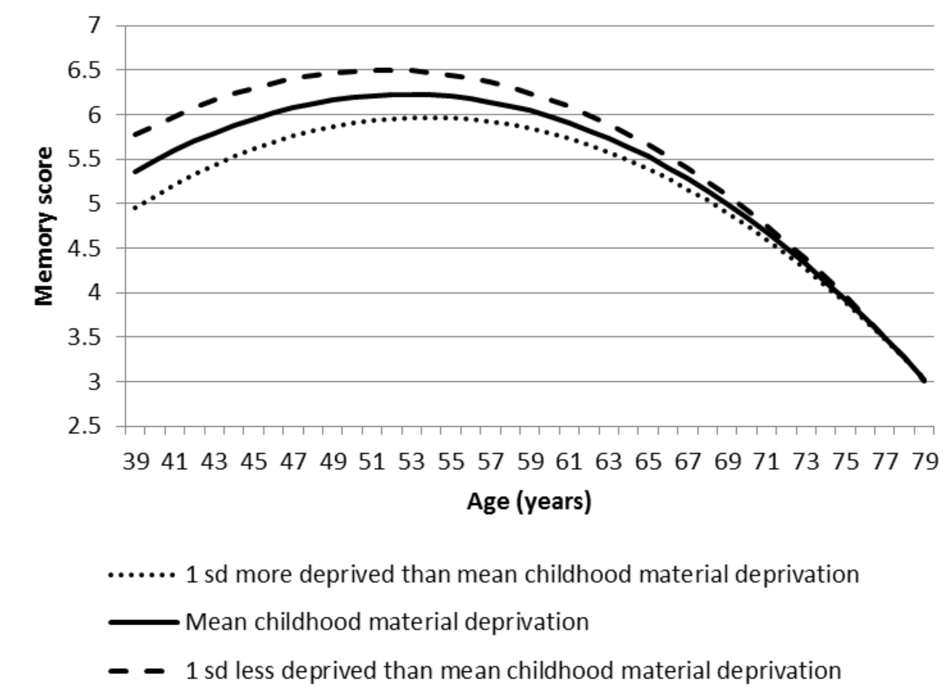


Figure 8.2: Model 14: Predicted memory trajectory by childhood material deprivation for male participants born in 1940 taking the cognitive test for the first time at phase 5

Model 15: Model 11 + β_{10} *child occ SEP*$_{j}$

Model 16: Model 11 + β_{10} *child occ SEP*$_{j}$+ β_{11} *female*$_{j}$**child occ SEP*$_{j}$

Model 17: Model 11 + β_{10} *child occ SEP*$_{j}$+ β_{11} *age*$_{ij}$**child occ SEP*$_{j}$

Participants whose fathers were in a non-manual occupational SEP had memory scores 0.20 (95% CI: 0.09, 0.32) points higher than participants whose fathers were employed in manual occupations (Table 8.2). Neither the interaction between father's occupational SEP and gender (Model 16) nor father's occupational SEP and age (Model 17) were significant.

Model 18: Model 11 + β_{10}*school cert.*$_j$ + β_{11}*GCSE/A-Level*$_j$ + β_{12}*Degree*$_j$

Model 19: Model 11 + β_{10}*school cert.*$_j$ + β_{11}*GCSE/A-Level*$_j$ + β_{12}*Degree*$_j$ + β_{13}*female*$_j$**school cert.*$_j$ + β_{14}*female*$_j$**GCSE/A-Level*$_j$ + β_{15}*female*$_j$**Degree*$_j$

Model 20: Model 11 + β_{10}*school cert.*$_j$ + β_{11}*GCSE/A-Level*$_j$ + β_{12}*Degree*$_j$ + β_{13}*age*$_{ij}$**school cert.*$_j$ + β_{14}*age*$_{ij}$**GCSE/A-Level*$_j$ + β_{15}*age*$_{ij}$**Degree*$_j$

The third SEP variable considered was educational qualifications (Table 8.3), which was a significant predictor of memory score (Model 18), with increasing coefficients for increasing levels of educational qualifications. Neither the interaction between educational qualifications and gender (Model 19, p=0.186) nor the interaction between educational qualifications and age (Model 20, p=0.054) were significant.

Model 21: Model 11 + β_{10}*Senior/Higher Exec*$_j$ + β_{11}*Unified Grades 1-6*$_j$

Model 22: Model 11 + β_{10}*Senior/Higher Exec*$_j$ + β_{11}*Unified Grades 1-6*$_j$ + β_{12}*female*$_j$** Senior/Higher Exec*$_j$+ β_{13}*female*$_j$**Unified Grades 1-6*$_j$

Model 23: Model 11 + β_{10}*Senior/Higher Exec*$_j$ + β_{11}*Unified Grades 1-6*$_j$ + β_{12}*age*$_{ij}$** Senior/Higher Exec*$_j$ + β_{13}*age*$_{ij}$**Unified Grades 1-6*$_j$

The final SEP variable considered was occupational SEP at phase 1 of the study (Table 8.3). Occupational SEP was a significant predictor of memory score (Model 21), with increasing coefficients for increasing occupational status. The interaction term between gender and occupational SEP (Model 22) was not significant (p=0.609), however the interaction between age and occupational SEP (Model 23) was significant (p=0.014). Figure 8.3 shows the predicted mean memory scores by occupational SEP for male participants born in 1940, who were taking the cognitive tests for the first time. The gap between the clerical and other occupations decreased with increasing age.

Table 8.2: Testing the effect of SEP variables on the Whitehall II memory trajectory

	Model 11		Model 12		Model 14		Model 15	
	Coef. (95% CI)	p-value	Coef. (95% CI)	p-value	Coef. (95% CI)	p-value	Coef. (95% CI)	p-value
Fixed effects								
Age (centred at 50)	0.01 (0.02)	0.744	0.01 (0.02)	0.767	0.02 (0.02)	0.320	0.01 (0.02)	0.735
Age squared (centred at 50)	-0.005 (0.001)	<0.001	-0.005 (0.001)	<0.001	-0.005 (0.001)	<0.001	-0.005 (0.001)	<0.001
Practice effects: baseline – none								
Practice effects_1	0.48 (0.06)	<0.001	0.48 (0.06)	<0.001	0.48 (0.06)	<0.001	0.48 (0.06)	<0.001
Practice effects_2	0.67 (0.09)	<0.001	0.66 (0.09)	<0.001	0.66 (0.09)	<0.001	0.67 (0.09)	<0.001
Practice effects_3	0.67 (0.13)	<0.001	0.65 (0.13)	<0.001	0.66 (0.13)	<0.001	0.67 (0.13)	<0.001
Practice effects_4	0.98 (0.18)	<0.001	0.96 (0.18)	<0.001	0.96 (0.18)	<0.001	0.97 (0.18)	<0.001
Female	0.23 (0.07)	0.001	0.24 (0.07)	<0.001	0.24 (0.07)	<0.001	0.24 (0.07)	<0.001
Year of birth (Centred 1940)	0.06 (0.02)	0.001	0.06 (0.02)	0.003	0.06 (0.02)	0.004	0.06 (0.02)	0.002
Year of birth (1940) * Age (50)	-0.01 (0.002)	<0.001	-0.01 (0.002)	<0.001	-0.01 (0.002)	<0.001	-0.01 (0.002)	<0.001
Childhood material deprivation			0.18 (0.05)	<0.001	0.29 (0.07)	<0.001		
Child. mat. dep * Age					-0.01 (0.01)	0.040		
Father's occ. SEP: non-manual							0.20 (0.06)	0.001
Constant	6.68 (0.16)	<0.001	6.39 (0.18)	<0.001	6.20 (0.20)	<0.001	6.56 (0.16)	<0.001
Random effects								
var(age (centred at 50))	2.36e-20 (1.71e-20)		6.14e-13 (4.52e-13)		3.92e-22 (2.73e-22)		1.87e-20 (1.37e-20)	
var(constant)	2.25 (0.08)		2.24 (0.08)		2.23 (0.08)		2.24 (0.08)	
var (Residual)	2.96 (0.05)		2.96 (0.05)		2.96 (0.05)		2.96 (0.05)	
BIC	45838.28		45832.98		45838.03		45835.93	

Table 8.3: Testing the effect of educational qualifications (Model 18), occupational grade at phase 1 (Model 21), and occupational grade with age interaction (Model 23)

	Model 18		Model 21		Model 23	
	Coef. (95% CI)	p-value	Coef. (95% CI)	p-value	Coef. (95% CI)	p-value
Fixed effects						
Age (centred at 50)	0.02 (0.02)	0.423	0.02 (0.02)	0.412	0.04 (0.02)	0.086
Age squared (centred at 50)	-0.005 (0.001)	<0.001	-0.005 (0.001)	<0.001	-0.005 (0.001)	<0.001
Practice effects: baseline – none						
Practice effects_1	0.46 (0.06)	<0.001	0.45 (0.06)	<0.001	0.44 (0.06)	<0.001
Practice effects_2	0.63 (0.09)	<0.001	0.61 (0.09)	<0.001	0.61 (0.09)	<0.001
Practice effects_3	0.62 (0.13)	<0.001	0.60 (0.13)	<0.001	0.60 (0.13)	<0.001
Practice effects_4	0.92 (0.18)	<0.001	0.90 (0.18)	<0.001	0.91 (0.18)	<0.001
Female	0.42 (0.07)	<0.001	0.67 (0.07)	<0.001	0.66 (0.07)	<0.001
Year of birth (Centred 1940)	0.06 (0.02)	0.003	0.08 (0.02)	<0.001	0.08 (0.02)	<0.001
Year of birth (1940) * Age (50)	-0.01 (0.002)	<0.001	-0.01 (0.002)	<0.001	-0.01 (0.002)	<0.001
Educational qualifications: baseline - no qualifications		<0.001				
School certificate	0.42 (0.16)	0.008				
GCSE/A-Levels	0.84 (0.10)	<0.001				
Degree	1.33 (0.11)	<0.001				
Occupational grade: baseline – clerical				<0.001		<0.001
Senior and Higher Executive Officers			0.82 (0.09)	<0.001	1.10 (0.14)	<0.001
Unified Grades 1-6			1.48 (0.10)	<0.001	1.76 (0.15)	<0.001
Occupational grade*Age (centred at 50)						0.014
Senior and Higher Executive Officers*Age					-0.03 (0.01)	0.004
Unified Grades 1-6*Age					-0.03 (0.01)	0.010
Constant	5.69 (0.19)	<0.001	5.57 (0.18)	<0.001	5.34 (0.20)	<0.001
Random effects						
var(age (centred at 50))	1.78e-19 (1.36e-19)		1.56e-19 (1.19e-19)		1.05e-20 (6.90e-21)	
var(constant)	2.11 (0.08)		2.06 (0.08)		2.05 (0.08)	
var (Residual)	2.96 (0.05)		2.96 (0.05)		2.96 (0.05)	
BIC	45696.95		45626.44		45636.4	

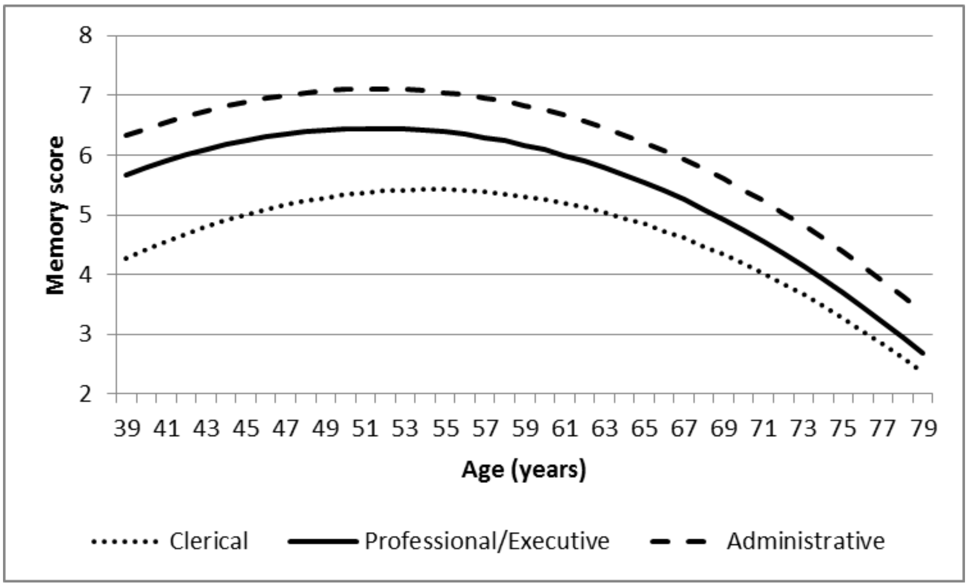


Figure 8.3: Model 23: Predicted memory trajectory by phase 1 occupational grade for male participants born in 1940 taking the cognitive test for the first time at phase 5

Model 24: Model 11 + β_{10}*child mat. dep.*$_{.j}$ + β_{11}*school cert.*$_{.j}$ + β_{12}*GCSE/A-Level*$_j$ + β_{13}*Degree*$_j$

Model 25: Model 11 + β_{10}*child mat. dep.*$_{.j}$ + β_{11}*school cert.*$_{.j}$ + β_{12}*GCSE/A-Level*$_j$ + β_{13}*Degree*$_j$ + β_{14}*Senior/Higher Exec*$_j$ + β_{15}*Unified Grades 1-6*$_j$

Model 26: Model 11 + β_9*child mat. dep.*$_{.j}$ + β_{10}*age*$_{ij}$**child mat. dep.*$_{.j}$ + β_{11}*school cert.*$_{.j}$ + β_{12}*GCSE/A-Level*$_j$ + β_{14}*Degree*$_j$ + β_{15}*Senior/Higher Exec*$_j$ + β_{16}*Unified Grades 1-6*$_j$

Model 27: Model 11 + β_9*child mat. dep.*$_{.j}$ + β_{10}*age*$_{ij}$**child mat. dep.*$_{.j}$ + β_{11}*school cert.*$_{.j}$ + β_{12}*GCSE/A-Level*$_j$ + β_{13}*Degree*$_j$ + β_{14}*age*$_{ij}$**school cert.*$_{.j}$ + β_{15}*age*$_{ij}$**GCSE/A-Level*$_j$ + β_{16}*age*$_{ij}$**Degree*$_j$ + β_{17}*age*$_{ij}$**Senior/Higher Exec*$_j$ + β_{18}*age*$_{ij}$**Unified Grades 1-6*$_j$

A significant effect of childhood material deprivation remained on the intercept after adjusting for educational qualifications (Model 24), although the effect had been attenuated (0.11 (95% CI: 0.02, 0.20) points increase per unit increase in childhood material deprivation score). The effect of childhood material deprivation was further attenuated by additionally adjusting for occupational SEP at phase 1 (Model 25), but remained a significant predictor. Childhood material deprivation was a significant predictor of rate of cognitive decline in Model 26, however there did not remain an effect of childhood material deprivation on the rate of memory decline after adjusting for later life SEP (Model 27, p=0.050). Occupational grade was a significant predictor of rate of memory decline (Model 27, p=0.0497), with those in clerical positions declining slightly slower (Figure 8.4).

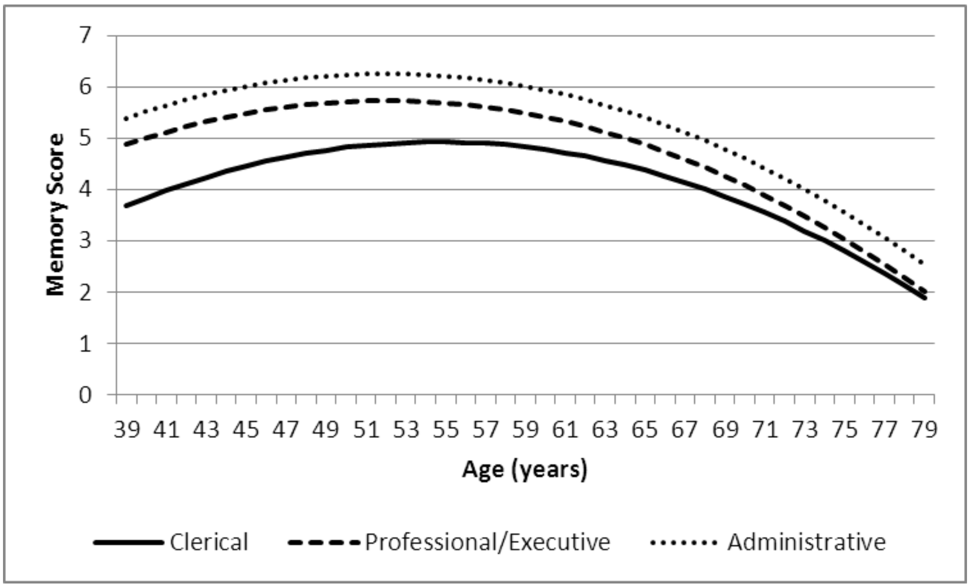


Figure 8.4: Model 27: Predicted memory trajectory by phase 1 occupational grade for male participants born in 1940 taking the cognitive test for the first time at phase 5

Model 28: Model 11 + β_{10}*child occ SEP*$_j$ + β_{11}*school cert.*$_j$ + β_{12}*GCSE/A-Level*$_j$ + β_{13}*Degree*$_j$

Model 28 concluded that the effect found in Model 15 for father's occupational SEP during childhood was fully attenuated when adjusting for educational qualifications.

To summarise, childhood material deprivation was a significant predictor of rate of memory decline before adjusting for the effect of later life SEP on the rate of memory decline, but after this adjustment it was no longer significant. Childhood occupational SEP was only a significant predictor of memory score before adjustment for later life SEP, and was not a significant predictor of rate of memory decline. In the fully adjusted model only occupational SEP was a significant predictor of rate of memory decline, with those in a clerical position experiencing a slightly slower rate of memory decline.

Table 8.4: Testing the effect of childhood material deprivation (Model 24) and father's occupational SEP (Model 26) after adjusting for educational qualifications, and occupational grade at phase 1 for childhood material deprivation

	Model 24		Model 26		Model 28	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Fixed effects						
Age (centred at 50)	0.02 (0.02)	0.295	0.04 (0.02)	0.076	0.02 (0.02)	0.423
Age squared (centred at 50)	-0.01 (0.001)	<0.001	-0.01 (0.001)	<0.001	-0.005 (0.001)	<0.001
Practice effects: baseline – none						
Practice effects_1	0.44 (0.06)	<0.001	0.44 (0.06)	<0.001	0.46 (0.06)	<0.001
Practice effects_2	0.59 (0.09)	<0.001	0.59 (0.09)	<0.001	0.63 (0.09)	<0.001
Practice effects_3	0.57 (0.13)	<0.001	0.57 (0.13)	<0.001	0.62 (0.13)	<0.001
Practice effects_4	0.88 (0.18)	<0.001	0.88 (0.18)	<0.001	0.92 (0.18)	<0.001
Female	0.69 (0.07)	<0.001	0.69 (0.07)	<0.001	0.42 (0.07)	<0.001
Year of birth (Centred 1940)	0.07 (0.02)	<0.001	0.07 (0.02)	0.001	0.06 (0.02)	0.004
Year of birth (1940) * Age (50)	-0.01 (0.002)	<0.001	-0.01 (0.002)	<0.001	-0.01 (0.002)	<0.001
Childhood material deprivation	0.09 (0.05)	0.038	0.22 (0.07)	0.002		
Child mat. Dep. * Age (50)			-0.01 (0.005)	0.022		
Father's occ. SEP: non-manual					0.03 (0.06)	0.639
Educational qualifications: baseline - no qualifications		<0.001		<0.001		<0.001
School certificate	0.24 (0.16)	0.134	0.24 (0.16)	0.130	0.42 (0.16)	0.009
GCSE/A-Levels	0.55 (0.11)	<0.001	0.55 (0.11)	<0.001	0.84 (0.10)	<0.001
Degree	0.78 (0.12)	<0.001	0.78 (0.12)	<0.001	1.31 (0.11)	<0.001
Occupational grade: baseline – clerical		<0.001		<0.001		
Senior and Higher Executive Officers	0.62 (0.10)	<0.001	0.61 (0.10)	<0.001		
Unified Grades 1-6	1.13 (0.11)	<0.001	1.13 (0.11)	<0.001		
Constant	5.06 (0.20)	<0.001	4.85 (0.22)	<0.001	5.68 (0.19)	<0.001
Random effects						
var(age (centred at 50))	6.44e-20 (4.95e-20)		2.68e-20 (1.89e-20)		1.77e-19 (1.35e-19)	
var(constant)	2.02 (0.07)		2.01 (0.07)		2.11 (0.08)	
var (Residual)	2.96 (0.05)		2.96 (0.05)		2.96 (0.05)	
BIC	45609.25		45613.28		45705.99	

8.3.3 Multiple Imputation

The four steps laid out in section 2.4.1 were followed to develop the imputation model for these analyses. The variables included in the first step are those in the models of interest, see Table 8.5. There were three interaction terms between age and each SEP variable, as an interaction was needed for age at each of the three phases.

The variables which were considered in the second step are presented in Appendix 20: Table A20.1. As before, only variables from earlier phases were tested to investigate whether they were predictive of missingness. In the third step, those variables which were a significant predictor of missingness for at least one of the variables in the model of interest were tested to examine whether they were associated with any of the variables in the model of interest (Appendix 20: Table A20.2). Only three variables from Appendix 20: Table A20.2 were not predictive of missingness for any of the variables of interest: the belief that it is safer to trust no one, the belief that you are not easily angered, and childhood emotional deprivation.

As the imputation model was already large, only variables from phase 9 were considered as auxiliary variables in the final step. The variables considered are in Appendix 20: Table A20.3. The final list of variables included in the imputation model is presented in Table 8.5.

Table 8.5: Variables included in the final imputation model

Phase	Predictive of missingness and observed value	Phase	Predictive of missingness and observed value	Phase	Included in model of interest
1	Age finished full time education	5	Ever told high blood pressure	1	Father's occupational SEP
1	Accommodation type	5	Ever diagnosed with cancer	5	Childhood material deprivation
1	Age mother finished full time education	5	Deprivation score	5	Educational qualifications
1	State of health in the last year	5	AH4 score	1	Occupational grade
1	Any longstanding illnesses?	5	Mill Hill test score	5, 7, 9	Memory score
1	Smoking status	5	Verbal fluency - S words	5, 7, 9	Age (centred at 50)
1	Job satisfaction	5	Verbal fluency - animals	5, 7, 9	Age (centred at 50) squared
1	Usually pressed for time	5	How financially secure do you feel in next 10 years	5, 7, 9	The number of times the cognitive tests had previously been taken
1	Believe no one cares much about you	5	To what extent do you feel you might as well give up because you can't make things better for yourself	1	Gender
1	Isolation score	7	Last recorded occupational grade	1	Year of birth (centred at 1940)
1	Depression case from GHQ	7	General health		Age*father's occupational SEP
3	Last recorded occupational grade	7	Health stops you from doing what you want to do		Age*childhood material deprivation
3	Marital status	7	CASP score		Age*educational qualifications
3	AH4 score	7	Clinic or home visit		Age*occupational grade
3	Mill Hill test score	7	MMSE score		
3	Verbal fluency - S words	7	Still at civil service	**Phase**	**Auxiliary variables**
3	Verbal fluency – animals	7	Marital status	9	AH4 score
3	Job involves travel away from home	7	AH4 score	9	Mill Hill test score
3	Memory score	7	Mill Hill test score	9	Verbal fluency - S words
4	Ever told had depression	7	Verbal fluency - S words	9	Verbal fluency – animals
4	Ever told had anxiety	7	Verbal fluency - animals	9	Last recorded occupational SEP
5	Last recorded occupational grade				

Comparing complete case and multiple imputation results

The multiple imputation analyses were carried out including (MI) and excluding (MID) imputed memory scores.

The majority of the multiple imputation analyses reached the same conclusions as the complete case analyses; therefore only the differences are discussed here. The only difference in the significance of any of the unadjusted SEP variables on the intercept was for father's occupational SEP; father's occupational SEP was a significant predictor of memory score in the complete case analysis (p=0.001), but not significant in the MI (p=0.102) or MID (p=0.078) analyses.

There were also differences related to the significance of the SEP variables on the rate of memory decline. In Model 14 the interaction between childhood material deprivation and age was significant in the complete case analysis (p=0.040), but not in the MI (p=0.644) or MID (p=0.507) analyses. The results from Model 20, containing the interaction between educational qualifications and age, were the other way round, with the interaction not significant in the complete case analysis (p=0.535), but significant in the MI (p=0.028) and MID (p<0.001) analyses.

8.4 Life course

It is plausible that the influence of life course SEP is different for the intercept and for the slope (rate of cognitive decline). For example, an accumulation model might be the most appropriate model for the intercept, whereas SEP might have no effect on rate of decline; alternatively the critical period model may be the most appropriate for cognitive decline if the rate of decline is only influenced by adult SEP. Therefore, the life course methodology was carried out as described in Section 6.3, but using a multilevel model. It was applied first in relation to the intercept, then in relation to the slope by including the interaction terms between each SEP term and age. The methodology was also applied to test the life course models for both the intercept and slope, constraining the same life course model to model the intercept and slope. The same basic model was used for the life course models as the regular multilevel models (Model 11, section 8.3.2).

Life course models - influencing the intercept

When childhood material deprivation was used as the measure of childhood SEP (Table 8.6), each of the life course models was significantly different from the saturated model. This implies that none of the hypothesised life course models were appropriate, and that the relationship was more complicated than any of the models allowed; therefore the saturated model was accepted as the model which provided the best fit. However when father's occupational SEP during childhood was used as the childhood SEP measure, both the adult accumulation and mutually adjusted accumulation models were not significantly different from the saturated model, with the adult accumulation model having the lower BIC value.

The multiple imputation analyses produced the same conclusions for both measures of childhood SEP, for both MI and MID; that only the mutually adjusted accumulation model was not significantly different from the saturated model (see Appendix 21: Table A21.1). This is different to the complete case result.

Life course models - influencing the slope

When childhood material deprivation was used as the measure of childhood SEP (Table 8.7), all of the life course models were significantly different from the saturated model, implying that none of the models were sufficiently complex to explain the relationship, and that the saturated model was required.

When father's occupational SEP was used as the measure of childhood SEP (Table 8.7), none of the life course models were significantly different from the saturated model. The no effect model had the lowest BIC value, implying that SEP was not associated with the rate of memory decline.

The same conclusions were reached for father's occupational SEP, for both the MI and MID analyses, which found that none of the life course hypotheses were significantly different from the saturated model (see Appendix 21: Table A21.2). However the results using childhood material deprivation were very different to the results found in the complete case analyses; in the MI analyses none of the models were significantly different to the saturated model, and in the MID analyses the accumulation, adult accumulation, mutually adjusted accumulation, early adulthood critical period and intra-generational social mobility models were not significantly different to the saturated model (see Appendix 21: Table A21.2).

Life course models - influencing both the intercept and the slope

The third set of life course models tested constrained the same life course model to support both the intercept and the slope. When childhood material deprivation was used as the measure of childhood SEP (Table 8.8), all of the models were significantly different from the saturated model, implying that none of the life course models were complex enough to explain the relationship between life course SEP and both memory score and memory decline. However when father's occupational SEP was used as the measure of childhood SEP (Table 8.8), the adult accumulation and mutually adjusted accumulation models were not significantly different from the saturated model, with the adult accumulation model providing the best fit to the data.

For both measures of childhood SEP, the MI and MID analyses concluded that the mutually adjusted accumulation model was not significantly different from the saturated model. For father's occupational SEP the MI analyses also found that the adult critical period model was not significantly different from the saturated model (Appendix 20: Table A20.3). As explained earlier, it is not possible to identify which model provided the best fit between two life course models in the MI analyses which were both not significantly different from the saturated model.

When the BICs of the models which were not significantly different from the saturated model were compared out of those influencing the intercept, the slope, and both the intercept and slope, the lowest BIC was found for the adult accumulation model for the intercept only model.

Table 8.6: Testing the life course hypotheses for the intercept, using childhood material deprivation (left) and father's occupational SEP (right) as the childhood SEP measure

		Childhood material deprivation		Father's occupational SEP		
Hypothesis	df	Chi squared statistic	P-value	Chi squared statistic	P-value	BIC
No effect	7	201.84	<0.0001	190.02	<0.0001	
Accumulation	6	45.45	<0.0001	42.87	<0.0001	
Adult accumulation	6	23.44	0.0007	12.09	**0.0599**	**45681.04**
Mutually adjusted accumulation	4	15.67	0.0035	6.72	**0.1517**	45694.20
Critical Period						
Childhood	6	193.45	<0.0001	177.99	<0.0001	
Early adulthood	6	96.50	<0.0001	84.97	<0.0001	
Adulthood	6	50.17	<0.0001	38.75	<0.0001	
Social Mobility						
Inter generational	5	171.22	<0.0001	163.93	<0.0001	
Intra generational	5	173.93	<0.0001	162.18	<0.0001	
Any mobility	5	146.04	<0.0001	144.34	<0.0001	
Any mobility with 3 way interaction	4	21.08	0.0003	27.89	<0.0001	

Table 8.7: Testing the life course hypotheses for the slope, using childhood material deprivation (left) and father's occupational SEP (right) as the childhood SEP measure

		Childhood material deprivation		Father's occupational SEP		
Hypothesis	df	Chi squared statistic	P-value	Chi squared statistic	P-value	BIC
No effect	7	24.73	0.0008	6.14	**0.5237**	**45724.54**
Accumulation	6	17.40	0.0079	4.55	**0.6027**	45732.22
Adult accumulation	6	23.19	0.0007	4.73	**0.5785**	45732.40
Mutually adjusted accumulation	4	9.98	0.0408	3.37	**0.4976**	45749.57
Critical Period						
Childhood	6	12.69	0.0483	5.67	**0.4617**	45733.33
Early adulthood	6	21.53	0.0015	3.63	**0.7271**	45731.30
Adulthood	6	24.64	0.0004	5.97	**0.4271**	45733.63
Social Mobility						
Inter generational	5	16.50	0.0056	5.67	**0.3393**	45742.10
Intra generational	5	22.18	0.0005	4.39	**0.4952**	45740.68
Any mobility	5	17.75	0.0033	6.03	**0.3029**	45741.77
Any mobility with 3 way interaction	4	14.01	0.0073	4.25	**0.3738**	45750.89

Table 8.8: Testing the life course hypotheses for the intercept and slope, using childhood material deprivation (left) and father's occupational SEP (right) as the childhood SEP measure

		Childhood material deprivation		Father's occupational SEP		
Hypothesis	df	Chi squared statistic	P-value	Chi squared statistic	P-value	BIC
No effect	14	227.22	<0.0001	196.22	<0.0001	
Accumulation	12	63.45	<0.0001	47.38	<0.0001	
Adult accumulation	12	47.05	<0.0001	17.05	**0.1476**	**45689.12**
Mutually adjusted accumulation	8	27.94	0.0005	8.54	**0.3826**	45717.68
Critical Period						
Childhood	12	211.05	<0.0001	183.72	<0.0001	
Early adulthood	12	118.17	<0.0001	87.78	<0.0001	
Adulthood	12	74.93	<0.0001	44.77	<0.0001	
Social Mobility						
Inter generational	10	184.00	<0.0001	168.19	<0.0001	
Intra generational	10	197.22	<0.0001	166.39	<0.0001	
Any mobility	10	165.04	<0.0001	150.35	<0.0001	
Any mobility with 3 way interaction	8	35.58	<0.0001	31.88	<0.0001	

8.5 Comparison of results

The findings from the two methods provide fairly consistent results for the intercept. When father's occupational SEP was used as the measure of childhood SEP, educational qualifications and occupation were significant predictors of the intercept of the memory trajectory, and the adult accumulation model was supported using the life course hypotheses. When childhood material deprivation was used as the measure of childhood SEP, all three of the SEP measures were significant predictors of the intercept, and the saturated model was supported by the life course models, implying that the mutually adjusted accumulation model, which contains all three of the SEP measures without any constraints on their coefficients, was not sufficiently complex.

However the results were less consistent for the slope of the memory trajectory. When father's occupational SEP was used as the measure of childhood SEP only occupational SEP was predictive of the rate of decline in the initial analyses, whereas in the life course analyses the no effect model was not significantly different from the saturated model, and had the lowest BIC, implying that SEP had no impact on the rate of memory decline. When childhood material deprivation was used as the measure of SEP, only occupational SEP was a significant predictor of rate of memory decline in the adjusted model, but all of the life course models were different to the saturated model, implying they were not complex enough.

There is a large difference in the results of the life course models for slope for the two childhood SEP measures; for father's occupational SEP none of the life course models were significantly different from the saturated model, whereas for childhood material deprivation all of the models were significantly different to the saturated model. This is likely to be due to the difference in how well the saturated model fits the data.

8.6 Discussion

<u>Main findings</u>

There was an effect of each SEP variable on the intercept of the memory trajectory before mutually adjusting for SEP at other stages of the life course, but whether there was an effect of each SEP variable on the rate of memory decline was not as clear, as the results varied according to the SEP measure considered. There was a significant effect of childhood material deprivation and occupational grade on rate of memory decline, with faster rates of decline for participants with more advantaged childhood material conditions and a slightly slower rate of decline for participants with a clerical occupational at phase 1.

Childhood material deprivation was a significant predictor of the intercept of the memory trajectory after adjusting for later life SEP, but father's occupational SEP was not. The effect of childhood material deprivation on the rate of memory decline (p=0.050) was fully attenuated by adjustment for the interactions between age and educational qualifications, and age and occupational SEP.

Overall similar results were found when using multiple imputation, although father's occupational SEP was not a significant predictor of memory score in the MI or MID analyses, and childhood material deprivation was not significantly associated with rate of memory decline. However the interaction between educational qualifications and age was significant in the MI and MID analyses, unlike the complete case analysis.

The life course results differed by childhood SEP measure. When childhood material deprivation was used the saturated model was chosen as the appropriate life course model for the intercept, the slope, and the intercept and slope together. When father's occupational SEP was used the adult accumulation model was chosen for the intercept and intercept and slope together, and the no effect model provided the best fit when considering the slope.

Comparison with other studies

Childhood SEP and cognitive decline

Of the four studies discussed in section 1.2.7, three found no association between childhood SEP and cognitive decline in older age, and the fourth (52) found a slight increase in odds of cognitive decline for those participants whose father was a farmer compared to white-collar workers. This study used the odds of experiencing the worst 10% of change in cognitive score as the outcome measure, whereas the other three studies investigated change in score. Two of the studies only had 5 years of follow up data, so it is possible that five years was not sufficient to identify a difference in rate of change by childhood SEP. However the Whitehall II study used in the analyses in this chapter investigated decline over 10 years, and found evidence of an effect of childhood material deprivation on rate of cognitive decline in the complete case analysis, although no effect of father's occupational SEP. The sample for the study which found an effect of childhood SEP (52) comprised nurses from the Nurses' Health Study, who have higher qualifications than the general population, as do members of the Whitehall II cohort. It is therefore possible that this effect would only be observed in a subsample of the population.

The results found in this chapter relating to the relationship between childhood SEP and the intercept of the memory trajectory differ to those found in Chapter 4, when the effect of childhood SEP after adjusting for later life SEP was investigated with respect to crystallized cognitive function. In the Whitehall II analyses in Chapter 4 childhood SEP remained significant for men but not women, and in the NSHD analyses childhood material deprivation was fully attenuated by later life SEP, but a significant effect of father's occupational SEP remained; the results from the NSHD in Chapter 4 are the opposite of the results found in this chapter. As in some of the literature discussed, different conclusions were reached for different cognitive measures and different measures of SEP.

Adult SEP and cognitive decline

In the literature contradictory results were found when considering whether educational qualifications were associated with the rate of cognitive decline; in Anstey and Christensen's review (61) neither of the studies that used fluid measures of cognitive function found a significant association between educational qualifications and rate of cognitive decline. Many studies have found that the significance of a relationship between educational qualifications and cognitive decline depended on the measure of cognitive function used (66;67). Using the Whitehall II study, Singh-Manoux et al (69) found that educational qualifications did not affect the rate of cognitive decline but that occupational SEP did affect the rate of cognitive decline. This is in line with the complete case results found in this Chapter, though educational qualifications were a significant predictor of cognitive decline when multiple imputation was implemented.

Life course SEP and cognitive decline

The only study investigating the effect of life course SEP on cognitive decline used the MMSE, a measure of mental status, and investigated cognitive decline as a binary outcome, whether or not decline (defined as a drop of 3 points) took place (91). This study by Long et al (91) investigated the accumulation hypothesis, and found evidence to support it, with a more disadvantaged accumulated SEP score associated with increased risk of experiencing decline. However the life course hypotheses for cognitive decline had not previously been compared.

Methodological discussion

The life course approach has not previously been implemented to investigate decline; the methodology explained by Mishra et al (89) was developed for continuous outcomes in generalised linear models. Although some extensions are relatively straightforward, such as an extension to binary outcomes, it was not immediately clear how to investigate the life course effect on decline. The most straight forward extension was to include the additional terms in the model in the same way as in a linear regression model, and then additionally include each term with an age interaction in order to investigate the life course effect on rate of decline. Testing the constraints on the coefficients for both the intercept and slope at the same time tested whether one life course model was appropriate for modelling both the intercept and slope.

Multilevel models allow for missingness that is MAR, yet there were still some differences between the complete case and MI/MID results. This is likely to be related to the fact that not all the variables which make the missingness MAR are included in the multilevel model, which is the condition required for the multilevel model to account for MAR missingness (186).

Limitations

Although it was decided to use occupational SEP at phase 1, before the cognitive tests were introduced to the study, occupational SEP could be fitted as a time-varying variable. The variable used to measure childhood material deprivation was made of four variables, some of which were likely to have a different prevalence over time, such as having an outside toilet and access to a car; this may be important as the participants were born between 1930 and 1952. The proportion going to university also varied greatly over the twenty year period from 1950 to 1970, the range over which Whitehall participants would have attended university (171). These analyses were only carried out on the Whitehall II participants as three waves of memory data were available for the Whitehall II participants. However this limits the generalizability of the results, as the Whitehall II study is an occupational cohort, and therefore not representative of the general population.

Strengths

This study investigated the effect of life course SEP on cognitive decline, using a measure of cognitive function that was not designed to identify cognitive impairment. In the only previous study investigating life course SEP and cognitive decline, the measure of cognitive function used was the MMSE, which measures cognitive impairment, and only the accumulation hypothesis was considered. This study used the life course methodology developed by Mishra et al (89), allowing the accumulation, critical period and social mobility hypotheses to be compared.

Conclusions and implications

Each of the four SEP variables considered had an impact on the intercept of the memory trajectory, although the effect of father's occupational SEP was fully attenuated by educational qualifications, unlike childhood material deprivation, which remained significant after adjustment for educational qualifications and occupational SEP. Childhood material deprivation and occupational SEP were also associated with the rate of memory decline in the complete case analysis, although the effect on the slope of childhood material deprivation was fully attenuated by adjustment for later life SEP. As the cognitive score at a given age results from both the peak cognitive score reached and the rate of cognitive decline, predictors of both the intercept and the slope are important.

Although multilevel modelling allows for missingness that is MAR, this relies on all the variables which make the missingness MAR being included in the multilevel model; as this is unlikely to occur it is also important to account for missing data in these situations. As discussed above, some different conclusions were drawn between the complete case and multiple imputation analyses.

Chapter 9: Conclusions and Discussion

This study has confirmed that childhood SEP is associated with crystallized cognitive function in late adulthood, but that some, or all, of this association, depending on the SEP measure used, may operate through later life SEP, as well as childhood cognitive function. It has demonstrated the utility of using recently developed life course methodology for investigating alternative life course hypotheses; accumulation, critical period and social mobility. It has extended the use of this methodology to investigate effects on not only the level of crystallized cognitive function but also on the trajectory of fluid cognitive function. There were also important implications of the method used to account for missing data; for example in unadjusted analyses the association of childhood material deprivation on rate of memory decline depended on the missing data methodology used; with an effect found in the complete case analysis, but not a multiple imputation analysis. No effect of childhood SEP was found on the rate of cognitive decline after adjusting for later life SEP.

Many of the results varied by missing data method; the simulation study suggested the multiple imputation results were most trustworthy, with the highest coverage under MAR and MNAR missingness. Therefore the multiple imputation results are focussed on where appropriate below.

9.1 Summary of main findings

Childhood SEP and adult crystallized cognitive function

The first objective of Aim 1 was to investigate the relationship between childhood SEP and adult crystallized cognitive function, before and after adjustment for later life SEP. In unadjusted analyses in Chapter 4, both measures of childhood SEP were associated with crystallized cognitive function in adulthood. The findings relating childhood SEP to cognitive function in adulthood after adjustment for later life SEP were inconsistent; conclusions varied depending on the measure of childhood SEP used, the gender of participants and cohort. In cases where conclusions varied according to the missing data methodology, the Heckman selection model usually resulted in a different conclusion compared with the complete case and multiple imputation analyses.

Missing data methodology, using a simulation study

The second objective of Aim 1 was to carry out a simulation study, examining the performance of the three selected methods of dealing with missing data; complete case, multiple imputation and Heckman selection models. The Heckman selection method did not perform well in the simulation study in Chapter 5 when the missing data was MAR or MNAR. This is likely to be due to the complex nature of the missing data in the NSHD and Whitehall II. In the simulation study multiple imputation performed well for MAR missingness; however although it was found to be the best method of the three considered for MNAR missingness, the bias was still beyond the acceptable limit for some of the variables, showing that none of the methods considered accounted adequately for missing data under a missing data mechanism that cannot be ruled out when considering cognitive function in old age.

Life course SEP and adult crystallized cognitive function

The last objective of Aim 1 was to investigate the life course hypotheses on crystallized cognitive function. In the majority of the life course analyses in Chapter 6, an accumulation model provided the best fit to the data, often including all three time points but sometimes only the two adult time points. Again, the results depended on the SEP variables used, cohort and gender. As in earlier analyses, when there were differences in results between the missing data methods, the Heckman selection model produced different results to the other two approaches.

Childhood SEP and memory trajectory

The objectives in the first half of Aim 2 were to investigate the impact of childhood SEP on both the intercept and rate of decline of memory trajectories, before and after adjusting for later life SEP. In the unadjusted multiple imputation analyses in Chapter 8 childhood material deprivation was a significant predictor of the intercept of the memory trajectory, but father's occupational SEP was not. Childhood material deprivation remained a significant predictor of memory intercept after adjustment for later life SEP. Neither of the childhood SEP measures were significant predictors of rate of memory decline in the unadjusted analyses, although childhood material deprivation was a significant predictor in the complete case analysis.

Life course SEP and memory trajectory

The fourth objective of aim 2 was to investigate the life course hypotheses with respect to memory decline. When father's occupational SEP was used as the measure of childhood SEP, the adult accumulation model was the life course model identified in Chapter 8, whereas all of the life course models were significantly different to the saturated model when childhood material deprivation was the measure of childhood SEP. This meant that none of the life course models were complex enough to explain the relationship between SEP and the memory intercept. The effect of life course SEP on the memory slope depended on the childhood SEP measure used, with none of the life course models sufficiently complex when childhood material deprivation was used, and no effect of SEP when father's occupational SEP was used. Different conclusions were reached in many of these analyses when multiple imputation was used.

9.2 Relevance of thesis

Life expectancy has been increasing in Europe since 1950 (187), leading to an increased number of elderly people, and poor cognitive function is one of the most disabling conditions in old age (184). As cognitive function in old age depends on both the peak cognitive function reached earlier in life and the rate of cognitive decline, it is important to investigate which characteristics predict both cognitive function earlier in life and rate of cognitive decline. It is also important to know when cognitive decline begins, as interventions are more likely to succeed if they are implemented when individuals first start to experience cognitive decline (184).

There is also evidence that cognitive decline does not only occur immediately before a diagnosis of dementia, with lower scores found in those with a diagnosis of dementia up to ten years prior to the diagnosis (188). It is therefore important to investigate the trajectory of cognitive function throughout the life course. The work in this thesis extends the previous work, described in Chapter 1.

Identifying when in the life course social inequalities in cognition occur may lead to better targeting of interventions to improve cognition of those of lower socioeconomic position. The Marmot review (189) emphasized the importance of reducing the social gradient in health, focussing not only on the most disadvantaged, but on individuals throughout the gradient, scaled according to their level of disadvantage. In order for this to occur, methods of identifying the level of disadvantage are required. Life course SEP is one measure of disadvantage for adults, and investigating the life course models highlighted the relative importance of each SEP variable at each stage of the life course. This work has found that SEP at all three stages of the life course are important predictors of cognitive function. Hence policies aimed at reducing social inequalities at all life stages could be effective.

The work presented here could also be used as a starting point from which to investigate the characteristics and behaviours that explain the inequalities in cognitive decline. This information could be used to design interventions which could then be tested. For example, some occupations involve more routine work than others, leading to a narrower range of cognitive tasks being carried out during the working life. Owen et al (190) found that ‘brain training’ activities only improved performance in the cognitive tasks that had been practiced, but did not transfer to other closely linked cognitive tasks. It is therefore likely that different aspects of SEP impact on different cognitive functions; for example those with higher educational qualifications may have higher scores on tests such as the NART due to exposure to a wider range of vocabulary.

As participants in longitudinal datasets grow older, opportunities become available to investigate the life course influences of cognitive function and cognitive decline. This thesis has taken advantage of the data offered by two such studies. However, missing data is very common in epidemiological datasets, especially in such longitudinal datasets. Missing data are often ignored, and complete case analyses are carried out. Recently there has been an increased focus on missing data, and methodology for accounting for missing data has been implemented in common statistical packages, including Stata, R, SAS and SPSS. It is often unclear exactly how the statistical software carries out the analyses after the code has been entered. This can lead to the inappropriate use of missing data methodology, and has even led to incorrect results being published, and later corrected (116). Moreover most published articles that deal with missing data through the use of imputation methods do not provide details of how the imputation analyses were carried out. It is thus not possible to assess whether the imputation has been carried out appropriately.

This thesis has considered in detail the implementation of multiple imputation and Heckman selection in two complex longitudinal datasets and has shown that considerable work is required prior to running either a multiple imputation or Heckman selection analysis correctly. To identify an appropriate imputation model it is first necessary to identify which variables will be in the model of interest, including the form of the variables, such as squared and interaction terms. The next step is to test which variables from the data available are predictive of missing data in any of the variables in the model of interest, in order to make sure that the MAR assumption is met. Then the variables that were predictive of missing data are tested for their predictive power of the variables in the model of interest. Finally, auxiliary variables are included, which improve the prediction of the variables in the imputation model. It is better to include auxiliary variables which are not causes of missingness rather than risk missing any (121). As shown in the analyses in this thesis, this can lead to large imputation models. When implementing multiple imputation by chained equations in Stata, the correct type of regression must be specified for each variable in the imputation model, for example logistic regression for binary variables, and ordered logistic regression for ordered categorical variables. It is also necessary to choose how many imputations to carry out, and whether to passively calculate imputed values of squared or interaction terms, so a squared term is equal to the initial value squared, or whether to allow the squared term to have a value that is not equal to the initial value squared. Although passive calculations have not been investigated in this thesis, research has shown that constraining the value to equal the initial value squared introduces bias (122).

There are no clear guidelines for how to choose a selection model for a Heckman selection analysis. It is necessary to include a variable which is a significant predictor of selection, but is not an independent variable in the model of interest. However some consider a stricter condition to be necessary (134); requiring a variable associated with selection, but not associated with the dependent variable in the model of interest. The selection model satisfying the stricter condition performed slightly better in the simulation study in Chapter 5. From the work carried out above, a wide range of variables should be tested as predictors of selection. Variables that are missing a large proportion of data should be dropped, as only the complete cases are used in the selection model. The criterion of the largest proportion of significant variables is sensible for identifying a final selection model, as the inverse Mills ratio is calculated using all the variables included in the selection model.

The simulation study carried out in Chapter 5 compared the results from complete case, multiple imputation and Heckman selection analyses, and demonstrated the advantages of using multiple imputation when analysing longitudinal datasets, where numerous appropriate variables are available for the imputation model, improving the likelihood that the missingness is MAR.

9.3 Implications of findings

Epidemiological

In the NSHD father's occupational SEP remained a significant predictor of crystallized cognitive function independently of later life SEP, whereas childhood household amenities did not. Father's occupational SEP may be indicative of advantageous effects on the household environment that improves cognitive stimulation that childhood household amenities do not. Previous research has suggested that the relationship between childhood SEP and childhood cognitive development is associated with levels of access to services as well as positive and negative experiences (191). Multiple aspects of the home environment have also been considered, such as learning stimulation, parental responsiveness and the number of books on the shelves (191); factors which are likely to be associated with the father's occupation. A measure of chaos in the household was found to be correlated with childhood cognitive ability, after controlling for SEP (192).

This would imply that the home environment was more important in relation to childhood cognitive development than the material conditions. Hence, advice could be provided to parents regarding ways of improving the level of cognitive stimulation provided in the household; this already exists to a certain extent through encouraging parents to read with their children through organisations such as the National Literacy Trust (193). Further work is required to investigate the pathways through which the childhood SEP variables act (section 9.5), and the results from this would provide more suggestions of potential ways to reduce the difference between those in the more and less advantaged situations.

One mechanism that has been suggested is genetic factors; intelligence is highly heritable (161), and the father's level of intelligence is likely to be associated with his later life SEP. There is evidence of gene-environment interactions influencing cognitive decline, such as the interaction between APOE-e4 status and education (194). Behavioural and social factors, such as alcohol consumption and smoking status, have been found on the pathway between childhood SEP and adult cognitive function, but these would not explain the relationship that exists between childhood SEP and childhood cognitive function (195). Anaemia caused by iron-deficiency, as well as malnutrition in general, has been suggested as one mechanism which could affect the brain's development; studies in numerous countries have found that severe iron-deficiency in infancy is associated with lower test scores, indicating that the deficiency at a critical period of growth may have caused irreversible abnormalities (196). Many factors interact with childhood SEP when considered with respect to later life outcomes, such as environmental lead exposure, which has been associated with cognitive deficits due to its toxic impact on the developing nervous system, with a worse impact among those from a lower SEP background (197). The health and development of preterm children is more likely to be affected if the child is from a low SEP background (37).

It is feasible that the mechanisms may vary depending on the childhood SEP measure. One study that has demonstrated the variation in pathways linking different childhood SEP measures and childhood cognitive ability investigated potential mechanisms through which mother's education, occupation and income related to childhood academic achievement (198). The pathways considered were parenting practices, school behaviour, and skill activities at home. There was no direct effect of income on academic achievement, the direct effect of education was mediated by skill activities at home, and the direct effect of occupation was not mediated.

The life course methodology developed by Mishra et al (89) represents progress from the previous methodology available for investigating the life course hypotheses, by identifying one model that provides the best fit to the data, after considering whether each model was significantly different to the saturated model. Both steps are important; identifying whether each model is significantly different from the saturated model, and identifying the best of those models which are not significantly different from the saturated model. Accumulation hypotheses were supported when investigating crystallized cognitive function, again highlighting the importance of SEP at each stage of the life course. This thesis presents one of the first uses of this methodology in a practical situation. It is through the application of new methodologies that potential issues are likely to be discovered, and extensions developed.

The life course methodology was very flexible, allowing additional hypotheses to be tested. Additional accumulation and social mobility life course hypotheses were tested. An adult accumulation model was added, to allow SEP at more than one time point to influence the outcome, without requiring childhood SEP to have an impact, as well as a mutually adjusted accumulation model, which removed the constraint that all of the SEP variables have the same coefficient from the accumulation model. The additional social mobility model added a three-way interaction term to the social mobility model which had the same effect of upward mobility, regardless of when it occurred, and similarly the same effect of downward mobility, regardless of when it occurred. The three-way interaction distinguished the two groups who were not social mobile; those who remained in the lower SEP category at all three time points, and those who remained in the higher SEP category at all three time points. The observed scores showed that these groups had very different cognitive scores, in both the NSHD and Whitehall II study.

The effect on findings of weighting the SEP measures to allow for the proportion of time spent in each stage of the life course was considered. In the Whitehall II study the older participants had spent a much greater proportion of their life in the later adult stage of the life course compared to younger participants; weighting allowed the impact of this on the life course model to be investigated. In the examples used here, in general weighting had little impact on the results. However, this may not be the case in other cohorts or with other outcomes. Further, only one possible weighting strategy was examined for each dataset, and there are other weighting combinations which would also be realistic. For example, when determining the period of the life course which educational qualifications influence, the period will depend on how long each individual was in full time education, with participants who were not in full time education for as long likely to be influenced by their occupational SEP from an earlier age than those who remained in full time education for longer.

Very few studies had investigated the impact of childhood SEP on cognitive decline (section 1.2.7). This study examined the effect of two childhood SEP variables, with different conclusions reached depending on how missing data were accounted for. Father's occupational SEP was not a significant predictor of rate of memory decline in either the complete case or multiple imputation analysis; however childhood material deprivation was a significant predictor in the complete case analysis, but not the multiple imputation analysis. Although the simulation study did not consider the missing data methods in a multilevel environment, multilevel modelling assumes the missing data are MAR when all the variables which make the missingness MAR are included in the model; as that is not the case in this situation, the multiple imputation results are more trustworthy.

In chapter 8, it was demonstrated how the life course methodology can also be applied to situations beyond the linear regression analyses that it was initially developed for. The life course methodology was applied to a multilevel model, allowing the life course hypotheses to be investigated with respect to trajectories, rather than outcomes at only one time point. When implementing the life course methodology using multilevel models it is possible to test the life course effect on the intercept and the life course effect on the slope, as well as constraining the same life course model for both the intercept and the slope, to investigate whether there is one life course model determining the whole shape of the trajectory.

The life course hypotheses had not previously been compared with respect to memory decline, and the results were dependent on the childhood SEP variable considered. When childhood material deprivation was included in the model, none of the life course models were complex enough to describe the relationship with either the intercept or slope, so the saturated model was required. However the adult accumulation model was the life course model identified to model the intercept of the memory trajectory when father's occupational SEP was included in the model, with no effect of SEP at any stage of the life course found on the slope. These new findings highlight the large differences that can be found when different SEP measures are considered.

Missing data

It is important to investigate the reasons behind any missing data, and to do everything possible in the planning stages to limit the amount of missing data. However well planned the study is though, some missing data are inevitable and hence will need to be dealt with in the analyses.

As shown in Figure 4.1 (NSHD) and Figure 4.3 (Whitehall II), there can be a large difference in the results depending on the missing data method applied. In Figure 4.1 there were differences in the size of the regression coefficients and therefore the magnitude of the effect of childhood SEP, whereas in Figure 4.3 there was also a difference in whether childhood SEP was a significant predictor of adult cognitive function after adjusting for later life SEP (with only the Heckman selection model finding a significant association). In the analyses carried out in Chapter 4 the results were mainly different for the Heckman selection models, whereas the complete case and MID results were similar.

Although it was not possible to test which model provided the best fit to the data for MID in the life course analyses, there were differences between the missing data methods in whether individual life course models were significantly different from the saturated model, and different final models were selected between the complete case and Heckman selection analyses.

For the multilevel models carried out in Chapter 8, only complete case and MID models were compared. There was one difference in significance between the complete case and MID results where father's occupational SEP was a significant predictor of memory score in the complete case analysis, but not in the MID analysis. Similarly childhood material deprivation was a significant predictor of rate of memory decline in the complete case analysis, but not the MID analysis.

These are some of the examples found in this thesis where different conclusions were reached depending on the method used to account for missing data. Differences also occurred in the magnitude of the effect size found, as well as the significance of the variables. These situations highlight the importance of appropriately accounting for missing data, and putting in the required thought when producing an imputation or selection model to produce results that can be trusted. Otherwise results may be produced which lead to incorrect conclusions being drawn, which in turn may have serious consequences when trying to initiate interventions based on the flawed results, or otherwise make use of the results.

Multiple imputation can now be implemented in many statistical packages, and is relatively easy to carry out. This simplicity may also be a disadvantage, as it is thus easy to carry out without the appropriate preparation. It requires time and thought to identify potential variables for the imputation model, and then test whether each variable is associated with missingness and the variables in the model of interest. It is necessary to consider the complexities of the analysis before finalising the imputation model, such as how to treat squared and interaction terms, as described in section 9.2. Sufficient time must be allowed to carry out each stage of the imputation process described in section 2.4. Additionally, if missing data is anticipated when setting up a study, additional data which other studies have found predictive of dropout should be collected at baseline, so that appropriate variables are available for carrying out a multiple imputation analysis, increasing the probability of the missing data mechanism being MAR.

Heckman selection models were not suitable for the complex missing data found in the longitudinal cohort studies analysed. Numerous variables were associated with missing data in both the NSHD and Whitehall II studies. However, the selection part of the Heckman model, used to calculate the inverse Mills ratio, is carried out on complete data only. Therefore a limited number of variables can be included in the selection model when the missing data are not restricted to the dependent variable, in order to have a reasonable sample size for the selection model. Previous work has found that Heckman's selection model performed better than multiple imputation when only one independent variable had missing data, but that it performed worse when missing data occurred in many of the independent variables (199). Although the Heckman model will only converge with a limited number of variables in the selection model when using the command in Stata, it is possible to carry out a Heckman selection analysis in two stages, first the selection model, then calculating the inverse Mills ratio, and inserting the inverse Mills ratio into the analysis of interest. This would allow for a larger range of variables to be included in the selection model, although the issue of a complete case analysis being carried out in the selection model would remain.

Multiple imputation performed well for MAR missingness in the simulation study, although the Heckman selection models did not perform well when the missing data mechanism was MAR or MNAR. In longitudinal cohort studies with complex patterns of missing data, where a selection model is unlikely to be appropriate and where MNAR is likely, sensitivity analyses would be advisable. Such analyses would investigate the effect of different assumptions for the missing dependent variable values. This is discussed in more detail in section 9.5.

9.4 Strengths and limitations

The strengths and limitations of each study have been addressed in the discussion section of each chapter; this section will therefore focus on overarching strengths and limitations.

The use of two large cohort studies to investigate the effect of life course SEP on cognitive function is a major strength of the study, as the studies have different strengths and weaknesses. The NSHD collected childhood SEP data prospectively, whereas in the Whitehall II study the childhood data were collected retrospectively. The participants in the Whitehall II study were different ages when the retrospective data was collected, and the participant's age may impact on the reliability of the retrospective data. As the Whitehall II participants were born in a range of years, the effect of having a certain childhood SEP position or level of childhood material deprivation may have been different depending on birth year. In future work this could be addressed by considering an interaction between year of birth and childhood SEP, although the relationship may well not be linear.

The results from the two cohort studies are not directly comparable as there were differences in the measures used for both crystallized cognitive function and SEP, such as the childhood material deprivation/household conditions variables. Although one of the studies was a birth cohort study, the sample was not updated to allow for immigration, and the makeup of England, Scotland and Wales is very different now to how it was in 1946. The other cohort used in this thesis was an occupational cohort, and therefore a healthy worker cohort, which limits the generalizability of the results. However, if SEP inequalities are evident even in this more socially homogenous cohort of civil servants, the inequalities are likely to be evident in a population cohort, where a wider range of SEP levels exist (200). The importance of even relatively small differences in SEP is highlighted in the Whitehall II study.

A major strength of this thesis is the focus on methodology for dealing with missing data. Missing data is important in all studies, but especially longitudinal studies, where the same variables are often found to be predictive of who remains in and who drops out of the study at each phase. It is especially important in this study, as both SEP and cognitive function are predictive of dropout. If only the complete dataset is analysed, and certain characteristics do predict dropout, then the sample analysed is unlikely to represent the population of interest, limiting the usefulness of the study. Similarly, when modelling cognitive decline it is important to account for missing data, otherwise decline is likely to be underestimated due to increased dropout amongst those who experience the most severe decline.

The two main methods of analysing MNAR missing data involve pattern mixture models and selection models. Pattern mixture models are not appropriate when many different missing data patterns exist, as is the case in both the NSHD and Whitehall II. Therefore a selection model was chosen to be compared to the complete case and multiple imputation analyses. Heckman selection models are among the most common selection model, and despite their main usage being in the field of economics, they have also been used successfully in the social sciences (201;202).

The performance of the three missing data methods was compared using a simulation study, in order to understand better the analyses in the 2 cohorts. The simulation study followed the guidelines set out by Burton et al (172), whereas sufficient details are rarely provided in published simulation studies (172). The simulated data was as representative of the NSHD as possible and was more complex than many previous simulation studies; simulation studies are often simplified and contain only a few variables, whereas the simulation study in Chapter 5 contained sufficient variables to realistically simulate missingness under each of the three missing data mechanisms. The simulation study also used an imputation model developed following the methodology of Carpenter and Plewis (113), containing sufficient variables to have a high chance of meeting the assumption of MAR in the observed dataset the simulations were based on.

As there are no clear guidelines for choosing a selection model, two methods of choosing a Heckman selection model were compared. Three datasets were created from each simulated dataset in order to compare how each of the three missing data methods performed under each of the three missing data mechanisms. The simulation study had a clearly defined aim, and a sample size calculation was carried out, in order to ensure that there was the required level of accuracy for the interpretation of the results. The three missing data methods were compared using a measure of accuracy, bias and coverage, as the results may differ across criteria (121). The advantage of this is that the results of the simulation study can be trusted, and the missing data methods have been fairly compared for the situation under investigation. Without the simulation study, comparison of results from different methods is difficult as the true results remain unknown.

A limitation of any simulation study based on an observed dataset is that that it is based only on the data that were observed, rather than a complete dataset. Therefore all the measurements in the simulation study, such as the means and standard deviations, and relationships, such as correlations and the equations to predict dropout, are based by necessity on only the observed data. However this should not influence the interpretation of the simulation results, as all the missing data techniques were applied to the same simulated datasets.

This thesis did not investigate methodology for MNAR in the multilevel models, as it was not possible to implement the Heckman selection models for multilevel models using the same command as for the linear regression Heckman selection analyses in Stata.

This thesis could only consider cognitive decline over a period of ten years due to the data available; however there were sufficient data to show that cognitive decline is measurable before it may be noticed by people themselves, with memory decline observed from the mid-forties in chapter 8. Future waves of the Whitehall II study will enable decline over longer periods to be investigated, and future waves of the NSHD will enable cognitive decline to be investigated in a birth cohort study. Measures of both crystallized and fluid cognitive function were available. In the Whitehall II study SEP predictors of the intercept were investigated for both crystallized cognitive function (measured using the NART, chapter 4) and fluid cognitive function (measured using memory, chapter 8). Different conclusions were reached for the two measures of cognitive function, demonstrating that the effect of SEP differs by both SEP measures and cognitive measures.

There is on-going debate as to whether adjustment for later life SEP represents an overadjustment when investigating the relationship between childhood SEP and an outcome later in life. The analyses above first considered the total effect of childhood SEP, which some consider the most relevant effect (165). However, for others (203) it is also of interest to know whether an effect of childhood SEP on later life outcomes remains while holding constant the level of later life SEP. It is logical to focus the available resources on those individuals at increased risk; if an effect of childhood SEP remains after adjustment for later life SEP, those individuals with lower SEP during childhood should be prioritised over those in the same adult SEP category but who were in the higher SEP category during childhood. The effect of overadjustment would typically be to bias coefficients towards the null (166). However it is widely acknowledged that even if childhood SEP does not remain a significant predictor after adjusting for later life SEP, childhood SEP is always an important consideration for health in later life as SEP tracks through the life course (165). A related issue to that of overadjustment is collider bias, where the association between two variables is affected by conditioning on a common effect (167). Although there is the potential for collider bias to be introduced by the addition of adult SEP to the model, the collider variable (adult SEP) is only influenced by exposures and not by the outcome, due to the later time at which the outcome was measured. Therefore the impact of collider bias should not be large in this case (168;169).

Potential confounders or the pathways through which SEP at each stage of the life course may act on cognitive function or cognitive decline beyond later life SEP were not considered here. However this work is an important starting point, and future work could investigate the pathways, enabling more concrete implications and interventions to be developed. One method to do this would be to use latent growth curve models, as both cognitive function and SEP are latent variables which cannot be directly measured. Finally, in the analyses which considered cognitive function at one point in time, linear regression methods were used. Although both distributions were relatively normally distributed, an alternative analysis method would be to use quantile regression, which does not require the assumption of normality. The impact of using a different analysis method could be explored in future work.

9.5 Future work

The next step for this work is to investigate the pathways through which the various SEP measures affect cognitive function and cognitive decline, whilst maintaining a focus on missing data. Potential pathways include parental encouragement, characteristics of the school attended, childhood nutrition, health, and lifestyle factors such as diet, exercise and smoking status, as well as biological markers and genetic influences. Pathways could be investigated using latent growth curve modelling, in which SEP and cognitive function from more stages of the life course could be incorporated into the models in the NSHD analyses. Understanding the pathways that childhood SEP act through would also help with understanding the practical differences between being in the different categories for father's occupational SEP and childhood material deprivation. Additionally, from the analyses carried out in this thesis it is not possible to tell whether it is the educational qualifications themselves that have an effect on cognitive function, or the experiences that they represent, such as the networking opportunities at university, which may help with achieving a higher occupational SEP. Exploring the pathways could help clarify the aspects of SEP at each stage of the life course which influence cognitive function and cognitive decline, and the relative importance of these pathways. For example, by identifying the pathways each measure of childhood SEP acts through, it would be possible to develop potential interventions to reduce the social gradient, focussing on the variables in the pathway that are not experienced to the same level by those participants from lower SEP backgrounds, such as nutrition.

As the participants age and more data are collected it will also be possible to explore cognitive decline over a longer period, and investigate whether the same results are found, and whether SEP at each time point has a larger influence. The rate of memory decline differs by age group (184), and it is possible that different factors influence cognitive decline at the beginning of decline, including the age at which decline starts, and the rate of decline later in life. SEP factors from throughout the life course have been found in this thesis to influence the trajectory, but a further step would involve investigating whether different cognitive trajectories were found for participants with different SEP trajectories. With longer follow up and increasing age of participant, more deaths will occur. Hence, missing data due to death and dropout may need to be considered separately, as although some of the underlying predictors of missingness are the same for both, there are also likely to be differences. Sensitivity analyses could be investigated for situations where the missing data mechanism is likely to be MNAR, but the missing data mechanism is too complex for a selection model to be effective. For example, it could be assumed that each participant who had dropped out had the lowest observed cognitive score; another option would be to assume each participant had the highest observed cognitive score, although in practice this is unlikely. In this way the robustness of the findings to different assumptions can be assessed.

It would be of interest to compare the results attained from using the Heckman command in Stata and those attained from carrying out a Heckman selection model in two stages. This would allow a much larger range of variables to be included in the selection model, without the problems of convergence experienced when using the Heckman command in Stata. It would also be of interest to investigate the situations in which Heckman selection models perform adequately; whether the missing data occurred in only the dependent variable or both the dependent and independent variables, varying the true number of variables that predict selection, and the amount and mechanism of the missing data in the variables used for the selection model. Alternative joint modelling approaches could also be investigated.

Due to the simplification of making the missing data monotone, it was not possible to investigate the results of multiple imputation compared to multiple imputation, then deletion in the simulation study, however different conclusions were reached when applying the methods to observed data; therefore more work is required to clarify in which situations each of the methods are appropriate. It has been suggested that whether to use MI or MID depends on how much additional information is in the imputed outcome; von Hippel (123) implies that unless the information is 'quite substantial' then MID is advantageous; however it is not clear how to judge this. Young and Johnson have demonstrated that using MID rather than MI may be unnecessary in some situations using an observed dataset, and point out that in situations where an imputed dataset for a public survey is released, indicators would have to be provided for each variable to show exactly which values were imputed (204). Therefore a simulation study investigating MI and MID using different groups of auxiliary variables could be carried out, in order to further the current knowledge on which situations would benefit from use of an MID analysis over an MI analysis.

Paradata, such as method of data collection, or the number of attempts that were required to contact a participant, has recently become an area of increasing interest (205), and one area in which paradata could have an important impact is in the area of missing data; specifically paradata could be added to imputation models or be used in selection models. Future work could examine any benefit of adding paradata to these models.

Future work could model Heckman selection models in the multilevel setting using the `gllamm` command (206), or by carrying out each stage separately; first modelling a selection model, calculating the inverse Mills ratio from the selection model, and then running the model of interest. As Heckman selection models are designed for missing data in the dependent variable, and the dependent variable is measured multiple times in the longitudinal analyses, requiring a multilevel model in this thesis, the selection model would model selection into the multilevel model, for which a participant only requires the dependent variable to be measured at one time point. The ability of the model to deal with selective attrition beyond the selection into the model is currently limited by the assumptions that such models make regarding the nature of attrition over time (Sean Clouston, personal communication).

9.6 Conclusions

In general, findings suggest that SEP at each of the three stages of the life course affect crystallized cognitive function later in life. The results did however differ by SEP measure, cohort, gender, and missing data method. The results for memory decline were similarly dependent on the missing data methodology implemented, but in the fully adjusted model only occupational SEP influenced the rate of memory decline.

Missing data is a key issue in the analysis of data from longitudinal studies, and it has been shown that results can vary (in terms of significance of variables and model selection) depending on the approach taken. Multiple imputation is a very powerful method of allowing for missing data when implemented correctly, but it does require thought and preparation to carry out properly in packages such as Stata. The Heckman selection model was found to be inappropriate in situations with the complexity of missingness seen in the two longitudinal cohorts studied, and should thus be used with extreme caution in similar settings.

References

(1) Kuh D, Cooper R, Hardy R, Guralnik J, Richards M. Lifetime Cognitive Performance is Associated With Midlife Physical Performance in a Prospective National Birth Cohort Study. Psychosomatic Medicine 2009 Jan;71(1):38-48.

(2) Hatch SL, Jones PB, Kuh D, Hardy R, Wadsworth MEJ, Richards M. Childhood cognitive ability and adult mental health in the British 1946 birth cohort. Social Science & Medicine 2007 Jun;64(11):2285-96.

(3) Kuh D, Richards M, Hardy R, Butterworth S, Wadsworth MEJ. Childhood cognitive ability and deaths up until middle age: a post-war birth cohort study. International Journal of Epidemiology 2004 Apr;33(2):408-13.

(4) Korten AE, Jorm AF, Jiao Z, Letenneur L, Jacomb PA, Henderson AS, et al. Health, cognitive, and psychosocial factors as predictors of mortality in an elderly community sample. Journal of Epidemiology and Community Health 1999 Feb;53(2):83-8.

(5) Calvin CM, Deary IJ, Fenton C, Roberts BA, Der G, Leckenby N, et al. Intelligence in youth and all-cause-mortality: systematic review with meta-analysis. International Journal of Epidemiology 2011 Jun;40(3):626-44.

(6) Van Gelder BM, Tijhuis MAR, Kalmijn S, Giampaoli S, Kromhout D. Decline in cognitive functioning is associated with a higher mortality risk. Neuroepidemiology 2007;28(2):93-100.

(7) Lubinski D. Cognitive epidemiology: With emphasis on untangling cognitive ability and socioeconomic status. Intelligence 2009 Nov;37(6):625-33.

(8) Singh-Manoux A. Commentary: Is it time to redefine cognitive epidemiology? International Journal of Epidemiology 2010 Oct;39(5):1369-71.

(9) Small BJ, Backman L. Longitudinal trajectories of cognitive change in preclinical Alzheimer's disease: A growth mixture modeling analysis. Cortex 2007 Oct;43(7):826-34.

(10) Earles JL, Salthouse TA. Interrelations of Age, Health, and Speed. Journals of Gerontology Series B-Psychological Sciences and Social Sciences 1995 Jan;50(1):33-41.

(11) Tomalski P, Johnson MH. The effects of early adversity on the adult and developing brain. Current Opinion in Psychiatry 2010 May;23(3):233-8.

(12) Deary I. Why do intelligent people live longer? Nature 2008 Nov 13;456(7219):175-6.

(13) Osler M, Batty GD. Commentary: Influence of early life intelligence test performance on later health: do lower scoring children become less healthy adults? International Journal of Epidemiology 2004 Apr;33(2):414-5.

(14) Whalley LJ, Deary IJ. Longitudinal cohort study of childhood IQ and survival up to age 76. British Medical Journal 2001 Apr 7;322(7290):819-22.

(15) Gottfredson LS. Intelligence: Is it the epidemiologists' elusive "Fundamental cause" of social class inequalities in health? Journal of Personality and Social Psychology 2004 Jan;86(1):174-99.

(16) Singh-Manoux A, Ferrie JE, Lynch JW, Marmot M. The role of cognitive ability (intelligence) in explaining the association between socioeconomic position and health: Evidence from the Whitehall II prospective cohort study. American Journal of Epidemiology 2005 May 1;161(9):831-9.

(17) Batty GD, Der G, Macintyre S, Deary L. Does IQ explain socioeconomic inequalities in health? Evidence from a population based cohort study in the west of Scotland. British Medical Journal 2006 Mar 11;332(7541):580-3.

(18) Bosma H, van Boxtel MPJ, Kempen GIJM, van Eijk JTM, Jolles J. To what extent does IQ 'explain' socio-economic variations in function? Bmc Public Health 2007 Jul 25;7.

(19) Neisser U, Boodoo G, Bouchard TJ, Boykin AW, Brody N, Ceci SJ, et al. Intelligence: Knowns and unknowns. American Psychologist 1996 Feb;51(2):77-101.

(20) Deary IJ, Batty GD. Cognitive epidemiology. Journal of Epidemiology and Community Health 2007 May;61(5):378-84.

(21) Craik FIM, Bialystok E. Cognition through the lifespan: mechanisms of change. Trends in Cognitive Sciences 2006 Mar;10(3):131-8.

(22) Christensen H. What cognitive changes can be expected with normal ageing? Australian and New Zealand Journal of Psychiatry 2001 Dec;35(6):768-75.

(23) McGurn B, Starr JM, Topfer JA, Pattie A, Whiteman MC, Lemmon HA, et al. Pronunciation of irregular words is preserved in dementia, validating premorbid IQ estimation. Neurology 2004 Apr 13;62(7):1184-6.

(24) Wilson RS, Beckett LA, Barnes LL, Schneider JA, Bach J, Evans DA, et al. Individual differences in rates of change in cognitive abilities of older persons. Psychology and Aging 2002 Jun;17(2):179-93.

(25) Christensen H, Henderson AS, Griffiths K, Levings C. Does ageing inevitably lead to declines in cognitive performance? A longitudinal study of elite academics. Personality and Individual Differences 1997 Jul;23(1):67-78.

(26) Rabbitt P. Does It All Go Together When It Goes - the 19Th Bartlett Memorial Lecture. Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology 1993 Aug;46(3):385-434.

(27) Terrera GM, Matthews F, Brayne C. A comparison of parametric models for the investigation of the shape of cognitive change in the older population. Bmc Neurology 2008 May 16;8.

(28) Schaie KW. The Course of Adult Intellectual-Development. American Psychologist 1994 Apr;49(4):304-13.

(29) Starr JM, Whalley LJ, Inch S, Shering PA. The Quantification of the Relative Effects of Age and NART-Predicted IQ on Cognitive Function in Healthy Old-People. International Journal of Geriatric Psychiatry 1992 Mar;7(3):153-7.

(30) Deary IJ, Whiteman MC, Starr JM, Whalley LJ, Fox HC. The impact of childhood intelligence on later life: Following up the Scottish Mental Surveys of 1932 and 1947. Journal of Personality and Social Psychology 2004 Jan;86(1):130-47.

(31) Ho SC, Woo J, Sham A, Chan SG, Yu ALM. A 3-year follow-up study of social, lifestyle and health predictors of cognitive impairment in a Chinese older cohort. International Journal of Epidemiology 2001 Dec;30(6):1389-96.

(32) Richards M, Shipley B, Fuhrer R, Wadsworth MEJ. Cognitive ability in childhood and cognitive decline in mid-life: longitudinal birth cohort study. British Medical Journal 2004 Mar 6;328(7439):552-554B.

(33) Bourne VJ, Fox HC, Deary IJ, Whalley LJ. Does childhood intelligence predict variation in cognitive change in later life? Personality and Individual Differences 2007;42:1551-9.

(34) Gow AJ, Johnson W, Pattie A, Whiteman MC, Starr J, Deary IJ. Mental ability in childhood and cognitive aging. Gerontology 2008;54(3):177-86.

(35) Gow AJ, Johnson W, Pattie A, Brett CE, Roberts B, Starr JM, et al. Stability and Change in Intelligence From Age 11 to Ages 70, 79, and 87: The Lothian Birth Cohorts of 1921 and 1936. Psychology and Aging 2011 Mar;26(1):232-40.

(36) Rabbitt P, Chetwynd A, McInnes L. Do clever brains age more slowly? Further exploration of a nun result. British Journal of Psychology 2003 Feb;94:63-71.

(37) Bradley RH, Corwyn RF. Socioeconomic status and child development. Annual Review of Psychology 2002;53:371-99.

(38) Mercy JA, Steelman LC. Familial Influence on the Intellectual Attainment of Children. American Sociological Review 1982;47(4):532-42.

(39) White KR. The Relation Between Socio-Economic Status and Academic-Achievement. Psychological Bulletin 1982;91(3):461-81.

(40) Sirin SR. Socioeconomic status and academic achievement: A meta-analytic review of research. Review of Educational Research 2005;75(3):417-53.

(41) Bradley RH, Corwyn RF, WhitesideMansell L. Life at home: Same time, different places - An examination of the HOME Inventory in different cultures. Early Development & Parenting 1996 Dec;5(4):251-69.

(42) Firkowska A, Ostrowska A, Sokolowska M, Stein Z, Susser M, Wald I. Cognitive-Development and Social-Policy. Science 1978;200(4348):1357-62.

(43) Najman JM, Aird R, Bor W, O'Callaghan M, Williams GM, Shuttlewood GJ. The generational transmission of socioeconomic inequalities in child cognitive development and emotional health. Social Science & Medicine 2004 Mar;58(6):1147-58.

(44) Jefferis BJMH, Power C, Hertzman C. Birth weight, childhood socioeconomic environment, and cognitive development in the 1958 British birth cohort study. British Medical Journal 2002 Aug 10;325(7359):305-8.

(45) Lawlor DA, Batty GD, Morton SMB, Deary IJ, Macintyre S, Ronalds G, et al. Early life predictors of childhood intelligence: evidence from the Aberdeen children of the 1950s study. Journal of Epidemiology and Community Health 2005 Aug;59(8):656-63.

(46) Tong S, Baghurst P, Vimpani G, Mcmichael A. Socioeconomic position, maternal IQ, home environment, and cognitive development. Journal of Pediatrics 2007 Sep;151(3):284-8.

(47) Li L. Analysis of early life influences on cognitive development in childhood using multilevel models. Quaderni di Statistica 2008;10:99-113.

(48) Feinstein L. Inequality in the early cognitive development of children in the 1970 cohort. Economica 2003 Feb;70(277):73-97.

(49) Cerhan JR, Folsom AR, Mortimer JA, Shahar E, Knopman DS, McGovern PG, et al. Correlates of cognitive function in middle-aged adults. Gerontology 1998 Mar;44(2):95-105.

(50) Fuhrer R, Head J, Marmot MG. Social position, age, and memory performance in the Whitehall II study. Socioeconomic Status and Health in Industrial Nations 1999;896:359-62.

(51) Cagney KA, Lauderdale DS. Education, wealth, and cognitive function in later life. Journals of Gerontology Series B-Psychological Sciences and Social Sciences 2002 Mar;57(2):163-72.

(52) Lee S, Kawachi I, Berkman LF, Grodstein F. Education, other socioeconomic indicators, and cognitive function. American Journal of Epidemiology 2003 Apr 15;157(8):712-20.

(53) Zhao JH, Brunner EJ, Kumari M, Singh-Manoux A, Hawe E, Talmud PJ, et al. APOE polymorphism, socioeconomic status and cognitive function in mid-life - The Whitehall II Longitudinal Study. Social Psychiatry and Psychiatric Epidemiology 2005 Jul;40(7):557-63.

(54) Fritsch T, McClendon MJ, Smyth KA, Lerner AJ, Friedland RP, Larsen JD. Cognitive functioning in healthy aging: The role of reserve and lifestyle factors early in life. Gerontologist 2007 Jun;47(3):307-22.

(55) Hatch SL, Feinstein L, Link BG, Wadsworth MEJ, Richards M. The continuing benefits of education: Adult education and midlife cognitive ability in the British 1946 birth cohort. Journals of Gerontology Series B-Psychological Sciences and Social Sciences 2007 Nov;62(6):S404-S414.

(56) Wilson RS, Hebert LE, Scherr PA, Barnes LL, de Leon CFM, Evans DA. Educational attainment and cognitive decline in old age. Neurology 2009 Feb 3;72(5):460-5.

(57) Jorm AF, Rodgers B, Henderson AS, Korten AE, Jacomb PA, Christensen H, et al. Occupation type as a predictor of cognitive decline and dementia in old age. Age and Ageing 1998 Jul;27(4):477-83.

(58) Gallacher JEJ, Elwood PC, Hopkinson C, Rabbitt PMA, Stollery BT, Sweetnam PM, et al. Cognitive function in the Caerphilly study: Associations with age, social class, education and mood. European Journal of Epidemiology 1999 Feb;15(2):161-9.

(59) Aneshensel CS, Ko MJ, Chodosh J, Wight RG. The Urban Neighborhood and Cognitive Functioning in Late Middle Age. Journal of Health and Social Behavior 2011 Jun;52(2):163-79.

(60) Osler M, McGue M, Christensen K. Socioeconomic position and twins' health: a life-course analysis of 1266 pairs of middle-aged Danish twins. International Journal of Epidemiology 2007 Feb;36(1):77-83.

(61) Anstey K, Christensen H. Education, activity, health, blood pressure and apolipoprotein E as predictors of cognitive change in old age: A review. Gerontology 2000 May;46(3):163-77.

(62) Valenzuela MJ, Sachdev P. Brain reserve and cognitive decline: a non-parametric systematic review. Psychological Medicine 2006 Aug;36(8):1065-73.

(63) Albert MS, Savage CR, Blazer D, Jones K, Berkman L, Seeman T, et al. Predictors of cognitive change in older persons: MacArthur studies of successful aging. Psychology and Aging 1995 Dec;10(4):578-89.

(64) Christensen H, Korten AE, Jorm AF, Henderson AS, Jacomb PA, Rodgers B. Education and decline in cognitive performance: Compensatory but not protective. International Journal of Geriatric Psychiatry 1997 Mar;12(3):323-30.

(65) Deary IJ, MacLennan WJ, Starr JM. Is age kinder to the initially more able?: Differential ageing of a verbal ability in the healthy old people in Edinburgh study. Intelligence 1998;26(4):357-75.

(66) Leibovici D, Ritchie K, Ledesert B, Touchon J. Does education level determine the course of cognitive decline? Age and Ageing 1996 Sep;25(5):392-7.

(67) Ardila A, Ostrosky-Solis F, Rosselli M, Gomez C. Age-related cognitive decline during normal aging: The complex effect of education. Archives of Clinical Neuropsychology 2000 Aug;15(6):495-513.

(68) Schmand B, Smit J, Lindeboom J, Smits C, Hooijer C, Jonker C, et al. Low education is a genuine risk factor for accelerated memory decline and dementia. Journal of Clinical Epidemiology 1997 Sep;50(9):1025-33.

(69) Singh-Manoux A, Marmot M, Glymour MM, Sabia S, Kivimaki M, Dugravot A. Does Cognitive Reserve Shape Cognitive Decline? Annals of Neurology 2011.

(70) Dugravot A, Gueguen A, Kivimaki M, Vahtera J, Shipley M, Marmot MG, et al. Socioeconomic position and cognitive decline using data from two waves: what is the role of the wave 1 cognitive measure? Journal of Epidemiology and Community Health 2009 Aug;63(8):675-80.

(71) Lord FM. A Paradox in Interpretation of Group Comparisons. Psychological Bulletin 1967;68(5):304-5.

(72) Plewis I. Statistical methods for understanding cognitive growth: A review, a synthesis and an application. British Journal of Mathematical & Statistical Psychology 1996 May;49:25-42.

(73) Kaplan GA, Turrell G, Lynch JW, Everson SA, Helkala EL, Salonen JT. Childhood socioeconomic position and cognitive function in adulthood. International Journal of Epidemiology 2001 Apr;30(2):256-63.

(74) Maurer J. Height, education and later-life cognition in Latin America and the Caribbean. Economics & Human Biology 2010 Jul;8(2):168-76.

(75) Everson-Rose SA, de Leon CFM, Bienias JL, Wilson RS, Evans DA. Early life conditions and cognitive functioning in later life. American Journal of Epidemiology 2003 Dec 1;158(11):1083-9.

(76) Wilson RS, Scherr PA, Hoganson G, Bienias JL, Evans DA, Bennett DA. Early life socioeconomic status and late life risk of Alzheimer's disease. Neuroepidemiology 2005;25(1):8-14.

(77) Turrell G, Lynch JW, Kaplan GA, Everson SA, Helkala EL, Kauhanen J, et al. Socioeconomic position across the lifecourse and cognitive function in late middle age. Journals of Gerontology Series B-Psychological Sciences and Social Sciences 2002 Jan;57(1):S43-S51.

(78) Zhang ZX, Plassman BL, Xu Q, Zahner GEP, Wu B, Gai MY, et al. Lifespan influences on mid- to late-life cognitive function in a Chinese birth cohort. Neurology 2009 Jul 21;73(3):186-94.

(79) Wilson RS, Scherr PA, Bienias JL. Socioeconomic characteristics of the community in childhood and cognition in old age. Experimental Aging Research 2005 Oct;31(4):393-407.

(80) Packard CJ, Bezlyak V, Mclean JS, Batty GD, Ford I, Burns H, et al. Early life socioeconomic adversity is associated in adult life with chronic inflammation, carotid atherosclerosis, poorer lung function and decreased cognitive performance: a cross-sectional, population-based study. Bmc Public Health 2011 Jan 17;11.

(81) Richards M, Sacker A. Lifetime antecedents of cognitive reserve. Journal of Clinical and Experimental Neuropsychology 2003 Aug;25(5):614-24.

(82) Singh-Manoux A, Richards M, Marmot M. Socioeconomic position across the lifecourse: How does it relate to cognitive function in mid-life? Annals of Epidemiology 2005 Sep;15(8):572-8.

(83) Johnson W, Gow AJ, Corley J, Starr JM, Deary IJ. Location in cognitive and residential space at age 70 reflects a lifelong trait over parental and environmental circumstances: The Lothian Birth Cohort 1936. Intelligence 2010 Jul;38(4):402-11.

(84) Britton A, Singh-Manoux A, Marmot M. Alcohol consumption and cognitive function in the Whitehall II study. American Journal of Epidemiology 2004 Aug 1;160(3):240-7.

(85) Richards M, Wadsworth MEJ. Long term effects of early adversity on cognitive function. Archives of Disease in Childhood 2004 Oct;89(10):922-7.

(86) Kuh D, Ben-Shlomo Y, Lynch J, Hallqvist J, Power C. Life course epidemiology. Journal of Epidemiology and Community Health 2003 Oct;57(10):778-83.

(87) Hallqvist J, Lynch J, Bartley M, Lang T, Blane D. Can we disentangle life course processes of accumulation, critical period and social mobility? An analysis of disadvantaged socio-economic positions and myocardial infarction in the Stockholm Heart Epidemiology Program. Social Science & Medicine 2004 Apr;58(8):1555-62.

(88) Lynch J. Social position and health. Annals of Epidemiology 1996 Jan;6(1):21-3.

(89) Mishra G, Nitsch D, Black S, De Stavola B, Kuh D, Hardy R. A structured approach to modelling the effects of binary exposure variables over the life course. International Journal of Epidemiology 2009 Apr;38(2):528-37.

(90) Luo Y, Waite LJ. The impact of childhood and adult SES on physical, mental, and cognitive well-being in later life. Journals of Gerontology Series B-Psychological Sciences and Social Sciences 2005 Mar;60(2):S93-S101.

(91) Long JA, Ickovics JR, Gill TM, Horwitz RI. The cumulative effects of social class on mental status decline. Journal of the American Geriatrics Society 2001 Jul;49(7):1005-7.

(92) Lynch JW, Kaplan GA, Cohen RD, Kauhanen J, Wilson TW, Smith NL, et al. Childhood and Adult Socioeconomic-Status As Predictors of Mortality in Finland. Lancet 1994 Feb 26;343(8896):524-7.

(93) Batty GD, Lawlor DA, Macintyre S, Clark H, Leon DA. Accuracy of adults' recall of childhood social class: findings from the Aberdeen children of the 1950s study. Journal of Epidemiology and Community Health 2005 Oct;59(10):898-903.

(94) Enders CK. The performance of the full information maximum likelihood estimator in multiple regression models with missing data. Educational and Psychological Measurement 2001 Oct;61(5):713-40.

(95) Matthews FE, Chatfield M, Freeman C, McCracken C, Brayne C. Attrition and bias in the MRC cognitive function and ageing study: an epidemiological investigation. Bmc Public Health 2004 Apr 27;4.

(96) Smith GD, Bartley M, Blane D. The Black Report on Socioeconomic Inequalities in Health 10 Years on. British Medical Journal 1990 Aug 18;301(6748):373-7.

(97) Pamuk ER. Social-Class Inequality in Mortality from 1921 to 1972 in England and Wales. Population Studies-A Journal of Demography 1985;39(1):17-31.

(98) Bartley M. Health Inequality: an introduction to theories, concepts and methods. Polity; 2004.

(99) Chapin FS. A quantitative scale for rating the home and social environment of middle-class families in an urban community: A first approximation to the measurement of socioeconomic status. Journal of Educational Psychology 1928;19:99-111.

(100) Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: Concepts, methodologies, and guidelines. Annual Review of Public Health 1997;18:341-78.

(101) Grundy E, Holt G. The socioeconomic status of older adults: How should we measure it in studies of health inequalities? Journal of Epidemiology and Community Health 2001 Dec;55(12):895-904.

(102) Galobardes B, Shaw M, Lawlor DA, Lynch JW, Smith GD. Indicators of socioeconomic position (part 1). Journal of Epidemiology and Community Health 2006 Jan;60(1):7-12.

(103) Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M, et al. Socioeconomic status in health research - One size does not fit all. Jama-Journal of the American Medical Association 2005 Dec 14;294(22):2879-88.

(104) White KR. The Relation Between Socio-Economic Status and Academic-Achievement. Psychological Bulletin 1982;91(3):461-81.

(105) Macintyre S, Mckay L, Der G, Hiscock R. Socio-economic position and health: what you observe depends on how you measure it. Journal of Public Health Medicine 2003 Dec;25(4):288-94.

(106) O'Reilly D. Standard indicators of deprivation: do they disadvantage older people? Age and Ageing 2002 May;31(3):197-202.

(107) Glymour MM, Manly JJ. Lifecourse Social Conditions and Racial and Ethnic Patterns of Cognitive Aging. Neuropsychology Review 2008 Sep;18(3):223-54.

(108) Deary IJ, Corley J, Gow AJ, Harris SE, Houlihan LM, Marioni RE, et al. Age-associated cognitive decline. British Medical Bulletin 2009 Dec;92(1):135-52.

(109) Clarke P, Hardy R. Methods for Handling Missing Data. In: Pickles A, Maughan B, Wadsworth M, editors. Epidemiological Methods in Life Course Research. Oxford University Press; 2007. p. 157-79.

(110) Ryu E, Couper MP, Marans RW. Survey incentives: Cash vs. in-kind; face-to-face vs. mail; Response rate vs. nonresponse error. International Journal of Public Opinion Research 2006;18(1):89-106.

(111) Little RJ, Rubin DB. Statistical Analysis with Missing Data. John Wiley & Sons, Inc.; 2002.

(112) Kenward MG, Carpenter J. Multiple imputation: current perspectives. Stat Methods Med Res 2007 Jun;16(3):199-218.

(113) Carpenter J, Plewis I. Analysing Longitudinal Studies with Non-response: Issues and statistical methods. In: Willians M, Vogt WP, editors. The SAGE Handbook of Innovation in Social Research Methods. SAGE Publications Ltd; 2011.

(114) Carpenter J, Kenward MG. Missing Data 2009Available from: URL: http://www.lshtm.ac.uk/msu/missingdata/

(115) Plewis I, Carpenter J. Course notes: Handling missing data in longitudinal studies (CCSR). 2008.
Ref Type: Pamphlet

(116) Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. British Medical Journal 2009 Jun 29;339.

(117) Rubin D. Multiple imputation for nonresponse in surveys. John Wiley & Sons; 1987.

(118) Royston P. Multiple imputation of missing values. The Stata Journal 2004;4(3):227-41.

(119) Royston P. Multiple imputation of missing values: Update of ice. The Stata Journal 2005;5(4):527-36.

(120) Van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. Statistics in Medicine 1999 Mar 30;18(6):681-94.

(121) Collins LM, Schafer JL, Kam CM. A comparison of inclusive and restrictive strategies in modern missing data procedures. Psychological Methods 2001 Dec;6(4):330-51.

(122) von Hippel PT. How to Impute Interactions, Squares and Other Transformed Variables. Sociological Methodology 2009, Vol 39 2009;39:265-91.

(123) von Hippel PT. Regression with Missing Ys: An Improved Strategy for Analyzing Multiply Imputed Data. Sociological Methodology 2007, Vol 37 2007;37:83-117.

(124) Schafer JL. Multiple imputation: a primer. Statistical Methods in Medical Research 1999 Mar;8(1):3-15.

(125) Bodner TE. What Improves with Increased Missing Data Imputations? Structural Equation Modeling-A Multidisciplinary Journal 2008 Oct;15(4):651-75.

(126) Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? - Some practical clarifications of multiple imputation theory. Prevention Science 2007 Sep;8(3):206-13.

(127) Allison P. Missing Data. Thousand Oaks, CA: Sage; 2002.

(128) Mare RD, Chen MD. Further Evidence on Sibship Size and Educational Stratification. American Sociological Review 1986 Jun;51(3):403-12.

(129) Statacorp. Stata 11 Base Reference Manual. College Station, TX: Stata Press; 2009.

(130) Briggs DC. Causal inference and the Heckman model. Journal of Educational and Behavioral Statistics 2004;29(4):397-420.

(131) Winship C, Mare RD. Models for Sample Selection Bias. Annual Review of Sociology 1992;18:327-50.

(132) Lalonde RJ. Evaluating the Econometric Evaluations of Training-Programs with Experimental-Data. American Economic Review 1986 Sep;76(4):604-20.

(133) Powers DE, Rock DA. Effects of coaching on SAT I: Reasoning test scores. Journal of Educational Measurement 1999;36(2):93-118.

(134) Sartori AE. An estimator for some binary-outcome selection models without exclusion restrictions. Political Analysis 2003;11(2):111-38.

(135) Douglas JWB. The Home and The School. Panther Books Ltd; 1976.

(136) Wadsworth MEJ, Butterworth SL, Hardy RJ, Kuh DJ, Richards M, Langenberg C, et al. The life course prospective design: an example of benefits and problems associated with study longevity. Social Science & Medicine 2003 Dec;57(11):2193-205.

(137) Wadsworth M, Kuh D, Richards M, Hardy R. Cohort profile: The 1946 National Birth Cohort (MRC National Survey of Health and Development). International Journal of Epidemiology 2006 Feb;35(1):49-54.

(138) Wadsworth MEJ, Mann SL, Rodgers B, Kuh DJL, Hilder WS, Yusuf EJ. Loss and Representativeness in A 43 Year Follow-Up of A National Birth Cohort. Journal of Epidemiology and Community Health 1992 Jun;46(3):300-4.

(139) David Rose. Official Social Classifications in the UK. Social Reseach Update 1995;(9).

(140) Crawford JR, Parker DM, Stewart LE, Besson JAO, Delacey G. Prediction of WAIS IQ with the National Adult Reading Test - Cross-Validation and Extension. British Journal of Clinical Psychology 1989 Sep;28:267-73.

(141) Marmot MG, Shipley MJ, Rose G. Inequalities in Death - Specific Explanations of A General Pattern. Lancet 1984;1(8384):1003-6.

(142) van Rossum CTM, Shipley MJ, van de Mheen H, Grobbee DE, Marmot MG. Employment grade differences in cause specific mortality. A 25 year follow up of civil servants from the first Whitehall study. Journal of Epidemiology and Community Health 2000 Mar;54(3):178-84.

(143) Whitehall II. website, 2012 [cited 2010 Aug 8];Available from: URL: http://www.ucl.ac.uk/whitehallII/study-phases

(144) Ferrie JE, Kivimaki M, Singh-Manoux A, Shortt A, Martikainen P, Head J, et al. Non-response to baseline, non-response to follow-up and mortality in the Whitehall II cohort. International Journal of Epidemiology 2009 Jun;38(3):831-7.

(145) Stansfeld S, Head J, Bartley M, Fonagy P. Social position, early deprivation and the development of attachment. Social Psychiatry and Psychiatric Epidemiology 2008 Jul;43(7):516-26.

(146) Kubinger KD. On artificial results due to using factor analysis for dichotomous variables. Psychology Science 2003;45(1):106-10.

(147) Marmot MG, Shipley MJ. Do socioeconomic differences in mortality persist after retirement? 25 year follow up of civil servants from the first Whitehall study. British Medical Journal 1996 Nov 9;313(7066):1177-80.

(148) Raven JC. Guide to Using the Mill Hill Vocabulary Scale with the Progressive Matrices Scales. Oxford, England: H. K. Lewis & Co.; 1958.

(149) Berney LR, Blane DB. Collecting retrospective data: Accuracy of recall after 50 years judged against historical records. Social Science & Medicine 1997 Nov;45(10):1519-25.

(150) Ferrer E, Salthouse TA, Stewart WF, Schwartz BS. Modeling age and retest processes in longitudinal studies of cognitive abilities. Psychology and Aging 2004 Jun;19(2):243-59.

(151) Rabbitt P, Diggle P, Smith D, Holland F, McInnes L. Identifying and separating the effects of practice and of cognitive ageing during a large longitudinal study of elderly community residents. Neuropsychologia 2001;39(5):532-43.

(152) Thorvaldsson V, Hofer SA, Berg S, Johansson B. Effects of repeated testing in a longitudinal age-homogeneous study of cognitive aging. Journals of Gerontology Series B-Psychological Sciences and Social Sciences 2006 Nov;61(6):348-54.

(153) Duff K, Beglinger LJ, Moser DJ, Paulsen JS, Schultz SK, Arndt S. Predicting Cognitive Change in Older Adults: The Relative Contribution of Practice Effects. Archives of Clinical Neuropsychology 2010 Mar;25(2):81-8.

(154) VanderWeele TJ. On the Relative Nature of Overadjustment and Unnecessary Adjustment. Epidemiology 2009 Jul;20(4):496-9.

(155) Persons JB, Burns DD, Perloff JM. Predictors of Dropout and Outcome in Cognitive Therapy for Depression in A Private-Practice Setting. Cognitive Therapy and Research 1988 Dec;12(6):557-75.

(156) Smart RG, Gray G. Multiple Predictors of Dropout from Alcoholism-Treatment. Archives of General Psychiatry 1978;35(3):363-7.

(157) Richards M, Hardy R, Wadsworth MEJ. Does active leisure protect cognition? Evidence from a national birth cohort. Social Science & Medicine 2003 Feb;56(4):785-92.

(158) Sturgis P, Read S, Allum N. Does intelligence foster generalized trust? An empirical test using the UK birth cohort studies. Intelligence 2010 Jan;38(1):45-54.

(159) Dodge KA, Pettit GS, Bates JE. Socialization Mediators of the Relation Between Socioeconomic-Status and Child Conduct Problems. Child Development 1994 Apr;65(2):649-65.

(160) Mcloyd VC. Socioeconomic disadvantage and child development. American Psychologist 1998 Feb;53(2):185-204.

(161) Plomin R. Genetics and general cognitive ability. Nature 1999 Dec 2;402(6761):C25-C29.

(162) Douglas JWB. Ability and Adjustment of Children Who Have Had Measles. British Medical Journal 1964;2(542):1301-3.

(163) Koletzko B, Aggett PJ, Bindels JG, Bung P, Ferre P, Gil A, et al. Growth, development and differentiation: a functional food science approach. British Journal of Nutrition 1998 Aug;80:S5-S45.

(164) Parks SE, Housemann RA, Brownson RC. Differential correlates of physical activity in urban and rural adults of various socioeconomic backgrounds in the United States. Journal of Epidemiology and Community Health 2003 Jan;57(1):29-35.

(165) Glymour M, Ertel K, Berkman L. What can life-course epidemiology tell us about inequalities in old age? In: Antonucci TC, Jackson JS, editors. Annual Review of Gerontology and Geriatrics: Life course perspectives on late-life health inequalities.New York: 2009.

(166) Schisterman EF, Cole SR, Platt RW. Overadjustment Bias and Unnecessary Adjustment in Epidemiologic Studies. Epidemiology 2009 Jul;20(4):488-95.

(167) Glymour MM. Using causal diagrams to understand common problems in social epidemiology. In: Oakes M, Kaufman J, editors. Methods in Social Epidemiology. 2006.

(168) Chaix B, Leal C, Evans D. Neighborhood-level Confounding in Epidemiologic Studies Unavoidable Challenges, Uncertain Solutions. Epidemiology 2010 Jan;21(1):124-7.

(169) Greenland S. Quantifying biases in causal models: Classical confounding vs collider-stratification bias. Epidemiology 2003 May;14(3):300-6.

(170) Fox-Wasylyshyn SM, El-Masri MM. Handling missing data in self-report measures. Research in Nursing & Health 2005 Dec;28(6):488-95.

(171) Robertson SL. Globalising UK Higher Education. the Centre for Learning and Life Chances in Knowledge Economies and Societies at: http://www.llakes.org; 2010.

(172) Burton A, Altman DG, Royston P, Holder RL. The design of simulation studies in medical statistics. Statistics in Medicine 2006 Dec 30;25(24):4279-92.

(173) Schafer JL, Graham JW. Missing data: Our view of the state of the art. Psychological Methods 2002 Jun;7(2):147-77.

(174) Kristman VL, Manno M, Cote P. Methods to account for attrition in longitudinal data: do they work? A simulation study. Eur J Epidemiol 2005;20(8):657-62.

(175) Marshall A, Altman DG, Royston P, Holder RL. Comparison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study. Bmc Medical Research Methodology 2010 Jan 19;10.

(176) Young R, Johnson DR. A Comparison of Four Methods for Handling Missing Secondary Respondent Data. 2011 Jun 6; 2011.

(177) White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. Statistics in Medicine 2010 Dec 10;29(28):2920-31.

(178) Steyerberg EW, van Veen M. Imputation is beneficial for handling missing data in predictive models. Journal of Clinical Epidemiology 2007 Sep;60(9):979.

(179) Spicker P. An introduction to social policy 2012 [cited 2012 Jan 25];Available from: URL: http://www2.rgu.ac.uk/publicpolicy/introduction/education.htm

(180) Stata help file for mim [computer program]. 2011.

(181) Li KH, Raghunathan TE, Rubin DB. Large-Sample Significance Levels from Multiply Imputed Data Using Moment-Based Statistics and An F-Reference Distribution. Journal of the American Statistical Association 1991 Dec;86(416):1065-73.

(182) Buis ML. Multiple imputation and BIC 2007Available from: URL: http://www.stata.com/statalist/archive/2007-04/msg00346.html

(183) Hall BH, Mairesse J, Turner L. Identifying age, cohort and period effects in scientific research productivity: discussion and illustration using simulated and actual data on French physicists. 2005.
Ref Type: Unpublished Work

(184) Singh-Manoux A, Kivimaki M, Glymour MM, Elbaz A, Berr C, Ebmeier KP, et al. Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. British Medical Journal 2012 Jan 5;344.

(185) UCLA: Academic Technology Services SCG. Stata FAQ: How can I perform multiple imputation on longitudinal data using ICE? 2012 [cited 2012 Jan 11];Available from: URL: http://www.ats.ucla.edu/stat/stata/faq/mi_longitudinal.htm

(186) Carpenter J. Missing data in multilevel models: perspectives on available methodology. 2004.
Ref Type: Slide

(187) Leon DA. Trends in European life expectancy: a salutary view. International Journal of Epidemiology 2011 Apr;40(2):271-7.

(188) Elias MF, Beiser A, Wolf PA, Au R, White RF, D'Agostino RB. The preclinical phase of Alzheimer disease - A 22-year prospective study of the Framingham cohort. Archives of Neurology 2000 Jun;57(6):808-13.

(189) Marmot M, Allen J, Goldblatt P, Boyce T, McNeish D, Grady M, et al. The Marmot Review: Fair Society, Healthy lives. Strategic review of health inequalities in England post 2010. London: The Marmot Review; 2010.

(190) Owen AM, Hampshire A, Grahn JA, Stenton R, Dajani S, Burns AS, et al. Putting brain training to the test. Nature 2010 Jun 10;465(7299):775-U6.

(191) Petrill SA, Pike A, Tom P, Plomin R. Chaos in the home and socioeconomic status are associated with cognitive development in early childhood: Environmental mediators identified in a genetic design. Intelligence 2004;32(5):445-60.

(192) Pike A, Iervolino AC, Eley TC, Price TS, Plomin R. Environmental risk and young children's cognitive and behavioral development. International Journal of Behavioral Development 2006 Jan;30(1):55-66.

(193) National Literacy Trust 2012 [cited 2012 Mar 10];Available from: URL: http://www.literacytrust.org.uk/

(194) Seeman TE, Huang MH, Bretsky P, Crimmins E, Launer L, Guralnik JM. Education and APOE-e4 in longitudinal cognitive decline: MacArthur studies of successful aging. Journals of Gerontology Series B-Psychological Sciences and Social Sciences 2005 Mar;60(2):74-83.

(195) Lupien SJ, King S, Meaney MJ, Mcewen BS. Can poverty get under your skin? Basal cortisol levels and cognitive function in children from low and high socioeconomic status. Development and Psychopathology 2001;13(3):653-76.

(196) Saloojee H, Pettifor JM. Iron deficiency and impaired child development - The relation may be causal, but it may not be a priority for intervention. British Medical Journal 2001 Dec 15;323(7326):1377-8.

(197) Needleman HL, Bellinger D. The Health-Effects of Low-Level Exposure to Lead. Annual Review of Public Health 1991;12:111-40.

(198) DeGarmo DS, Forgatch MS, Martinez CR. Parenting of divorced mothers as a link between social status and boys' academic outcomes: Unpacking the effects of socioeconomic status. Child Development 1999 Sep;70(5):1231-45.

(199) Krug G. Missing data in the record linkage of process and survey data : An empirical comparison of selected missing data techniques . Institute for Employment Research; 2009.

(200) Chandola T, Ferrie J, Sacker A, Marmot M. Social inequalities in self reported health in early old age: follow-up of prospective cohort study. British Medical Journal 2007 May 12;334(7601):990-993B.

(201) Kim DH, Schneider B. Social capital in action: Alignment of parental support in adolescents' transition to postsecondary education. Social Forces 2005 Dec;84(2):1181-206.

(202) Sleath B, Shih YCT. Sociological influences on antidepressant prescribing. Social Science & Medicine 2003 Mar;56(6):1335-44.

(203) McKenzie SK, Carter K, Blakely T, Collings S. The association of childhood socio-economic position and psychological distress in adulthood: is it mediated by adult socio-economic position? Longitudinal and Life Course Studies 2010;1(4):339-58.

(204) Young R, Johnson DR. Imputing the Missing Y's:Implications for Survey Producers and Survey Users. 2010.

(205) Nicolaas G. Survey paradata: a review. ESRC National Centre for Research Methods; 2011 Jan.

(206) Rabe-Hesketh S, Skrondal A, Pickles A. Multilevel selection models using gllamm 2002 [cited 2012 Mar 10];Available from: URL: http://www.stata.com/meeting/2dutch/select.pdf

(207) Lang IA, Llewellyn DJ, Langa KM, Wallace RB, Huppert FA, Melzer D. Neighborhood deprivation, individual socioeconomic status, and cognitive function in older people: Analyses from the English Longitudinal Study of Ageing. Journal of the American Geriatrics Society 2008 Feb;56(2):191-8.

(208) Christensen H, Hofer SM, Mackinnon AJ, Korten AE, Jorm AF, Henderson AS. Age is no kinder to the better educated: absence of an association investigated using latent growth techniques in a community sample. Psychological Medicine 2001 Jan;31(1):15-28.

(209) Tucker-Drob EA, Johnson KE, Jones RN. The Cognitive Reserve Hypothesis: A Longitudinal Examination of Age-Associated Declines in Reasoning and Processing Speed. Developmental Psychology 2009 Mar;45(2):431-46.

Appendices

Appendix 1: Papers on childhood cognitive function and later life cognitive decline

Reference	Sample (N)	Earlier measures of cognitive function used (retrospective/ prospective)	Later life measures of cognitive function used	Analysis method	Other variables considered	Overall Conclusions	Comments
Rabbitt, Chetwynd, McInnes, 2003 (36)	ESRC/MRC Manchester-Newcastle longitudinal study of cognitive change in later life, Age 49+ (N=3,263)	Youthful AH4 scores estimated from their Mill Hill vocabulary test scores taken later in life (Retrospective)	Heim AH4 group intelligence test, Mill Hill A vocabulary test	Linear regression (conditional model of change)	Age, gender, occupational class	Very weak positive association	- estimated childhood cognitive function from a test taken in adulthood - large age range
Richards, Shipley, Fuhrer, Wadsworth, 2004 (32)	National Survey of Health and Development England, Wales and Scotland (N=2,058)	Age 15: Heim AH4 test, Watts-Vernon reading test (Prospective)	Age 43 and 53: verbal memory and timed visual search	Linear regression (conditional model of change)	Education, occupation, NART at age 53, range of health variables	Verbal memory: Negative association for both AH4 and Watts-Vernon. Visual search: Men: AH4: no association for timed visual search. Watts-Vernon: no association for timed visual search after adjusting for education and	- disproportionate loss to follow up of those with low cognitive ability.

						occupation Women: AH4 and Watts-Vernon: no association for timed visual search after adjusting for NART (beyond education and occupation).	
Bourne, Fox, Deary, Whalley, 2007 (33)	Lothian birth cohorts 1921 and 1936 (N=91 and 349)	Age 11: Moray House Test of general intelligence (Prospective)	Raven's Progressive Matrices	Linear regression (conditional model of change)	Sex, education, occupation, cohort, interval between testing sessions, smoking status, alcohol intake	- Negative association - Older cohort showed more cognitive decline	- complete case analysis - selective sampling for follow up from original samples - small sample size for 1921 cohort
Gow, Johnson, Pattie, Whiteman, Starr, Deary, 2008 (34)	Lothian birth cohort 1921 (N=321)	Age 11: Moray House Test of general intelligence (Prospective)	Ages 79 and 83: Raven's Progressive Matrices, Verbal fluency and logical memory tests	Linear regression (outcome: residual of conditional model of change) and latent growth curve modelling	Sex, education, occupational social class, smoking status, alcohol consumption.	- Linear regression model: negative association. - Latent growth curve model: no association. - results from latent growth curve model more accurate – linear regression can't completely account for test-specific variance	- linear regression model: complete case analysis - latent growth curve model: FIML (MAR assumption) - selective sampling for follow up from original 1921 sample - cognitive change measured over 3.1 – 5.8 years
Gow, Johnson, Pattie, Brett, Roberts, Starr, Deary, 2011 (35)	Lothian birth cohorts 1921 (N=550)	Age 11: Moray House Test of general intelligence (Prospective)	Moray House Test at ages 79 and 87	Growth curve model	Sex, social class, education, smoking status, alcohol consumption	No association	- selective sampling for follow up from original 1921 sample, and selective attrition

Appendix 2: Papers on adult SEP and adult cognitive function

Reference	Sample (N)	Measures of SEP used	Measures of cognitive function used	Analysis method	Other variables considered	Overall conclusions	Missing data methodology applied	Comments
Cerhan, Folsom, Mortimer, Shahar, Knopman, McGovern, Hays, Crum, Heiss, 1998 (49)	Atherosclerosis Risk in Communities (ARIC) Study America, Aged 45-64 (N = 13,913)	Education, occupation	Delayed Word Recall Test, Digit Subscale of the Wechsler Adult Intelligence Scale-Revisited, Word Fluency test of the Multilingual Aphasia Examination	ANCOVA	Age, sex, marital status, depression	- Cognitive function was positively correlated with education level. - Managers/professionals had better cognitive function than other occupations.	Complete case analysis (data from 2nd wave used)	Potentially other variables adjusted for; unclear in article
Jorm, Rodgers, Henderson, Korten, Jacomb, Christensen, Mackinnon, 1998 (57)	Males aged 70+ from Canberra and Queanbeyan (N = 531 for cross-sectional, N = 329 for longitudinal)	Main occupation, years of education	Episodic Memory Test, Symbol-Letter Modalities Test, NART	hierarchical multiple linear regression, ANOVA	Age, native English	Cross-sectional: 'realistic' occupations had poorer cognitive performance.	Complete case analysis	- Only males
Fuhrer, Head, Marmot, 1999 (50)	Whitehall II (N = 3,398)	Education, employment grade	Short-term verbal memory,	Linear regression	Age, stratified by gender	- The significant effect of education on memory is removed when employment grade is included in the model, but	Complete case analysis	- Only introduced cognitive testing around halfway through 3rd

						employment grade remains significant.		wave of Whitehall II, unsure whether those who have cognitive data represent the full sample.
Gallacher, Elwood, Hopkinson, Rabbitt, Stollery, Sweetnam, Brayne, Huppert, 1999 (58)	Caerphilly study, Aged 55-69 (N = 1,870)	Occupational social class, education	AH4, choice reaction time, CAMCOG, NART	Linear regression	Age, mood at time of testing	- Cognitive function is positively associated with social class and education. - Social class and education are closely related but also make substantial independent contributions to cognitive function.	Complete case analysis, despite non-responders being older and more likely to be in a manual social class	- Both education and social class treated as continuous variables.
Cagney, Lauderdale, 2002 (51)	Asset and Health Dynamics Among the Oldest Old (AHEAD) America Nationally representative Aged 70+ (N = 6,577)	Wealth, household income, education	Memory (delayed and immediate), working memory, knowledge, language and orientation	Least squares regression, using a cluster correction to allow for spousal pairs from the same household	Age, gender	- Income and net worth have a much smaller impact on cognitive function than education. - Education and other SEP factors are not interchangeable with respect to cognitive function.	Complete case analysis	- Cross sectional sample, so can't disentangle an age-effect from a cohort effect. - Excluded proxy respondents
Lee, Kawachi, Berkman, Grodstein, 2003	Nurses' Health Study, (N = 19,319) Females only	Educational attainment	TICS, delayed recall, East Boston memory Test, verbal fluency, digit	Logistic regression	Age, husband's education, median household	- Decreasing odds of low cognitive function with increasing education	Complete case analysis	- Only females - Highly educated sample

(52)			span backwards		income, diabetes, blood pressure, heart disease, use of vitamin E, aspirin, postmenopausal hormones, BMI, smoking, alcohol, antidepressants, age at menopause, SF-36			
Rabbitt, Chetwynd, McInnes, 2003 (36)	Part of the ESRC/MRC Manchester/Newcastle longitudinal study of cognitive change Aged 49-92 (N = 3,263)	Occupation	Heim AH4, Mill Hill A vocabulary test	Regression models	Age, gender	- Occupational category is positively associated with cognitive function.	Complete case analysis	- large age range
Zhao, Brunner, Kumari, Singh-Manoux, Hawe, Talmud, Marmot, Humphries, 2005 (53)	Whitehall II (N=6,004)	Education, employment grade	Memory, AH4, Mill Hill, Phonetic fluency, Semantic fluency	Linear regression	APOE genotype, GHQ score (binary)	- Education and occupation associated with all 5 cognitive tests	Complete case analysis	- phase 5 cognitive tests, not accounting for some participants having previously taken the tests at phase 3
Bosma, van Boxtel, Kempen, van Eijk,	Maastricht Aging Study (MAAS), Netherlands Aged 24-81	Occupation	Bother due to forgetfulness in daily life	ANOVA	Parental education, father's occupation,	- No association between occupational level and cognitive functioning at baseline	Complete case analysis	- wide age range - measure of cognitive

Jolles, 2007 (18)	(N=1,211)				childhood deprivation, delayed developmental milestones			function
Fritsch, McClendon, Smyth, Lerner, Friedland, Larsen, 2007 (54)	Cleveland Longitudinal Aging Studies of Students (N=349)	Occupation, education	TICS, Logical Memory A subtest of the Wechsler Memory Scale, verbal fluency, Timed Months of the Year Backwards Test	SEM	High school records, high school activity level (mental, physical, social), parental occupation,	- Those with higher education had higher cognitive function - Occupational demands were not predictive of cognitive function	complete case analysis	- biasing effects of more men leaving high school due to World War II.
Hatch, Feinstein, Link, Wadsworth, Richards, 2007 (55)	National Survey of Health and Development (N=1,934)	Education, adult education and training	NART, verbal memory, verbal fluency, letter search	Multivariate regression	Gender, cognitive ability at 8 and 26,	- Education associated with all cognitive measures in adulthood. - Continued education associated with verbal memory, verbal fluency and NART	FIML	- adjusted for childhood cognitive ability
Osler, McGue, Christensen, 2007 (60)	Random sample of middle-aged Danish twins (N = 2,532) Only used like-sex twins	Type of employment, vocational education, number of subordinates	Verbal fluency, forward digit span, backward digit span, immediate recall, delayed recall, speeded digit symbol task	Odds ratios, t-test, MANOVA, chi-squared	zygosity, rearing social class	- Higher social class twin had higher cognitive test scores, only statistically significant for dizygotic male twins.	complete case analysis	- twin study – often have more social support throughout life
Lang, Llewellyn, Langa, Wallace, Huppert,	ELSA Aged 50+ Urban (N = 7,216)	Index of Multiple Deprivation 2004, income, wealth,	Immediate and delayed verbal memory, prospective memory, verbal	Regression	Age, sex, smoking, alcohol, diabetes, hypertension/	- Neighbourhood deprivation in urban areas is negatively associated with cognitive function in	complete case analysis	- excludes residents in institutions, those with no cognitive data

Melzer, 2008 (207)		education	fluency, letter cancellation task.		high blood pressure, visual problems, hearing loss, depression	older adults independently of their individual socioeconomic circumstances and level of education. - Individual level deprivation is negatively associated with cognitive function		recorded and those with only proxy results
Wilson, Hebert, Scherr, Barnes, de Leon, Evans, 2009 (56)	Chicago Health and Aging Project (N=6,533)	Education	Immediate and delayed recall, oral form of Symbol Digit Modalities Test, MMSE (global score)	Mixed effects models	Race, self-report heart attack, diabetes, hypertension, stroke, cancer	- Higher level of education was related to baseline cognition	complete case analysis	- unknown learning effect
Aneshensel, Ko, Chodosh, Wight, 2011 (59)	3rd wave of Health and Retirement Study, 1990 U.S. Census (N=4,525)	Individual level: Education, household wealth, household income. Neighbourhood -level: measure including proportion of residents >25 without a high school degree, households receiving	Multidimensional measure based on TICS	Hierarchical linear regression	Gender, age, race/ethnicity, marital status, employment status, social integration, health conditions, depression, self-rated health, smoking, drinking	- Education significant predictor, household wealth, household income, neighbourhood deprivation not significant when all in the model, wealth became significant when cross-level interactions were added, including a significant interaction between neighbourhood-level socioeconomic disadvantage and	complete case analysis	- sample not representative - neighbourhood boundaries defined by census, may not be how residents experience neighbourhood

		public assistance income, living below poverty level, >16 unemployed.				wealth. - When modelled individually, socioeconomic disadvantage is statistically significant.		

Appendix 3: Papers on adult SEP and adult cognitive decline

Reference	Sample (N)	Measures of SEP used	Measures of cognitive function used	Analysis method	Other variables considered	Overall Conclusions	Missing data methodology applied	Comments
Albert, Jones, Savage, Berkman, Seeman, Blazer, Rowe, 1995 (63)	Three cohorts of the EPESE (established populations for epidemiologic studies of the elderly) Age 70-79 at baseline US (N=1,011)	Income, education	Language, nonverbal memory, verbal memory, conceptualization, visuospatial ability	LISREL (linear structural relations modelling technique), using adjusted change scores	Age, race, gender, lifestyle (inc. smoking and alcohol), physical (inc. BMI), psychosocial (inc. life satisfaction) and physiologic (inc. lung function) variables	Education was a direct predictor of cognitive change, income was not	Complete case analysis	- Cognitive change over 2-2.5 years.
Leibovici, Ritchie, Ledesert, Touchon, 1996 (66)	Eugeria longitudinal study of cognitive ageing Aged 60+, manifesting recent subclinical deterioration in at least one area of cognitive decline France (N = 283)	Education	NART and tests representing six cognitive domains: attention, primary memory, secondary memory, implicit memory, visuospatial ability, language.	Principal components analysis, regression	Age group	- Elderly people with a high level of education show least decline in tests with a high learned component. - Level of education makes little difference to the rate of decline in tests which have a higher 'nature' rather than	Complete case analysis	- Sample all manifesting recent subclinical deterioration in at least one area of cognitive decline - Uses those from longitudinal study that had completed two stages of the study.

						'nurture' component. - Age has a strong effect on the rate of cognitive decline.		
Christensen, Korten, Jorm, Henderson, Jacomb, Rodgers, MacKinnon, 1997 (64)	Longitudinal study in Canberra and Queanbeyan, supplemented by sample from nursing homes and age ≥92. Australia Aged 70+ (N=652)	Education	Symbol Letter Modalities Test, Episodic Memory Test, choice reaction time, NART, vocabulary, similarities, information.	Hierarchical linear regression	Age, gender, previous medical conditions, ADL, activity levels	- Education a significant predictor of change in MMSE, NART, vocabulary, similarities and information (low education associated with greater decline in similarities, vocabulary and information, and a lack of improvement on the NART).	Observed outcome data, complete case independent variables	- Gender difference in those who remained in study.
Schmand,Smit, Lindeboom, Smits, Hooijer, Jonker, Deelman , 1997 (68)	Longitudinal Aging Study Amsterdam (N=1,774), The Amsterdam Study of the Elderly (N=4,051) Age 65+	Education	MAAS: MMSE, verbal memory, delayed recall LASA: MMSE, Abbreviated Mental Test	ANOVA	Stratified by age cohort	- For older ages, more cognitive decline for those with lower education. - Accelerated decline sets in at earlier age for lower	Complete case analysis	- 4 year follow up, so 'longitudinal' results made up of different cohorts. -estimated missing memory scores using most relevant correlates

						levels of education.		
Deary, MacLennan, Starr, 1998 (65)	Healthy Old People in Edinburgh study, Age 70+ (N=387)	Education, occupation	NART	ANOVA	Age, regular medication, major illness	- Participants with lower levels of education had more cognitive decline, non-manual professionals had less decline than manual workers.	Complete case analysis	- healthy sample at baseline
Jorm, Rodgers, Henderson, Korten, Jacomb, Christensen, Mackinnon, 1998 (57)	Males aged 70+ from Canberra and Queanbeyan (N = 531 for cross-sectional, N = 329 for longitudinal)	Main occupation, years of education	Episodic Memory Test, Symbol-Letter Modalities Test, NART	hierarchical multiple linear regression, ANOVA Measuring change: difference score approach rather than conditional regression	Age, native English	The occupational differences are in pre-morbid ability rather than in cognitive decline.	Complete case analysis	- Only males
Ardila, Ostrosky-Solis, Rosselli, Gòmez, 2000 (67)	Volunteers recruited from community centres from 5 states of the Mexican Republic Age 16-85 (N=806)	Education	NEUROPSI (orientation, attention and concentration, verbal memory, language, conceptual functions, motor functions, recall)	ANOVA	Age	- Higher education group declined slower for recall of words. - Low education group increased in	Complete case analysis	- Huge age range - Cross-sectional study - unclear sampling strategy

						the backwards digit task, but the higher education group decreased slightly. -Higher education group decreased, lower education group remained constant in semantic verbal fluency test.		
Christensen, Hofer, Mackinnon, Korten, Jorm, Henderson, 2001 (208)	Sample from the electoral roll of Canberra and Queanbeyan, Aged 70+ (N=294)	Education, income	Vocabulary, Similarities, NART, Word recognition, Recall of 3 words, Address recall, The Symbol Letter Modalities Test	Latent growth modelling, ANOVA, regression	Gender, current medical symptoms, past illnesses, history of stroke	- No association between level of education and rate of cognitive decline	Uses direct maximum likelihood (uses all available data)	- Did not consider retest effects
Lee, Kawachi, Berkman, Grodstein, 2003 (52)	Nurses' Health Study, (N=15,594) Females only	Education, household income	TICS, delayed recall, East Boston memory Test, verbal fluency, digit span backwards	Logistic regression Measuring change: difference score approach	Father's occupation, age, history of diabetes, high blood pressure, heart disease, use of vitamin	- Decreasing odds of cognitive decline with increasing education - Household	Complete case analysis	- Only females - decline defined as worst 10% of decline.

				rather than conditional regression	E supplements, use of aspirin, use of postmenopausal hormones, body mass index, smoking, alcohol consumption, antidepressants, age at menopause SF-36	income did not significantly affect cognitive decline		
Rabbitt, Chetwynd, McInnes, 2003 (36)	Part of the ESRC/MRC Manchester/Newcastle longitudinal study of cognitive change Aged 49-92 (N = 3,263)	Occupation	Heim AH4, Mill Hill A vocabulary test	Regression models	Age, gender	- Rate of cognitive decline is the same across occupational categories	Complete case analysis	- large age range - cross-sectional study, so not ideal for estimating decline
Valenzuela, Sachdev, 2006 (62)	Review of 18 studies	Education (13 studies), Occupation (4 studies)				Education: overall large and significant effect of education on cognitive decline Occupation: overall effect non-significant.		
Dugravot, Guéguen, Kivimaki, Vaheta, Shipley,	Whitehall II UK (N=1,744)	Employment grade	Verbal memory, phonetic and semantic fluency	Compares results of ANOVA and ANCOVA.	(none)	ANOVA: no effect of SEP on cognitive decline. ANCOVA:	complete case analysis	- doesn't account for practice effects/different number of times taken the test

Marmot, Singh-Manoux, 2009 (70)						significantly greater cognitive decline in lower SEP groups ANCOVA adjusted for measurement error: no association between SEP and cognitive decline		- restricted sample (age 50-55 at baseline)
Tucker-Drob, Johnson, Jones. 2009 (209)	Control group of Advanced cognitive training for independent and vital elderly study (ACTIVE) US Aged 65-89 at baseline (N=690)	Education	Reasoning (word series, letter series, letter sets) and processing speed	Latent growth curves	Age	Education was not related to cognitive change	FIML	- only 5 years follow up data, 2 measures.
Wilson, Hebert, Scherr, Barnes, de Leon, Evans, 2009 (56)	Chicago Health and Aging Project (N=6,533)	Education	Immediate and delayed recall, oral form of Symbol Digit Modalities Test	Mixed effects models	Race, self-report heart attack, diabetes, hypertension, stroke, cancer	- Higher level of education was not related to rate of cognitive decline. - When education was allowed to be nonlinear, rate of decline started slightly	complete case analysis	- unknown learning effect

						faster but was slower during later years of the follow up for those with higher levels of education		
Singh-Manoux, Marmot, Glymour, Sabia, Kivimaki, Dugravot, 2011 (69)	Whitehall II (N=7,454)	Education, occupation	AH4, verbal memory, phonetic and semantic fluency, Mill Hill vocabulary test	Linear mixed models	Age, stratified by gender	- Greater cognitive decline in the high occupation group, except for Mill Hill test. - No association between education and decline.	Used all available data	- education and occupation categorised into high/intermediate /low. - SEP measures examined separately - Working population sample

Appendix 4: Papers on childhood SEP and adult cognitive function

Reference	Sample (N)	Measures of childhood SEP used (retrospective?)	Measures of cognitive function used	Analysis method	Other variables considered	Overall Conclusions	Missing data methodology applied	Comments
Kaplan, Turrell, Lynch, Everson, Helkala, Salonen, 2001 (73)	Kuopio Ischaemic Heart Disease Risk Factor Study, East Finland, Males only Middle-aged (N=496)	composite measure of parents' education and principal occupation, also investigated each component of composite measure. (Retrospective)	Trail Making Test, Selective Reminding Test, Verbal Fluency Test, Visual Reproduction Test.	General linear models, with and without adjusting for participant's education	Education	- Significant graded positive association between composite childhood SEP and cognitive function, before and after adjustment for education. - Father's education and mother's occupation: no association. - Father's occupation: positive association, fully attenuated when adjust for education. - Mother's education: significant with and without	- Complete case analysis	- Recall bias - disproportionate attrition - Only males

						adjustment for education.		
Turrell, Lynch, Kaplan, Everson, Helkala, Kauhanen, Salonen, 2002 (77)	Kuopio Ischemic Heart Disease Risk Factor Study Finland (N = 486)	Parents' education and principal occupation, summed and categorized into tertiles. (retrospective)	Trail Making Test, Selective Reminding Test, Verbal Fluency Test, Visual Reproduction Test and MMSE.	Linear regression	Education, income, history and incidence of stroke, ischemic heart disease, atherosclerosis and diabetes, hypertension, blood lipids, fibrinogen, glucose, and insulin, medications for control of hypertension and cholesterol	- Each indicator of SEP: childhood SEP, education and income was were significantly positively associated with cognitive function.	- complete case analysis	- Recall bias - disproportionate attrition - only males
Everson-Rose, de Leon, Bienias, Wilson, Evans, 2003 (75)	Chicago Health and Aging Project, Aged 65+ (N=4,398)	composite index of paternal and maternal educational attainment, paternal occupational prestige, self-reported family financial status when the respondent was a child. (Retrospective)	Oral version of the Symbol Digit Modalities Test (a test of perceptual speed), immediate and delayed recall portions of the East Boston Story.	Mixed-effects regression models	Age, sex, race, childhood cognitive milieu (how frequently someone in the home read to, told stories to or played games with the respondent as a child.)	- Childhood SEP positively associated with cognitive function, before and after adjustment for education (and 'cognitive milieu').	- Analyses limited to respondents with cognitive function test scores from at least 2 of 3 interviews	- Follow up only 5 years - Recall bias
Lee, Kawachi, Berkman, Grodstein, 2003 (52)	Nurses' Health Study, (N=15,594) Females only	father's occupation at age 16 (retrospective)	delayed recall, East Boston memory Test, verbal fluency, digit span	Logistic regression	Age, educational attainment, husband's education, median household income, diabetes,	- childhood SEP was not significantly associated with adult cognitive function after	- Complete case analysis	- Only females - High-educational cohort, all nurses

			backwards		blood pressure, heart disease, use of vitamin E, aspirin, postmenopausal hormones, BMI, smoking, alcohol, antidepressants, age at menopause, SF-36	adjusting for adult SEP.		
Richards, Sacker, 2003 (81)	National Survey of Health and Development England, Wales and Scotland (N = 2,933 – 4,500)	Paternal occupation. (prospective)	At 53: NART, verbal memory, timed visual search.	Path modelling	educational attainment, current or last occupation at age 43, cognitive ability at age 8,	- Showed independent paths from childhood cognition, educational attainment and adult occupation to cognitive function as measured by the NART. The path from father's occupation to the NART was significant but 'substantially unimportant'.	- adjusts for missing data using FIML	- disproportionate attrition
Singh-Manoux, Richards, Marmot , 2005 (82)	Whitehall II, England Civil servants (N = 7,830)	Latent variable composed of mother's education, father's education, father's occupational social class and an indicator of financial	Cognitive function latent variable, composed of 5 tests: verbal memory, AH4-I, Mill	SEM – direct effects model (all indirect effects constrained to be zero)	Age, education, occupation and income	- There is no direct effect of childhood SEP on adult cognition after adjusting for adult SEP	- adjusts for missing data using FIML	- retrospective childhood SEP - does not adjust for childhood cognitive function - limited sample

		circumstances in childhood. (retrospective)	Hill vocabulary test, phonemic and semantic verbal fluency	and indirect effects.		- The indirect effects model provides a better fit to the data.		as Whitehall II is a study of white-collar workers
Wilson, Scherr, Bienias, 2005 (79)	Chicago Health and Aging Project, Aged 65+ (N=4,392)	Individual SEP: composite measure of parents' years of schooling, father's occupation, family financial situation during childhood (retrospective)	Oral version of the Symbol Digit Modalities Test (a test of perceptual speed), immediate and delayed recall portions of the East Boston Story, MMSE.	Linear mixed-effects models	County SEP in childhood: county average Duncan socioeconomic index for head of household, literacy rate for ages 6+, proportion aged 6-13 in school. Years of schooling, occupation, how often moved prior to age 16, born outside Cook county	Childhood household SEP wasn't significant after adjusting for adult SEP, but county-level childhood SEP was significant.	uses data from those who completed at least one follow-up interview	- Recall bias - Follow up only 5 years
Wilson, Scherr, Hoganson, Bienias, Evans, Bennett, 2005 (76)	Religious Orders Study, US (N=859)	Household SEP in childhood – composite measure of: mean parents' years of schooling, father's occupation, number of children in the family. (Retrospective)	19 cognitive tests grouped into 5 functional domains: episodic memory, semantic memory, working memory, perceptual speed and visuospatial	Mixed-effects models	Age, sex, county level SEP in childhood - composite measure of: county average Duncan socioeconomic index for head of household, literacy rate for ages 6+, proportion aged 6-13 in school.	Both early life county level and household level SEP were positively associated with level of cognitive function, before and after adjustment for education.	uses data from those who completed at least one follow-up interview	- Number of children giving equal importance to other variables in forming household SEP variable - Limited population

			ability.					
Zhang, Plassman, Xu, Zahner et al, 2009 (78)	Birth cohort, born at Peking Union Medical College Hospital from 1921-1954 Age 50-82 (N=2,062)	father's occupation (retrospective).	Fuld object memory evaluation, Fuld verbal fluency, Weschsler intelligence scale for children – revised block design, WAIS-R digit span	Logit models	Birth size, birth order, maternal age, childhood nutrition, height, education, alcohol, smoking, recreational activites, physical activities, cholesterol, diabetes, stroke, waist circumference, blood pressure	- Childhood SEP remained significant after adjusting for education and other adult variables.	Complete case analysis	- Wide age range - study sample
Johnson, Gow, Corley, Starr, Deary, 2010 (83)	Lothian Birth Cohort 1936 Age ~70 (N=1,091)	living conditions during childhood, father's education, father's occupation (retrospective)	Moray House Test	Structural equation modelling	Education, principal occupation (or husband's) prior to retirement, childhood IQ, neighbourhood environmental quality at age 70	- none of the childhood SEP variables were significantly associated with adult cognitive function after adjusting for education, adult SEP and child IQ. - Education and occupation had small effects after full adjustment	Used all available data	- retrospective childhood SEP and education, collected at age 70. - categorised occupation by job title
Packard, Bezlyak, McLean, Batty et al, 2011	Psychosocial and Biological Determinants of Ill-Health	number of siblings, parents owned home, father's occupation, bullied as child, owned car,	Stroop colour-word task (executive function), Choice	Linear regression	Income, education, home ownership, , trunk length BMI, inflammatory	- Each measure of childhood SEP not a significant predictor of cognitive	- complete case analysis	- sample selection – high levels of non-response

(80)	study Scotland Age 35-64 (N=666)	overcrowding, leg length (retrospective)	Reaction Time, memory		markers, smoker, cholesterol, blood pressure, lung function	function when adjusted for other childhood SEP measures.		

Appendix 5: Papers on childhood SEP and adult cognitive decline

Reference	Sample (N)	Measures of SEP used (childhood SEP retrospective?)	Measures of cognitive function used	Analysis method	Other variables considered	Overall Conclusions	Missing data methodology applied	Comments
Everson-Rose, de Leon, Bienias, Wilson, Evans, 2003 (75)	Chicago Health and Aging Project, Aged 65+ (N=4,398)	Childhood SEP: paternal and maternal educational attainment, paternal occupational prestige, self-reported family financial status when the respondent was a child. (Retrospective)	Oral version of the Symbol Digit Modalities Test (a test of perceptual speed), immediate and delayed recall portions of the East Boston Story, MMSE.	Mixed-effects regression models	Age, sex, race, childhood cognitive milieu (how frequently someone in the home read to, told stories to or played games with the respondent as a child), education	- no association	- Analyses limited to respondents with cognitive function test scores from at least 2 of 3 interviews	- Follow up only 5 years - Recall bias
Lee, Kawachi, Berkman, Grodstein, 2003 (52)	Nurses' Health Study, (N=15,594) Females only	father's occupation, (Retrospective)	TICS, delayed recall, East Boston memory Test, verbal fluency, digit span backwards	Logistic regression Measuring change: difference score approach rather than conditional regression	Educational attainment, husband's education, median household income, age, history of diabetes, high blood pressure, heart disease, use of vitamin E supplements, use of aspirin, use of postmenopausal	- Father's occupation had a small but significant effect on cognitive decline	- Complete case analysis	- Only females - decline defined as worst 10% of decline.

					hormones, body mass index, smoking, alcohol consumption, antidepressants, age at menopause SF-36			
Richards, Wadsworth 2004 (85)	National Survey of Health and Development England, Wales and Scotland (N=1,339)	SEP in childhood: material home conditions at age 4, maternal management and understanding at age 4, parental divorce. (Retrospective)	At 53: NART, verbal memory, timed visual search. At 43: verbal memory, timed visual search, timed peg placement	Multiple linear regression	Adult leg length, smoking, GHQ-28, paternal social class, maternal education, birth order	- No association between childhood SEP and cognitive decline	- Complete case analysis	- Disproportionate attrition - Ratings for childhood SEP were subjective - cognitive change from 43-53
Wilson, Scherr, Bienias, 2005 (79)	Chicago Health and Aging Project, Aged 65+ (N=4,392)	County SEP in childhood: county average Duncan socioeconomic index for head of household, literacy rate for ages 6+, proportion aged 6-13 in school. Individual SEP in childhood: parents' years of schooling, father's occupation (Retrospective)	Oral version of the Symbol Digit Modalities Test (a test of perceptual speed), immediate and delayed recall portions of the East Boston Story, MMSE.	Linear mixed-effects models	How often moved before age 16, education, principal lifetime occupation, born in Cook county, age, sex, race	- Neither early life county level nor household level SEP were associated with cognitive decline, either before or after adjusting for later life SEP.	- Only uses data from those who completed at least one follow-up interview	- Recall bias - Follow up only 5 years

Appendix 6: Papers on life course SEP and cognitive function

Reference	Sample (N)	Life course model tested	Measures of SEP used (childhood SEP retospective?)	Measures of cognitive function used	Analysis method	Other variables considered	Overall Conclusions	Missing data methodology applied	Comments
Turrell, Lynch, Kaplan, Everson, Helkala, Kauhanen, Salonen, 2002 (77)	Kuopio Ischemic Heart Disease Risk Factor Study Finland (N = 486)	Accumulation, social mobility	Childhood SEP: parents' education and principal occupation, summed and categorized into tertiles. SEP in adulthood: education and income (Retrospective)	Trail Making Test, Selective Reminding Test, Verbal Fluency Test, Visual Reproduction Test and MMSE.	Linear regression, adjusting for SEP at each time-point, looking at the SEP trajectories, and looking at cumulative SEP.	History and incidence of stroke, ischemic heart disease, atherosclerosis and diabetes, hypertension, blood lipids, fibrinogen, glucose, and insulin, medications for control of hypertension and cholesterol	- upwardly mobile participants had higher cognitive scores than those who had a steady low SEP; those who were downwardly mobile had a lower cognitive score than those who had a steady high SEP. - Cumulative SEP was significantly positively associated with cognitive function. - Each indicator of	- complete case analysis	- retrospective childhood SEP - does not adjust for childhood cognitive function - disproportionate attrition - cross-sectional design - only males

							SEP: childhood SEP, education and income was were significantly positively associated with cognitive function.		
Luo, Waite, 2005 (90)	Health and Retirement Study USA Aged 50+ (N = 19,949)	Accumulation, Social mobility	Childhood SEP: parents' education, father's occupation, how well off the family were financially. Adult SEP: education and household income (Retrospective)	Self-rated memory, a series of tests based on a modified version of the Telephone Interview for Cognitive Status (TICS).	Linear regression, adjusting for SEP at different time-points, looking at SEP mobility, and looking at cumulative SEP.	Gender, race/ethnicity, age	-Stable high and upwardly mobile participants had higher self-rated memory and cognitive function. - There is a cumulative effect of SEP on self-rated memory and cognitive function.	- complete case analysis	- retrospective childhood SEP - does not adjust for childhood cognitive function - self rated memory
Hatch, Feinstein, Link, Wadsworth, Richards, 2007 (55)	National Survey of Health and Development (N=1,934)	Intra-generational social mobility (age 26 to age 53)	Childhood SEP: father's occupation. Education, adult education and training (adjusted for childhood	NART, verbal memory, verbal fluency, letter search	Multivariate regression		- occupational social mobility had a significant effect on verbal ability, verbal memory and speed and	- FIML	

			cognitive function) (Prospective)				concentration at age 53, but not verbal fluency, after adjusting for childhood cognition, education, cognition at age 26 and adult education.			

Appendix 7: NSHD and Whitehall II complete case analyses, Aim 1

Table A7.1: NSHD complete case model development for men, with childhood household amenities as the measure of childhood SEP, and the outcome NART at age 53

MEN (N=893)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.463		0.628		0.440		0.522
Childhood - lacking 2 amenities	2.72 (2.20)	0.217	1.77 (1.87)	0.343	1.75 (1.81)	0.336	1.85 (1.28)	0.148	1.66 (1.29)	0.197
Childhood - lacking 1 amenity	4.41 (2.31)	0.056	2.12 (2.00)	0.289	1.68 (1.95)	0.389	1.16 (1.44)	0.420	0.95 (1.45)	0.511
Childhood - lacking no amenities	6.53 (2.16)	0.003	2.60 (1.86)	0.163	2.21 (1.80)	0.220	1.84 (1.27)	0.149	1.65 (1.30)	0.205
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			6.01 (0.99)	<0.001	4.58 (1.03)	<0.001	2.53 (0.89)	0.004	2.43 (0.89)	0.007
Education - GCE 'A'-Level			7.82 (0.88)	<0.001	6.04 (0.95)	<0.001	3.20 (0.84)	<0.001	3.18 (0.84)	<0.001
Education - Degree			14.37 (0.83)	<0.001	11.38 (0.97)	<0.001	6.22 (0.96)	<0.001	6.33 (0.95)	<0.001
Baseline: Age 43 own SEP – manual										
Age 43 own SEP - non-manual					4.60 (0.84)	<0.001	3.02 (0.71)	<0.001	2.99 (0.72)	<0.001
Standardised age 8 cognitive score							4.39 (0.34)	<0.001	4.31 (0.31)	<0.001
Standardised age 8 cognitive score squared									-0.25 (0.23)	0.287
Constant	28.93 (2.10)	<0.001	26.17 (0.83)	<0.001	24.94 (1.77)	<0.001	28.61 (1.24)	<0.001	29.05 (1.35)	<0.001
Model fit										
R-squared	0.0429		0.2618		0.3052		0.4570		0.4580	
BIC	6561.573		6349.988		6302.718		6089.343		6094.5	

Table A7.2: NSHD complete case model development for women, with childhood household amenities as the measure of childhood SEP, and the outcome NART at age 53

WOMEN (N=955)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.008		0.956		0.952		0.977		0.967
Childhood - lacking 2 amenities	1.63 (1.87)	0.385	0.93 (1.68)	0.580	0.84 (1.66)	0.612	0.15 (1.32)	0.912	0.06 (1.31)	0.963
Childhood - lacking 1 amenity	3.48 (1.95)	0.074	0.77 (1.71)	0.653	0.53 (1.70)	0.756	-0.16 (1.36)	0.905	-0.30 (1.35)	0.826
Childhood - lacking no amenities	4.07 (1.84)	0.027	0.87 (1.65)	0.599	0.73 (1.64)	0.656	0.08 (1.30)	0.951	-0.04 (1.30)	0.974
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			8.18 (0.70)	<0.001	7.37 (0.71)	<0.001	4.16 (0.65)	<0.001	4.09 (0.65)	<0.001
Education - GCE 'A'-Level			13.18 (0.71)	<0.001	11.97 (0.75)	<0.001	7.09 (0.70)	<0.001	7.17 (0.70)	<0.001
Education - Degree			18.31 (1.01)	<0.001	16.95 (1.05)	<0.001	8.26 (1.20)	<0.001	8.97 (1.17)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.80 (0.75)	<0.001	0.97 (0.65)	0.137	0.88 (0.66)	0.181
Standardised age 8 cognitive score							4.93 (0.30)	<0.001	4.80 (0.28)	<0.001
Standardised age 8 cognitive score squared									-0.42 (0.18)	0.022
Constant	29.61 (1.76)	<0.001	26.08 (1.62)	<0.001	24.78 (1.63)	<0.001	29.62 (1.28)	<0.001	30.17 (1.32)	<0.001
Model fit										
R-squared	0.0187		0.3511		0.3662		0.5378		0.5406	
BIC	7027.73		6653.281		6637.648		6343.016		6344.159	

Table A7.3: NSHD complete case model development for men, with father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

MEN (N=893)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Father's occupational SEP - manual										
Father's occ. SEP - non-manual	6.16 (0.71)	<0.001	2.78 (0.74)	<0.001	2.30 (0.71)	<0.001	0.72 (0.66)	0.274	0.74 (0.66)	0.258
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		0.010		0.008
Age 26 own SEP - IIIM			0.23 (1.32)	0.863	0.90 (1.26)	0.477	-0.07 (1.01)	0.942	-0.23 (1.01)	0.821
Age 26 own SEP - IIINM			7.15 (1.36)	<0.001	5.10 (1.39)	<0.001	2.82 (1.21)	0.020	2.73 (0.23)	0.026
Age 26 own SEP - I and II			8.17 (1.31)	<0.001	6.17 (1.32)	<0.001	2.43 (1.14)	0.033	2.40 (1.14)	0.035
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					4.67 (0.90)	<0.001	2.97 (0.77)	<0.001	2.89 (0.78)	<0.001
Standardised age 8 cognitive score							4.76 (0.34)	<0.001	4.67 (0.31)	<0.001
Standardised age 8 cognitive score squared									-0.29 (0.22)	0.199
Constant	31.83 (0.47)	<0.001	29.11 (1.18)	<0.001	27.16 (1.18)	<0.001	31.18 (0.95)	<0.001	31.58 (1.01)	<0.001
Model fit										
R-squared	0.0796		0.2149		0.2524		0.4372		0.4386	
BIC	6513.041		6391.39		6354.531		6107.771		6112.357	

Table A7.4: NSHD complete case model development for women, with father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

WOMEN (N=955)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Father's occupational SEP - manual										
Father's occ. SEP - non-manual	6.41 (0.67)	<0.001	4.10 (0.65)	<0.001	3.89 (0.62)	<0.001	1.76 (0.52)	0.001	1.81 (0.51)	<0.001
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP - IIIM			1.40 (1.44)	0.333	1.14 (1.40)	0.416	0.89 (1.05)	0.397	0.92 (1.03)	0.375
Age 26 own SEP - IIINM			5.15 (0.95)	<0.001	3.43 (0.97)	<0.001	1.85 (0.83)	0.027	1.75 (0.83)	0.036
Age 26 own SEP - I and II			11.46 (1.06)	<0.001	9.32 (1.08)	<0.001	4.80 (0.90)	<0.001	4.90 (0.91)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					4.34 (0.78)	<0.001	1.59 (0.68)	0.020	1.52 (0.69)	0.028
Standardised age 8 cognitive score							5.50 (0.29)	<0.001	5.39 (0.27)	<0.001
Standardised age 8 cognitive score squared									-0.41 (0.18)	0.022
Constant	30.88 (0.44)	<0.001	26.46 (0.83)	<0.001	24.79 (0.90)	<0.001	29.80 (0.77)	<0.001	30.24 (0.82)	<0.001
Model fit										
R-squared	0.0858		0.2372		0.2728		0.5078		0.5105	
BIC	6946.363		6794.026		6755.219		6389.33		6390.974	

Table A7.5: NSHD complete case model development for women, with father's occupational SEP, educational qualifications and head of household occupational SEP, and outcome NART

WOMEN (N=955)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Father's occupational SEP - manual										
Father's occ. SEP - non-manual	6.41 (0.67)	<0.001	2.33 (0.60)	<0.001	2.26 (0.60)	<0.001	1.04 (0.49)	0.036	1.07 (0.49)	0.029
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			7.85 (0.70)	<0.001	7.21 (0.71)	<0.001	3.90 (0.65)	<0.001	3.79 (0.66)	<0.001
Education - GCE 'A'-Level			12.34 (0.74)	<0.001	11.24 (0.76)	<0.001	6.46 (0.69)	<0.001	6.49 (0.69)	<0.001
Education - Degree			17.15 (1.05)	<0.001	15.67 (1.11)	<0.001	7.33 (1.20)	<0.001	8.05 (1.16)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					2.48 (0.66)	<0.001	1.60 (0.57)	0.005	1.60 (0.57)	0.005
Standardised age 8 cognitive score							4.86 (0.30)	<0.001	4.72 (0.28)	<0.001
Standardised age 8 cognitive score squared									-0.45 (0.18)	0.012
Constant	30.88 (0.44)	<0.001	26.60 (0.53)	<0.001	25.65 (0.60)	<0.001	29.39 (0.54)	<0.001	29.81 (0.59)	<0.001
Model fit										
R-squared	0.0858		0.3607		0.3746		0.5438		0.5470	
BIC	6946.363		6625.456		6611.19		6316.806		6317.042	

Table A7.6: NSHD complete case model development for women, with father's occupational SEP, own occupational SEP at age 26 and head of household occupational SEP at age 43, and outcome NART

WOMEN (N=955)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Father's occupational SEP - manual										
Father's occ. SEP - non-manual	6.41 (0.67)	<0.001	4.10 (0.65)	<0.001	3.77 (0.64)	<0.001	1.66 (0.52)	0.002	1.71 (0.52)	0.001
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP - IIIM			1.40 (1.44)	0.333	0.75 (1.43)	0.601	0.62 (1.06)	0.560	0.64 (1.04)	0.538
Age 26 own SEP - IIINM			5.15 (0.95)	<0.001	4.02 (0.93)	<0.001	1.84 (0.80)	0.021	1.71 (0.80)	0.033
Age 26 own SEP - I and II			11.46 (1.06)	<0.001	9.55 (1.04)	<0.001	4.51 (0.88)	<0.001	4.59 (0.87)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					3.87 (0.68)	<0.001	2.21 (0.58)	<0.001	2.22 (0.58)	<0.001
Standardised age 8 cognitive score							5.49 (0.28)	<0.001	5.36 (0.27)	<0.001
Standardised age 8 cognitive score squared									-0.44 (0.17)	0.011
Constant	30.88 (0.44)	<0.001	26.46 (0.83)	<0.001	25.32 (0.85)	<0.001	29.75 (0.73)	<0.001	30.19 (0.77)	<0.001
Model fit										
R-squared	0.0858		0.2372		0.2726		0.5145		0.5177	
BIC	6946.363		6794.026		6755.552		6376.291		6376.891	

Table A7.7: NSHD complete case model development for men, with childhood household amenities, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

MEN (N=893)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.046		0.109		0.177		0.224
Childhood - lacking 2 amenities	2.72 (2.20)	0.217	2.43 (1.99)	0.222	2.29 (1.91)	0.233	2.11 (1.29)	0.102	1.95 (1.30)	0.134
Childhood - lacking 1 amenity	4.41 (2.31)	0.056	3.23 (2.10)	0.126	2.61 (2.04)	0.199	1.62 (1.440	0.258	1.42 (1.44)	0.326
Childhood - lacking no amenities	6.53 (2.16)	0.003	4.15 (1.95)	0.033	3.62 (1.88)	0.054	2.60 (1.26)	0.040	2.43 (1.28)	0.058
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		0.005		0.004
Age 26 own SEP - IIIM			0.14 (1.33)	0.917	0.83 (1.27)	0.515	-0.12 (1.02)	0.903	-0.25 (1.02)	0.806
Age 26 own SEP - IIINM			7.39 (1.39)	<0.001	5.27 (1.40)	<0.001	2.79 (1.21)	0.021	2.72 (1.22)	0.026
Age 26 own SEP - I and II			8.67 (1.30)	<0.001	6.56 (1.31)	<0.001	2.44 (1.13)	0.031	2.42 (1.13)	0.032
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					4.72 (0.89)	<0.001	2.97 (0.77)	<0.001	2.91 (0.77)	<0.001
Standardised age 8 cognitive score							4.78 (0.36)	<0.001	4.71 (0.31)	<0.001
Standardised age 8 cognitive score squared									-0.24 (0.23)	0.295
Constant	28.93 (2.10)	<0.001	26.51 (2.28)	<0.001	24.83 (2.18)	<0.001	29.28 (1.50)	<0.001	29.78 (1.61)	<0.001
Model fit										
R-squared	0.0429		0.2134		0.2517		0.4406		0.4415	
BIC	6561.573		6406.763		6368.979		6116.007		6121.313	

Table A7.8: NSHD complete case model development for women, with childhood household amenities, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

WOMEN (N=955)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.008		0.393		0.492		0.820		0.833
Childhood - lacking 2 amenities	1.63 (1.87)	0.385	1.78 (1.68)	0.287	1.52 (1.68)	0.367	0.43 (1.30)	0.742	0.33 (1.30)	0.797
Childhood - lacking 1 amenity	3.48 (1.95)	0.074	1.90 (1.74)	0.275	1.45 (1.73)	0.405	0.24 (1.35)	0.860	0.10 (1.35)	0.938
Childhood - lacking no amenities	4.07 (1.84)	0.027	2.57 (1.66)	0.121	2.19 (1.65)	0.186	0.78 (1.29)	0.546	0.66 (1.29)	0.606
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP - IIIM			1.37 (1.45)	0.345	1.11 (1.41)	0.431	0.88 (1.05)	0.404	0.90 (1.03)	0.384
Age 26 own SEP - IIINM			5.51 (0.97)	<0.001	3.74 (0.99)	<0.001	1.95 (0.84)	0.021	1.87 (0.84)	0.027
Age 26 own SEP - I and II			12.59 (1.04)	<0.001	10.35 (1.090	<0.001	5.14 (0.91)	<0.001	5.25 (0.91)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					4.48 (0.80)	<0.001	1.58 (0.69)	0.022	1.53 (0.70)	0.028
Standardised age 8 cognitive score							5.65 (0.29)	<0.001	5.55 (0.27)	<0.001
Standardised age 8 cognitive score squared									-0.38 (0.18)	0.037
Constant	29.61 (1.76)	<0.001	25.06 (1.71)	<0.001	23.62 (1.74)	<0.001	29.63 (1.34)	<0.001	30.15 (1.37)	<0.001
Model fit										
R-squared	0.0187		0.2089		0.2469		0.5027		0.5050	
BIC	7027.73		6842.534		6802.475		6412.945		6415.366	

Table A7.9: NSHD complete case model development for women, with childhood household amenities, educational qualifications and head of household occupational SEP, and outcome NART

WOMEN (N=955)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.008		0.956		0.931		0.982		0.973
Childhood - lacking 2 amenities	1.63 (1.87)	0.385	0.93 (1.68)	0.580	1.09 (1.66)	0.511	0.28 (1.32)	0.832	0.19 (1.31)	0.885
Childhood - lacking 1 amenity	3.48 (1.95)	0.074	0.77 (1.71)	0.653	0.90 (1.69)	0.596	0.00 (1.35)	1.000	-0.15 (1.35)	0.913
Childhood - lacking no amenities	4.07 (1.84)	0.027	0.87 (1.65)	0.599	0.92 (1.64)	0.575	0.16 (1.30)	0.903	0.03 (1.30)	0.982
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			8.18 (0.70)	<0.001	7.51 (0.71)	<0.001	4.01 (0.65)	<0.001	3.92 (0.66)	<0.001
Education - GCE 'A'-Level			13.18 (0.71)	<0.001	12.04 (0.73)	<0.001	6.78 (0.70)	<0.001	6.84 (0.69)	<0.001
Education – Degree			18.31 (1.01)	<0.001	16.77 (1.06)	<0.001	7.75 (1.21)	<0.001	8.47 (1.17)	<0.001
Baseline: Age 43 HoH SEP – manual										
Age 43 HoH SEP - non-manual					2.55 (0.67)	<0.001	1.62 (0.57)	0.005	1.62 (0.57)	0.005
Standardised age 8 cognitive score							4.93 (0.30)	<0.001	4.79 (0.28)	<0.001
Standardised age 8 cognitive score squared									-0.44 (0.18)	0.014
Constant	29.61 (1.76)	<0.001	26.08 (1.62)	<0.001	24.99 (1.62)	<0.001	29.37 (1.28)	<0.001	29.90 (1.30)	<0.001
Model fit										
R-squared	0.0187		0.3511		0.3659		0.5420		0.5450	
BIC	7027.73		6653.281		6638.125		6334.342		6334.833	

Table A7.10: NSHD complete case model development for women, with childhood household amenities, own occupational SEP at age 26 and head of household occupational SEP at age 43, and outcome NART

WOMEN (N=955)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.008		0.393		0.495		0.874		0.892
Childhood - lacking 2 amenities	1.63 (1.87)	0.385	1.78 (1.68)	0.287	1.92 (1.68)	0.252	0.61 (1.31)	0.643	0.50 (1.30)	0.701
Childhood - lacking 1 amenity	3.48 (1.95)	0.074	1.90 (1.74)	0.275	1.99 (1.73)	0.248	0.46 (1.35)	0.735	0.31 (1.34)	0.819
Childhood - lacking no amenities	4.07 (1.84)	0.027	2.57 (1.66)	0.121	2.45 (1.65)	0.139	0.86 (1.29)	0.508	0.73 (1.29)	0.573
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP – IIIM			1.37 (1.45)	0.345	0.69 (1.44)	0.632	0.60 (1.06)	0.572	0.61 (1.03)	0.552
Age 26 own SEP – IIINM			5.51 (0.97)	<0.001	4.30 (0.94)	<0.001	1.92 (0.81)	0.017	1.81 (0.81)	0.025
Age 26 own SEP - I and II			12.59 (1.04)	<0.001	10.51 (1.02)	<0.001	4.81 (0.88)	<0.001	4.90 (0.88)	<0.001
Baseline: Age 43 HoH SEP – manual										
Age 43 HoH SEP - non-manual					4.10 (0.70)	<0.001	2.26 (0.58)	<0.001	2.27 (0.58)	<0.001
Standardised age 8 cognitive score							5.63 (0.28)	<0.001	5.51 (0.26)	<0.001
Standardised age 8 cognitive score squared									-0.42 (0.18)	0.020
Constant	29.61 (1.76)	<0.001	25.06 (1.71)	<0.001	23.78 (1.73)	<0.001	29.41 (1.33)	<0.001	29.95 (1.37)	<0.001
Model fit										
R-squared	0.0187		0.2089		0.2487		0.5099		0.5127	
BIC	7027.73		6842.534		6800.141		6398.965		6400.421	

Table A7.11: Whitehall II complete case model development for men, with childhood material deprivation as the measure of childhood SEP, and the outcome Mill Hill score at phase 9

Men (N=2,440)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e)	p-value
Childhood material deprivation	0.43 (0.12)	<0.001	0.25 (0.11)	0.027	0.21 (0.11)	0.046
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			2.72 (0.44)	<0.001	2.09 (0.42)	<0.001
Education: O-Level/A-Level/National diploma/Certificate			3.21 (0.30)	<0.001	2.45 (0.30)	<0.001
Education: University degree			4.95 (0.31)	<0.001	3.55 (0.31)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					2.64 (0.38)	<0.001
Last occ. grade - Unified Grades 1-6					4.36 (0.38)	<0.001
Age (phase 9)	0.00 (0.01)	0.820	0.03 (0.01)	0.010	0.02 (0.01)	0.186
No. of times taken cognitive tests	0.38 (0.09)	<0.001	0.25 (0.09)	0.005	0.20 (0.08)	0.019
Constant	23.47 (0.95)	<0.001	18.64 (0.98)	<0.001	17.41 (0.97)	<0.001
Model fit						
R-squared	0.0114		0.1263		0.2020	
BIC	13168.94		12887.74		13034.00	

Table A7.12: Whitehall II complete case model development for women, with childhood material deprivation as the measure of childhood SEP, and the outcome Mill Hill score at phase 9

Women (N=826)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e)	p-value
Childhood material deprivation	0.77 (0.26)	0.003	0.19 (0.24)	0.428	-0.15 (0.22)	0.513
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			0.93 (0.71)	0.194	0.24 (0.66)	0.718
Education: O-Level/A-Level/National diploma/Certificate			3.57 (0.43)	<0.001	2.06 (0.42)	<0.001
Education: University degree			6.33 (0.49)	<0.001	3.48 (0.52)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					3.20 (0.37)	<0.001
Last occ. grade – Unified Grades 1-6					5.61 (0.48)	<0.001
Age (phase 9)	-0.14 (0.03)	<0.001	-0.02 (0.03)	0.559	0.01 (0.03)	0.600
No. of times taken cognitive tests	0.12 (0.24)	0.628	0.36 (0.22)	0.105	0.27 (0.20)	0.176
Constant	31.28 (2.22)	<0.001	20.09 (2.22)	<0.001	17.56 (2.08)	<0.001
Model fit						
R-squared	0.0443		0.2168		0.3237	
BIC	4944.36		4803.10		4687.26	

Appendix 8: Developing the Whitehall II imputation model, Aim 1

Table A8.1: Identifying variables for the Whitehall II imputation model

phase variable collected		Significant predictor of							
		missing Mill Hill test score at phase 9		missing last recorded grouped occupational grade at phase 7		missing educational qualifications		missing childhood material deprivation	
		men	**women**	**men**	**women**	**men**	**women**	**men**	**women**
1	Year of birth	✔	✔	x	✔	x	✔	x	✔
1	Age finished full time education	✔	✔	✔	✔	✔	✔	x	✔
1	Accommodation type	✔	✔	✔	✔	✔	✔	✔	✔
1	Age mother finished full time education	x	✔	x	x	x	x	x	x
1	State of health in the last year	✔	✔	✔	✔	✔	✔	✔	✔
1	Any longstanding illnesses?	✔	✔	x	✔	x	x	✔	x
1	Smoking status	✔	✔	✔	✔	✔	✔	✔	✔
1	Job satisfaction	x	x	x	x	x	x	x	x
1	Usually pressed for time	✔	✔	✔	✔	✔	✔	✔	✔
1	Believe no one cares much about you	✔	x	x	x	x	x	x	x
1	Believe it is safer to trust no one	x	x	x	x	x	x	x	x
1	Believe you are not easily angered	x	x	x	x	x	x	x	x
1	Grouped occupational grade	✔	✔	✔	✔	✔	✔	✔	✔
1	Isolation score	✔	✔	✔	✔	✔	✔	✔	✔
1	Depression case from GHQ	x	x	x	x	x	x	x	x
3	Marital status	✔	✔	✔	x	✔	✔	✔	✔
3	Memory score	✔	✔	x	x	x	✔	x	x

Table A8.1 continued

		Significant predictor of

phase variable collected		missing Mill Hill test score at phase 9		missing last recorded grouped occupational grade at phase 7		missing educational qualifications		missing childhood material deprivation	
		men	**women**	**men**	**women**	**men**	**women**	**men**	**women**
3	AH4 score	✔	✔	✔	✔	✔	✔	✔	✔
3	Mill Hill test score	✔	✔	✔	✔	✔	✔	✔	✔
3	Verbal fluency - S words	✔	✔	✔	✔	✔	✔	✔	✔
3	Verbal fluency – animals	✔	✔	✔	✔	✔	✔	✔	✔
3	Job involves travel away from home	✔	✔	✔	✔	✔	✔	✔	✔
3	Last recorded occupational grade	✔	✔	✔	✔	✔	✔	✔	✔
4	Ever told had depression	✔	x	✔	x	✔	x	✔	x
4	Ever told had anxiety	✔	x	✔	x	✔	x	✔	x
5	Childhood emotional deprivation	x	x	x	x	x	✔	-	-
5	Ever told high blood pressure	✔	✔	✔	x	-	-	-	-
5	Ever diagnosed with cancer	✔	x	x	✔	-	-	-	-
5	Deprivation score	✔	✔	✔	x	-	-	-	-
5	Memory score	✔	✔	✔	✔	-	-	-	-
5	AH4 score	✔	✔	✔	✔	-	-	-	-
5	Mill Hill test score	✔	✔	✔	✔	-	-	-	-
5	Verbal fluency - S words	✔	✔	✔	✔	-	-	-	-
5	Verbal fluency – animals	✔	✔	✔	✔	-	-	-	-
5	How financially secure do you feel in next 10 years	✔	x	✔	x	-	-	-	-

Table A8.1 continued

phase variable collected		Significant predictor of			
		missing Mill Hill test score at phase 9		missing last recorded grouped occupational grade at phase7	
		men	women	men	women
5	To what extent do you feel you might as well give up because you can't make things better for yourself	✔	x	✔	x
5	Last recorded occupational grade	✔	✔	✔	✔
7	General health	✔	✔	-	-
7	Health stops you from doing what you want to do	✔	✔	-	-
7	CASP score	✔	x	-	-
7	Clinic or home visit	✔	✔	-	-
7	MMSE score	✔	✔	-	-
7	Still at civil service	✔	✔	-	-
7	Marital status	x	x	-	-
7	Memory score	✔	✔	-	-
7	AH4 score	✔	✔	-	-
7	Mill Hill test score	✔	✔	-	-
7	Verbal fluency - S words	✔	✔	-	-
7	Verbal fluency – animals	✔	✔	-	-

Table A8.2: Step 2 of identifying variables for the Whitehall II imputation model - men

phase		Mill Hill test score at phase 9	Childhood material deprivation	Last recorded grouped occupational grade at phase 7	Educational qualifications	Father's occupational SEP in childhood
1	Year of birth	x	✓	✓	✓	✓
1	Age finished full time education	✓	✓	✓	✓	✓
1	Accommodation type	✓	✓	✓	✓	✓
1	State of health in the last year	✓	✓	✓	x	x
1	Any longstanding illnesses?	✓	✓	x	x	x
1	Smoking status	✓	✓	✓	✓	x
1	Usually pressed for time	✓	X	✓	✓	✓
1	Grouped occupational grade	✓	✓	✓	✓	✓
1	Isolation score	x	✓	✓	✓	✓
1	Believe no one cares much about you	✓	✓	✓	x	x
3	Marital status	✓	X	✓	x	x
3	Memory score	✓	✓	✓	✓	✓
3	AH4 score	✓	✓	✓	✓	✓
3	Mill Hill test score	✓	✓	✓	✓	✓
3	Verbal fluency - S words	✓	✓	✓	✓	✓
3	Verbal fluency – animals	✓	✓	✓	✓	✓
3	Job involves travel away from home	✓	✓	✓	✓	x
3	Last recorded occupational grade	✓	✓	✓	✓	✓
4	Ever told had depression	x	✓	✓	x	x

Table A8.2 continued

phase		Mill Hill test score at phase 9	Childhood material deprivation	Last recorded grouped occupational grade at phase 7	Educational qualifications	Father's occupational SEP in childhood
4	Ever told had anxiety	x	✔	✔	x	x
5	Ever told high blood pressure	✔	✔	✔	✔	x
5	Ever diagnosed with cancer	x	X	x	x	x
5	Deprivation score	✔	✔	✔	x	x
5	Memory score	✔	✔	✔	✔	✔
5	AH4 score	✔	✔	✔	✔	✔
5	Mill Hill test score	✔	✔	✔	✔	✔
5	Verbal fluency - S words	✔	✔	✔	✔	✔
5	Verbal fluency – animals	✔	✔	✔	✔	✔
5	To what extent do you feel you might as well give up because you can't make things better for yourself	✔	✔	✔	x	x
5	Childhood emotional deprivation	x	✔	x	✔	✔
5	How financially secure do you feel in next 10 years	✔	✔	✔	✔	x
5	Last recorded occupational grade	✔	✔	✔	✔	✔
7	General health	✔	✔	✔	✔	x
7	Health stops you from doing what you want to do	✔	✔	✔	✔	✔
7	Clinic or home visit	✔	x	x	✔	x
7	MMSE score	✔	✔	✔	✔	x

Table A8.2 continued

phase		Mill Hill test score at phase 9	Childhood material deprivation	Last recorded grouped occupational grade at phase 7	Educational qualifications	Father's occupational SEP in childhood
7	Still at civil service	✔	✔	x	✔	x
7	Memory score	✔	✔	✔	✔	✔
7	AH4 score	✔	✔	✔	✔	✔
7	Mill Hill test score	✔	✔	✔	✔	✔
7	Verbal fluency - S words	✔	✔	✔	✔	✔
7	Verbal fluency – animals	✔	✔	✔	✔	✔
7	CASP score	✔	✔	✔	x	x

Table A8.3: Step 2 of identifying variables for the Whitehall II imputation model - women

phase		Mill Hill test score at phase 9	Childhood material deprivation	Last recorded grouped occupational grade at phase 7	Educational qualifications	Father's occupational SEP in childhood
1	Year of birth	✔	✔	✔	✔	✔
1	Age finished full time education	✔	✔	✔	✔	✔
1	Accommodation type	✔	✔	✔	✔	✔
1	Age mother finished full time education	✔	✔	✔	✔	Model did not converge
1	State of health in the last year	✔	✔	✔	✔	✔
1	Any longstanding illnesses?	x	✔	x	x	x
1	Smoking status	✔	x	✔	✔	✔
1	Usually pressed for time	✔	x	✔	✔	✔
1	Grouped occupational grade	✔	✔	✔	✔	✔
1	Isolation score	x	✔	✔	✔	✔
3	Marital status	✔	✔	✔	✔	✔
3	Memory score	✔	x	✔	✔	✔
3	AH4 score	✔	✔	✔	✔	✔
3	Mill Hill test score	✔	✔	✔	✔	✔
3	Verbal fluency - S words	✔	✔	✔	✔	✔
3	Verbal fluency – animals	✔	✔	✔	✔	✔
3	Job involves travel away from home	✔	✔	✔	✔	✔

Table A8.3 continued

phase		Mill Hill test score at phase 9	Childhood material deprivation	Last recorded grouped occupational grade at phase 7	Educational qualifications	Father's occupational SEP in childhood
3	Last recorded occupational grade	✔	✔	✔	✔	✔
5	Ever told high blood pressure	✔	✔	✔	✔	✔
5	Ever diagnosed with cancer	x	x	x	x	x
5	Deprivation score	✔	✔	✔	x	x
5	Memory score	✔	✔	✔	✔	✔
5	AH4 score	✔	✔	✔	✔	✔
5	Mill Hill test score	✔	✔	✔	✔	✔
5	Verbal fluency - S words	✔	✔	✔	✔	✔
5	Verbal fluency – animals	✔	✔	✔	✔	✔
5	Childhood emotional deprivation	✔	✔	x	x	x
5	Last recorded occupational grade	✔	✔	✔	✔	✔
7	General health	✔	✔	✔	✔	✔
7	Health stops you from doing what you want to do	✔	✔	✔	x	x
7	Clinic or home visit	✔	x	✔	✔	x

Table A8.3 continued

phase		Mill Hill test score at phase 9	Childhood material deprivation	Last recorded grouped occupational grade at phase 7	Educational qualifications	Father's occupational SEP in childhood
7	MMSE score	✔	✔	✔	✔	✔
7	Still at civil service	x	✔	✔	✔	x
7	Memory score	✔	✔	✔	✔	✔
7	AH4 score	✔	✔	✔	✔	✔
7	Mill Hill test score	✔	✔	✔	✔	✔
7	Verbal fluency - S words	✔	✔	✔	✔	✔
7	Verbal fluency – animals	✔	✔	✔	✔	✔

Table A8.4: Potential auxiliary variables for the Whitehall II imputation models

Potential auxiliary variable	Hypothesised to improve the fit of:	Men	Women
Job satisfaction (phase 1)	Educational qualifications (phase 5)	No	Yes
Age mother finished full time education (phase 1)	Father's occupation in childhood (phase 1)	Yes	-
Childhood emotional deprivation (phase 5)	Childhood material deprivation (phase 5)	Yes	-
Feel might as well give up (phase 5)	Mill Hill test score (phase 9)	-	Yes
Marital status (phase 9)	Educational qualifications (phase 5)	Yes	Yes
Last recorded occupational grade (phase 9)	Mill Hill test score (phase 9)	Yes	Yes
Memory score (phase 9)	Mill Hill test score (phase 9)	Yes	Yes
AH4 score (phase 9)	Mill Hill test score (phase 9)	Yes	Yes
Verbal fluency - S words (phase 9)	Mill Hill test score (phase 9)	Yes	Yes
Verbal fluency - animals (phase 9)	Mill Hill test score (phase 9)	Yes	Yes
MMSE score (phase 9)	Mill Hill test score (phase 9)	Yes	Yes
General health (phase 9)	Mill Hill test score (phase 9)	Yes	Yes
Difficulty paying bills (phase 9)	Mill Hill test score (phase 9)	Yes	Yes

Appendix 9: Whitehall II MID model development, Aim 1

Table A9.1: Whitehall II MID model development for men, using childhood material deprivation, with outcome Mill Hill test

Men (N=4,357)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e)	p-value
Childhood material deprivation	0.55 (0.10)	<0.001	0.29 (0.09)	0.001	0.22 (0.09)	0.013
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			2.12 (0.39)	<0.001	1.56 (0.38)	<0.001
Education: O-Level/A-Level/National diploma/Certificate			3.33 (0.28)	<0.001	2.51 (0.28)	<0.001
Education: University degree			5.21 (0.27)	<0.001	3.77 (0.28)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					2.90 (0.27)	<0.001
Last occ. grade – Unified Grades 1-6					4.58 (0.28)	<0.001
Age (phase 9)	0.01 (0.01)	0.511	0.04 (0.01)	<0.001	0.02 (0.01)	0.016
No. of times taken cognitive tests	0.20 (0.07)	0.003	0.13 (0.06)	0.033	0.07 (0.06)	0.210
Constant	23.84 (0.71)	<0.001	18.55 (0.74)	<0.001	17.24 (0.73)	<0.001

Table A9.2: Whitehall II MID model development for men, using father's occupational SEP, with outcome Mill Hill test

Men (N=4,357)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e)	p-value
Baseline: Father's occupational SEP in childhood – manual						
Father's occ. SEP - non-manual	0.95 (0.12)	<0.001	0.48 (0.12)	<0.001	0.39 (0.11)	0.001
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			2.08 (0.39)	<0.001	1.53 (0.38)	<0.001
Education: O-Level/A-Level/National diploma/Certificate			3.30 (0.27)	<0.001	2.48 (0.27)	<0.001
Education: University degree			5.14 (0.27)	<0.001	3.71 (0.27)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					2.90 (0.27)	<0.001
Last occ. grade – Unified Grades 1-6					4.57 (0.28)	<0.001
Age (phase 9)	0.00 (0.01)	0.979	0.03 (0.01)	<0.001	0.02 (0.01)	0.032
No. of times taken cognitive tests	0.20 (0.07)	0.002	0.14 (0.06)	0.030	0.08 (0.06)	0.198
Constant	24.62 (0.67)	<0.001	19.13 (0.71)	<0.001	17.62 (0.70)	<0.001

Table A9.3: Whitehall II MID model development for women, using childhood material deprivation, with outcome Mill Hill test

Women (N=1,687)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e)	p-value
Childhood material deprivation	0.70 (0.21)	0.001	0.12 (0.18)	0.517	-0.15 (0.16)	0.355
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			0.77 (0.50)	0.124	0.27 (0.45)	0.553
Education: O-Level/A-Level/National diploma/Certificate			3.95 (0.34)	<0.001	2.09 (0.33)	<0.001
Education: University degree			6.84 (0.39)	<0.001	3.25 (0.40)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					4.15 (0.26)	<0.001
Last occ. grade – Unified Grades 1-6					6.87 (0.36)	<0.001
Age (phase 9)	-0.18 (0.02)	<0.001	-0.03 (0.02)	0.156	-0.01 (0.02)	0.602
No. of times taken cognitive tests	0.56 (0.15)	<0.001	0.64 (0.14)	<0.001	0.49 (0.12)	<0.001
Constant	32.44 (1.63)	<0.001	19.66 (1.64)	<0.001	17.47 (1.46)	<0.001

Table A9.4: Whitehall II MID model development for women, using father's occupational SEP, with outcome Mill Hill test

Women (N=1,687)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e)	p-value
Baseline: Father's occupational SEP in childhood – manual						
Father's occ. SEP - non-manual	2.15 (0.29)	<0.001	0.50 (0.30)	0.092	0.09 (0.26)	0.723
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			0.70 (0.50)	0.165	0.23 (0.45)	0.600
Education: O-Level/A-Level/National diploma/Certificate			3.81 (0.37)	<0.001	2.05 (0.34)	<0.001
Education: University degree			6.60 (0.42)	<0.001	3.17 (0.42)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					4.14 (0.26)	<0.001
Last occ. grade – Unified Grades 1-6					6.83 (0.36)	<0.001
Age (phase 9)	-0.17 (0.02)	<0.001	-0.03 (0.02)	0.108	-0.01 (0.02)	0.679
No. of times taken cognitive tests	0.58 (0.15)	<0.001	0.64 (0.14)	<0.001	0.49 (0.12)	<0.001
Constant	32.12 (1.51)	<0.001	19.96 (1.56)	<0.001	17.08 (1.40)	<0.001

Appendix 10: NSHD Multiple Imputation (MI) and Multiple Imputation, then Deletion (MID) model development, Aim 1

Table A10.1: NSHD MI model development for men, with father's occupational SEP, educational qualifications and own occupational SEP at age 43, and outcome NART

MEN MI (N=2,815)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP – manual										
Childhood SEP - non-manual	5.94 (0.55)	<0.001	2.14 (0.54)	<0.001	1.28 (0.53)	0.018	0.10 (0.48)	0.834	0.15 (0.47)	0.758
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			7.49 (0.76)	<0.001	5.98 (0.79)	<0.001	3.43 (0.75)	<0.001	3.20 (0.74)	<0.001
Education - GCE 'A'-Level			9.85 (0.66)	<0.001	7.64 (0.72)	<0.001	4.06 (0.67)	<0.001	4.00 (0.68)	<0.001
Education – Degree			15.23 (0.69)	<0.001	11.95 (0.81)	<0.001	6.31 (0.87)	<0.001	6.64 (0.86)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					5.56 (0.61)	<0.001	3.70 (0.56)	<0.001	3.57 (0.55)	<0.001
Standardised age 8 cognitive score							5.45 (0.30)	<0.001	5.38 (0.29)	<0.001
Standardised age 8 cognitive score squared									-0.83 (0.28)	0.005
Constant	30.13 (0.39)	<0.001	26.05 (0.49)	<0.001	24.35 (0.47)	<0.001	28.12 (0.46)	<0.001	28.81 (0.52)	<0.001

Table A10.2: NSHD MID model development for men, with father's occupational SEP, educational qualifications and own occupational SEP at age 43, and outcome NART

MEN MID (N=1,370)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP – manual										
Childhood SEP - non-manual	5.79 (0.61)	<0.001	2.33 (0.56)	<0.001	1.58 (0.55)	0.004	0.42 (0.51)	0.413	0.46 (0.51)	0.366
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			6.30 (0.78)	<0.001	4.84 (0.83)	<0.001	2.86 (0.75)	<0.001	2.69 (0.75)	<0.001
Education - GCE 'A'-Level			8.24 (0.72)	<0.001	6.40 (0.77)	<0.001	3.49 (0.71)	<0.001	3.46 (0.71)	<0.001
Education – Degree			13.94 (0.67)	<0.001	11.08 (0.78)	<0.001	6.24 (0.82)	<0.001	6.42 (0.81)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					4.86 (0.70)	<0.001	3.41 (0.62)	<0.001	3.33 (0.63)	<0.001
Standardised age 8 cognitive score							4.89 (0.33)	<0.001	4.88 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.57 (0.28)	0.042
Constant	31.73 (0.38)	<0.001	27.58 (0.49)	<0.001	26.07 (0.53)	<0.001	29.05 (0.50)	<0.001	29.52 (0.56)	<0.001

Table A10.3: NSHD MI model development for men, with father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

MEN MI (N=2,815)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	5.94 (0.55)	<0.001	2.39 (0.58)	<0.001	1.71 (0.54)	0.002	0.35 (0.50)	0.486	0.40 (0.50)	0.423
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		0.004		0.003
Age 26 own SEP – IIIM			2.71 (0.89)	0.003	2.29 (0.86)	0.009	0.78 (0.71)	0.275	0.64 (0.71)	0.370
Age 26 own SEP – IIINM			8.81 (0.96)	<0.001	5.67 (1.02)	<0.001	2.27 (0.87)	0.010	2.17 (0.88)	0.015
Age 26 own SEP - I and II			11.20 (0.93)	<0.001	7.65 (1.03)	<0.001	2.78 (0.86)	0.002	2.84 (0.86)	0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					6.32 (0.68)	<0.001	4.13 (0.57)	<0.001	3.97 (0.57)	<0.001
Standardised age 8 cognitive score							6.03 (0.30)	<0.001	5.96 (0.30)	<0.001
Standardised age 8 cognitive score squared									-0.83 (0.30)	0.007
Constant	30.13 (0.39)	<0.001	25.69 (0.75)	<0.001	24.00 (0.72)	<0.001	28.66 (0.64)	<0.001	29.39 (0.69)	<0.001

Table A10.4: NSHD MID model development for men, with father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

MEN MID (N=1,370)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP – manual										
Childhood SEP - non-manual	5.79 (0.61)	<0.001	2.52 (0.60)	<0.001	1.98 (0.58)	0.001	0.59 (0.53)	0.267	0.63 (0.53)	0.232
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		0.003		0.002
Age 26 own SEP – IIIM			0.85 (1.08)	0.432	1.19 (1.02)	0.242	0.13 (0.85)	0.878	-0.06 (0.86)	0.945
Age 26 own SEP – IIINM			7.57 (1.13)	<0.001	5.06 (1.19)	<0.001	2.54 (1.03)	0.014	2.42 (1.04)	0.019
Age 26 own SEP - I and II			9.17 (1.04)	<0.001	6.65 (1.07)	<0.001	2.80 (0.95)	0.003	2.77 (0.95)	0.003
Baseline: Age 43 own SEP – manual										
Age 43 own SEP - non-manual					5.32 (0.79)	<0.001	3.51 (0.70)	<0.001	3.40 (0.71)	<0.001
Standardised age 8 cognitive score							5.40 (0.32)	<0.001	5.39 (0.32)	<0.001
Standardised age 8 cognitive score squared									-0.57 (0.28)	0.044
Constant	31.73 (0.38)	<0.001	28.42 (0.93)	<0.001	26.44 (0.92)	<0.001	29.92 (0.76)	<0.001	30.49 (0.82)	<0.001

Table A10.5: NSHD MI model development for men, with childhood household amenities, educational qualifications and own occupational SEP at age 43, and outcome NART

MEN MI (N=2,815)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.343		0.795		0.791		0.8048
Childhood - lacking 2 amenities	0.09 (1.35)	0.945	-0.07 (1.18)	0.953	-0.14 (1.13)	0.902	0.45 (0.99)	0.654	0.30 (0.99)	0.765
Childhood - lacking 1 amenity	1.79 (1.43)	0.211	0.45 (1.31)	0.731	0.04 (1.27)	0.972	0.02 (1.12)	0.988	-0.26 (1.13)	0.821
Childhood - lacking no amenities	4.01 (1.33)	0.003	1.04 (1.22)	0.394	0.45 (1.18)	0.701	0.62 (1.07)	0.562	0.36 (1.07)	0.740
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			7.65 (0.78)	<0.001	6.06 (0.82)	<0.001	3.41 (0.76)	<0.001	3.21 (0.75)	<0.001
Education - GCE 'A'-Level			10.09 (0.66)	<0.001	7.74 (0.73)	<0.001	4.03 (0.69)	<0.001	3.99 (0.69)	<0.001
Education – Degree			15.86 (0.72)	<0.001	12.27 (0.85)	<0.001	6.23 (0.91)	<0.001	6.61 (0.90)	<0.001
Baseline: Age 43 own SEP – manual										
Age 43 own SEP - non-manual					5.67 (0.59)	<0.001	3.69 (0.54)	<0.001	3.57 (0.54)	<0.001
Standardised age 8 cognitive score							5.46 (0.29)	<0.001	5.40 (0.29)	<0.001
Standardised age 8 cognitive score squared									-0.83 (0.28)	0.005
Constant	29.56 (1.24)	<0.001	25.94 (1.09)	<0.001	24.39 (1.04)	<0.001	27.76 (0.93)	<0.001	28.65 (1.03)	<0.001

Table A10.6: NSHD MID model development for men, with childhood household amenities, educational qualifications and own occupational SEP at age 43, and outcome NART

MEN MID (N=1,370)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.361		0.716		0.651		0.732
Childhood - lacking 2 amenities	1.85 (1.65)	0.044	0.98 (1.44)	0.496	0.90 (1.39)	0.516	1.07 (1.08)	0.321	0.88 (1.08)	0.413
Childhood - lacking 1 amenity	3.55 (1.76)	0.001	1.32 (1.56)	0.396	0.93 (1.51)	0.535	0.61 (1.20)	0.609	0.36 (1.21)	0.766
Childhood - lacking no amenities	5.51 (1.61)	<0.001	1.93 (1.42)	0.174	1.38 (1.38)	0.316	1.19 (1.07)	0.267	0.95 (1.08)	0.378
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			6.49 (0.78)	<0.001	4.95 (0.83)	<0.001	2.87 (0.75)	<0.001	2.73 (0.75)	<0.001
Education - GCE 'A'-Level			8.47 (0.72)	<0.001	6.51 (0.77)	<0.001	3.47 (0.72)	<0.001	3.46 (0.72)	<0.001
Education – Degree			14.62 (0.66)	<0.001	11.48 (0.77)	<0.001	6.24 (0.82)	<0.001	6.46 (0.81)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					5.02 (0.70)	<0.001	3.43 (0.62)	<0.001	3.36 (0.62)	<0.001
Standardised age 8 cognitive score							4.93 (0.33)	<0.001	4.92 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.55 (0.28)	0.053
Constant	29.58 (1.54)	<0.001	26.61 (1.38)	<0.001	25.23 (1.33)	<0.001	28.18 (1.03)	<0.001	28.83 (1.10)	<0.001

Table A10.7: NSHD MI model development for men, with childhood household amenities own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

MEN MI (N=2,815)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.013		0.140		0.419		0.485
Childhood - lacking 2 amenities	0.09 (1.35)	0.945	0.21 (1.27)	0.869	0.06 (1.19)	0.961	0.58 (1.02)	0.570	0.44 (1.02)	0.668
Childhood - lacking 1 amenity	1.79 (1.43)	0.211	1.06 (0.39)	0.447	0.54 (1.32)	0.684	0.26 (1.14)	0.820	-0.02 (1.15)	0.989
Childhood - lacking no amenities	4.01 (1.33)	0.003	2.25 (1.34)	0.095	1.46 (1.27)	0.254	1.16 (1.09)	0.289	0.91 (1.09)	0.408
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		0.003		0.002
Age 26 own SEP – IIIM			2.81 (0.89)	0.002	2.35 (0.86)	0.007	0.78 (0.71)	0.275	0.65 (0.71)	0.367
Age 26 own SEP – IIINM			9.11 (0.94)	<0.001	5.89 (1.00)	<0.001	2.30 (0.86)	0.008	2.21 (0.86)	0.011
Age 26 own SEP - I and II			11.69 (0.89)	<0.001	8.01 (0.99)	<0.001	2.79 (0.84)	0.001	2.87 (0.84)	0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					6.34 (0.66)	<0.001	4.09 (0.56)	<0.001	3.95 (0.55)	<0.001
Standardised age 8 cognitive score							6.04 (0.30)	<0.001	5.98 (0.30)	<0.001
Standardised age 8 cognitive score squared									-0.82 (0.30)	0.010
Constant	29.56 (1.24)	<0.001	24.87 (1.31)	<0.001	23.54 (1.25)	<0.001	28.04 (1.06)	<0.001	28.96 (1.16)	<0.001

Table A10.8: NSHD MID model development for men, with childhood household amenities own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

MEN MID (N=1,370)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.020		0.117		0.293		0.376
Childhood - lacking 2 amenities	1.85 (1.65)	0.044	1.83 (1.53)	0.231	1.58 (1.47)	0.284	1.39 (1.09)	0.205	1.20 (1.09)	0.269
Childhood - lacking 1 amenity	3.55 (1.76)	0.001	2.45 (1.65)	0.137	1.88 (1.59)	0.237	1.06 (1.21)	0.383	0.80 (1.22)	0.511
Childhood - lacking no amenities	5.51 (1.61)	<0.001	3.51 (1.48)	0.018	2.75 (1.44)	0.056	1.87 (1.07)	0.081	1.64 (1.08)	0.128
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		0.002		0.001
Age 26 own SEP – IIIM			0.94 (1.09)	0.391	1.27 (1.03)	0.217	0.17 (0.86)	0.840	-0.00 (0.86)	0.996
Age 26 own SEP – IIINM			7.88 (1.14)	<0.001	5.32 (1.19)	<0.001	2.61 (1.02)	0.011	2.51 (1.03)	0.015
Age 26 own SEP - I and II			9.71 (1.04)	<0.001	7.10 (1.06)	<0.001	2.90 (0.95)	0.002	2.89 (0.94)	0.002
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					5.33 (0.80)	<0.001	3.47 (0.70)	<0.001	3.38 (0.70)	<0.001
Standardised age 8 cognitive score							5.42 (0.32)	<0.001	5.41 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.54 (0.29)	0.064
Constant	29.58 (1.54)	<0.001	26.26 (1.73)	<0.001	24.69 (1.65)	<0.001	28.58 (1.25)	<0.001	29.32 (1.33)	<0.001

Table A10.9: NSHD MI model development for women, with father's occupational SEP, educational qualifications and own occupational SEP at age 43, and outcome NART

WOMEN MI (N=2547)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	6.94 (0.53)	<0.001	2.45 (0.48)	<0.001	2.33 (0.48)	<0.001	1.01 (0.43)	0.022	1.07 (0.43)	0.015
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			9.28 (0.52)	<0.001	8.42 (0.56)	<0.001	4.46 (0.54)	<0.001	4.36 (0.54)	<0.001
Education - GCE 'A'-Level			13.18 (0.59)	<0.001	11.97 (0.65)	<0.001	6.54 (0.65)	<0.001	6.64 (0.65)	<0.001
Education – Degree			16.97 (0.91)	<0.001	15.68 (0.95)	<0.001	6.34 (1.09)	<0.001	7.26 (1.12)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.70 (0.59)	<0.001	1.05 (0.50)	0.036	0.93 (0.50)	0.067
Standardised age 8 cognitive score							6.18 (0.29)	<0.001	6.13 (0.29)	<0.001
Standardised age 8 cognitive score squared									-0.68 (0.25)	0.010
Constant	29.69 (0.31)	<0.001	25.69 (0.34)	<0.001	24.45 (0.46)	<0.001	28.65 (0.42)	<0.001	29.24 (0.50)	<0.001

Table A10.10: NSHD MID model development for women, with father's occupational SEP, educational qualifications and own occupational SEP at age 43, and outcome NART

WOMEN MID (N=1455)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	6.25 (0.58)	<0.001	2.18 (0.53)	<0.001	2.09 (0.52)	<0.001	1.07 (0.44)	0.015	1.11 (0.44)	0.011
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			8.45 (0.57)	<0.001	7.71 (0.59)	<0.001	4.40 (0.53)	<0.001	4.32 (0.53)	<0.001
Education - GCE 'A'-Level			12.50 (0.62)	<0.001	11.42 (0.67)	<0.001	6.75 (0.60)	<0.001	6.80 (0.61)	<0.001
Education – Degree			16.48 (0.90)	<0.001	15.27 (0.94)	<0.001	7.14 (1.02)	<0.001	7.76 (1.08)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.50 (0.62)	<0.001	1.05 (0.53)	0.051	0.96 (0.54)	0.077
Standardised age 8 cognitive score							5.62 (0.30)	<0.001	5.62 (0.29)	<0.001
Standardised age 8 cognitive score squared									-0.48 (0.27)	0.078
Constant	30.93 (0.37)	<0.001	26.60 (0.42)	<0.001	25.37 (0.53)	<0.001	28.83 (0.46)	<0.001	29.24 (0.53)	<0.001

Table A10.11: NSHD MI model development for women, with father's occupational SEP, educational qualifications and head of household occupational SEP at age 43, and outcome NART

WOMEN MI (N=2547)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	6.94 (0.53)	<0.001	2.45 (0.48)	<0.001	2.28 (0.46)	<0.001	0.94 (0.42)	0.028	1.00 (0.42)	0.019
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			9.28 (0.52)	<0.001	8.78 (0.55)	<0.001	4.44 (0.53)	<0.001	4.29 (0.52)	<0.001
Education - GCE 'A'-Level			13.18 (0.59)	<0.001	12.29 (0.70)	<0.001	6.38 (0.64)	<0.001	6.43 (0.65)	<0.001
Education – Degree			16.97 (0.91)	<0.001	15.94 (0.99)	<0.001	6.09 (1.07)	<0.001	7.02 (1.10)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					2.07 (0.65)	0.002	1.39 (0.49)	0.006	1.40 (0.51)	0.008
Standardised age 8 cognitive score							6.22 (0.28)	<0.001	6.16 (0.29)	<0.001
Standardised age 8 cognitive score squared									-0.72 (0.25)	0.006
Constant	29.69 (0.31)	<0.001	25.69 (0.34)	<0.001	24.90 (0.42)	<0.001	28.62 (0.40)	<0.001	29.17 (0.47)	<0.001

Table A10.12: NSHD MID model development for women, with father's occupational SEP, educational qualifications and head of household occupational SEP at age 43, and outcome NART

WOMEN MID (N=1455)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	6.25 (0.58)	<0.001	2.18 (0.53)	<0.001	2.06 (0.52)	<0.001	1.02 (0.44)	0.021	1.06 (0.44)	0.015
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			8.45 (0.57)	<0.001	8.02 (0.59)	<0.001	4.41 (0.53)	<0.001	4.29 (0.53)	<0.001
Education - GCE 'A'-Level			12.50 (0.62)	<0.001	11.69 (0.66)	<0.001	6.64 (0.59)	<0.001	6.66 (0.60)	<0.001
Education – Degree			16.48 (0.90)	<0.001	15.43 (0.95)	<0.001	6.90 (1.02)	<0.001	7.53 (1.07)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					1.93 (0.56)	0.001	1.27 (0.48)	0.009	1.27 (0.48)	0.009
Standardised age 8 cognitive score							5.66 (0.29)	<0.001	5.65 (0.29)	<0.001
Standardised age 8 cognitive score squared									-0.51 (0.26)	0.058
Constant	30.93 (0.37)	<0.001	26.60 (0.42)	<0.001	25.81 (0.49)	<0.001	28.84 (0.42)	<0.001	29.22 (0.47)	<0.001

Table A10.13: NSHD MI model development for women, with father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

WOMEN MI (N=2547)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	6.94 (0.53)	<0.001	4.37 (0.50)	<0.001	4.07 (0.49)	<0.001	1.63 (0.44)	<0.001	1.70 (0.44)	<0.001
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP – IIIM			2.36 (0.97)	0.016	1.87 (0.98)	0.058	1.04 (0.73)	0.156	1.01 (0.73)	0.168
Age 26 own SEP – IIINM			7.36 (0.64)	<0.001	5.80 (0.71)	<0.001	3.00 (0.56)	<0.001	2.87 (0.56)	<0.001
Age 26 own SEP - I and II			12.35 (0.77)	<0.001	10.37 (0.87)	<0.001	4.88 (0.70)	<0.001	4.94 (0.72)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					3.69 (0.63)	<0.001	1.26 (0.51)	0.015	1.18 (0.52)	0.025
Standardised age 8 cognitive score							6.90 (0.26)	<0.001	6.88 (0.27)	<0.001
Standardised age 8 cognitive score squared									-0.58 (0.25)	0.024
Constant	29.69 (0.31)	<0.001	24.55 (0.55)	<0.001	23.36 (0.59)	<0.001	28.53 (0.50)	<0.001	29.08 (0.56)	<0.001

Table A10.14: NSHD MID model development for women, with father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

WOMEN MID (N=1455)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	6.25 (0.58)	<0.001	4.11 (0.57)	<0.001	3.78 (0.55)	<0.001	1.76 (0.45)	<0.001	1.82 (0.45)	<0.001
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP – IIIM			1.10 (1.21)	0.364	0.67 (1.20)	0.575	0.83 (0.89)	0.348	0.85 (0.87)	0.332
Age 26 own SEP – IIINM			5.65 (0.78)	<0.001	4.07 (0.82)	<0.001	2.29 (0.67)	0.001	2.22 (0.67)	0.001
Age 26 own SEP - I and II			11.14 (0.89)	<0.001	9.11 (0.95)	<0.001	4.57 (0.78)	<0.001	4.65 (0.78)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					4.01 (0.68)	<0.001	1.56 (0.57)	0.006	1.49 (0.57)	0.009
Standardised age 8 cognitive score							6.49 (0.29)	<0.001	6.50 (0.29)	<0.001
Standardised age 8 cognitive score squared									-0.45 (0.26)	0.086
Constant	30.93 (0.37)	<0.001	26.43 (0.68)	<0.001	24.98 (0.71)	<0.001	29.01 (0.59)	<0.001	29.40 (0.65)	<0.001

Table A10.15: NSHD MI model development for women, with father's occupational SEP, own occupational SEP at age 26 and head of household occupational SEP at age 43, and outcome NART

WOMEN MI (N=2547)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	6.94 (0.53)	<0.001	4.37 (0.50)	<0.001	3.96 (0.48)	<0.001	1.50 (0.42)	0.001	1.58 (0.42)	<0.001
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP – IIIM			2.36 (0.97)	0.016	1.94 (0.97)	0.046	0.98 (0.72)	0.175	0.93 (0.72)	0.198
Age 26 own SEP – IIINM			7.36 (0.64)	<0.001	6.49 (0.71)	<0.001	3.05 (0.54)	<0.001	2.86 (0.53)	<0.001
Age 26 own SEP - I and II			12.35 (0.77)	<0.001	10.91 (0.86)	<0.001	4.75 (0.65)	<0.001	4.76 (0.67)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					3.18 (0.70)	<0.001	1.78 (0.50)	0.001	1.82 (0.51)	0.001
Standardised age 8 cognitive score							6.90 (0.26)	<0.001	6.87 (0.26)	<0.001
Standardised age 8 cognitive score squared									-0.64 (0.24)	0.012
Constant	29.69 (0.31)	<0.001	24.55 (0.55)	<0.001	23.60 (0.58)	<0.001	28.40 (0.52)	<0.001	28.96 (0.56)	<0.001

Table A10.16: NSHD MID model development for women, with father's occupational SEP, own occupational SEP at age 26 and head of household occupational SEP at age 43, and outcome NART

WOMEN MID (N=1455)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP – manual										
Childhood SEP - non-manual	6.25 (0.58)	<0.001	4.11 (0.57)	<0.001	3.73 (0.56)	<0.001	1.68 (0.46)	<0.001	1.73 (0.45)	<0.001
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP – IIIM			1.10 (1.21)	0.364	0.54 (1.19)	0.649	0.69 (0.88)	0.431	0.70 (0.86)	0.418
Age 26 own SEP – IIINM			5.65 (0.78)	<0.001	4.81 (0.78)	<0.001	2.43 (0.64)	<0.001	2.32 (0.64)	<0.001
Age 26 own SEP - I and II			11.14 (0.89)	<0.001	9.61 (0.90)	<0.001	4.50 (0.74)	<0.001	4.55 (0.75)	<0.001
Baseline: Age 43 HoH SEP – manual										
Age 43 HoH SEP - non-manual					3.30 (0.60)	<0.001	1.81 (0.49)	<0.001	1.83 (0.49)	<0.001
Standardised age 8 cognitive score							6.52 (0.29)	<0.001	6.52 (0.28)	<0.001
Standardised age 8 cognitive score squared									-0.50 (0.26)	0.055
Constant	30.93 (0.37)	<0.001	26.43 (0.68)	<0.001	25.35 (0.70)	<0.001	29.00 (0.57)	<0.001	29.39 (0.62)	<0.001

Table A10.17: NSHD MI model development for women, with childhood household amenities, educational qualifications and own occupational SEP at age 43, and outcome NART

WOMEN MI (N=2547)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.322		0.401		0.893		0.939
Childhood - lacking 2 amenities	1.27 (1.46)	0.387	0.90 (1.22)	0.461	0.85 (1.21)	0.485	0.02 (0.98)	0.98	-0.01 (0.99)	0.994
Childhood - lacking 1 amenity	3.94 (1.49)	0.008	1.60 (1.24)	0.197	1.50 (1.24)	0.229	0.37 (0.98)	0.711	0.28 (0.99)	0.780
Childhood - lacking no amenities	4.79 (1.47)	0.001	1.71 (1.22)	0.163	1.56 (1.22)	0.202	0.35 (0.99)	0.725	0.26 (1.00)	0.797
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			9.56 (0.52)	<0.001	8.67 (0.56)	<0.001	4.54 (0.55)	<0.001	4.45 (0.54)	<0.001
Education - GCE 'A'-Level			13.88 (0.57)	<0.001	12.63 (0.63)	<0.001	6.78 (0.64)	<0.001	6.90 (0.64)	<0.001
Education – Degree			18.01 (0.85)	<0.001	16.65 (0.88)	<0.001	6.68 (1.05)	<0.001	7.61 (1.07)	<0.001
Baseline: Age 43 own SEP – manual										
Age 43 own SEP - non-manual					2.75 (0.59)	<0.001	1.06 (0.50)	0.035	0.94 (0.50)	0.064
Standardised age 8 cognitive score							6.24 (0.29)	<0.001	6.20 (0.29)	<0.001
Standardised age 8 cognitive score squared									-0.65 (0.25)	0.014
Constant	28.38 (1.38)	<0.001	24.80 (1.17)	<0.001	23.61 (1.24)	<0.001	28.63 (1.02)	<0.001	29.26 (1.09)	<0.001

Table A10.18: NSHD MID model development for women, with childhood household amenities, educational qualifications and own occupational SEP at age 43, and outcome NART

WOMEN MID (N=1455)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.822		0.869		0.904		0.929
Childhood - lacking 2 amenities	0.47 (1.49)	0.753	0.25 (1.29)	0.847	0.20 (1.27)	0.876	-0.22 (1.03)	0.829	-0.24 (1.02)	0.818
Childhood - lacking 1 amenity	2.64 (1.57)	0.092	0.39 (1.32)	0.770	0.23 (1.31)	0.858	-0.13 (1.06)	0.906	-0.19 (1.06)	0.858
Childhood - lacking no amenities	3.69 (1.48)	0.013	0.74 (1.27)	0.563	0.61 (1.26)	0.630	-0.13 (1.01)	0.898	0.06 (1.01)	0.949
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			8.71 (0.57)	<0.001	7.95 (0.58)	<0.001	4.48 (0.52)	<0.001	4.41 (0.53)	<0.001
Education - GCE 'A'-Level			13.21 (0.59)	<0.001	12.09 (0.64)	<0.001	7.02 (0.59)	<0.001	7.09 (0.59)	<0.001
Education – Degree			17.43 (0.84)	<0.001	16.17 (0.89)	<0.001	7.47 (0.01)	<0.001	8.09 (1.07)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.56 (0.63)	<0.001	1.06 (0.54)	0.050	0.98 (0.55)	0.073
Standardised age 8 cognitive score							5.69 (0.29)	<0.001	5.69 (0.29)	<0.001
Standardised age 8 cognitive score squared									-0.45 (0.27)	0.094
Constant	30.48 (1.40)	<0.001	26.50 (1.25)	<0.001	25.32 (1.25)	<0.001	29.08 (0.99)	<0.001	29.51 (1.03)	<0.001

Table A10.19: NSHD MI model development for women, with childhood household amenities, educational qualifications and head of household occupational SEP at age 43, and outcome NART

WOMEN MI (N=2547)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.322		0.369		0.900		0.949
Childhood - lacking 2 amenities	1.27 (1.46)	0.387	0.90 (1.22)	0.461	0.94 (1.22)	0.443	0.07 (0.99)	0.947	0.03 (1.00)	0.976
Childhood - lacking 1 amenity	3.94 (1.49)	0.008	1.60 (1.24)	0.197	1.59 (1.24)	0.202	0.39 (0.99)	0.696	0.29 (1.00)	0.770
Childhood - lacking no amenities	4.79 (1.47)	0.001	1.71 (1.22)	0.163	1.67 (1.22)	0.174	0.37 (0.99)	0.707	0.27 (1.00)	0.787
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			9.56 (0.52)	<0.001	9.01 (0.55)	<0.001	4.50 (0.54)	<0.001	4.37 (0.53)	<0.001
Education - GCE 'A'-Level			13.88 (0.57)	<0.001	12.89 (0.67)	<0.001	6.58 (0.63)	<0.001	6.67 (0.63)	<0.001
Education – Degree			18.01 (0.85)	<0.001	16.84 (0.94)	<0.001	6.39 (1.03)	<0.001	7.33 (1.06)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					2.19 (0.66)	0.002	1.43 (0.50)	0.005	1.45 (0.51)	0.006
Standardised age 8 cognitive score							6.27 (0.28)	<0.001	6.21 (0.29)	<0.001
Standardised age 8 cognitive score squared									-0.69 (0.25)	0.008
Constant	28.38 (1.38)	<0.001	24.80 (1.17)	<0.001	23.94 (1.22)	<0.001	28.55 (1.04)	<0.001	29.15 (1.11)	<0.001

Table A10.20: NSHD MID model development for women, with childhood household amenities, educational qualifications and head of household occupational SEP at age 43, and outcome NART

WOMEN MID (N=1455)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.822		0.877		0.927		0.952
Childhood - lacking 2 amenities	0.47 (1.49)	0.753	0.25 (1.29)	0.847	0.31 (1.28)	0.806	-0.16 (1.03)	0.874	-0.18 (1.02)	0.862
Childhood - lacking 1 amenity	2.64 (1.57)	0.092	0.39 (1.32)	0.770	0.40 (1.31)	0.763	-0.06 (1.06)	0.954	-0.13 (1.06)	0.899
Childhood - lacking no amenities	3.69 (1.48)	0.013	0.74 (1.27)	0.563	0.71 (1.26)	0.576	0.16 (1.01)	0.874	0.09 (1.01)	0.932
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			8.71 (0.57)	<0.001	8.25 (0.58)	<0.001	4.47 (0.52)	<0.001	4.38 (0.53)	<0.001
Education - GCE 'A'-Level			13.21 (0.59)	<0.001	12.34 (0.63)	<0.001	6.89 (0.58)	<0.001	6.93 (0.58)	<0.001
Education – Degree			17.43 (0.84)	<0.001	16.29 (0.89)	<0.001	7.20 (1.01)	<0.001	7.83 (1.06)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					2.02 (0.57)	<0.001	1.30 (0.49)	0.008	1.31 (0.49)	0.007
Standardised age 8 cognitive score							5.72 (0.29)	<0.001	5.71 (0.28)	<0.001
Standardised age 8 cognitive score squared									-0.49 (0.26)	0.069
Constant	30.48 (1.40)	<0.001	26.50 (1.25)	<0.001	25.63 (1.24)	<0.001	29.03 (0.98)	<0.001	29.45 (1.01)	<0.001

Table A10.21: NSHD MI model development for women, with childhood household amenities own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

WOMEN MI (N=2547)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.024		0.044		0.644		0.695
Childhood - lacking 2 amenities	1.27 (1.46)	0.387	1.21 (1.34)	0.366	1.10 (1.33)	0.409	0.04 (1.02)	0.972	-0.00 (1.03)	0.999
Childhood - lacking 1 amenity	3.94 (1.49)	0.008	2.31 (1.34)	0.085	2.14 (1.35)	0.115	0.51 (1.03)	0.623	0.43 (1.04)	0.679
Childhood - lacking no amenities	4.79 (1.47)	0.001	2.81 (1.32)	0.035	2.56 (1.33)	0.056	0.62 (1.03)	0.548	0.55 (1.04)	0.598
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP – IIIM			2.31 (0.98)	0.019	1.80 (0.99)	0.072	1.01 (0.74)	0.175	0.98 (0.73)	0.185
Age 26 own SEP – IIINM			7.62 (0.65)	<0.001	5.95 (0.71)	<0.001	3.01 (0.57)	<0.001	2.89 (0.56)	<0.001
Age 26 own SEP - I and II			13.46 (0.75)	<0.001	11.27 (0.85)	<0.001	5.13 (0.71)	<0.001	5.21 (0.72)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					3.94 (0.63)	<0.001	1.31 (0.52)	0.012	1.24 (0.53)	0.020
Standardised age 8 cognitive score							7.03 (0.26)	<0.001	7.02 (0.26)	<0.001
Standardised age 8 cognitive score squared									-0.54 (0.26)	0.042
Constant	28.38 (1.38)	<0.001	23.38 (1.32)	<0.001	22.22 (1.38)	<0.001	28.53 (1.07)	<0.001	29.10 (1.15)	<0.001

Table A10.22: NSHD MID model development for women, with childhood household amenities own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

WOMEN MID (N=1455)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.087		0.148		0.491		0.527
Childhood - lacking 2 amenities	0.47 (1.49)	0.753	0.65 (1.34)	0.626	0.50 (1.33)	0.710	-0.13 (1.03)	0.898	-0.15 (1.02)	0.880
Childhood - lacking 1 amenity	2.64 (1.57)	0.092	1.31 (1.41)	0.351	1.00 (1.40)	0.475	0.20 (1.08)	0.850	0.14 (1.07)	0.896
Childhood - lacking no amenities	3.69 (1.48)	0.013	2.16 (1.34)	0.106	1.84 (1.33)	0.166	0.65 (1.02)	0.528	0.59 91.02)	0.563
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP – IIIM			1.01 (1.22)	0.408	0.57 (1.22)	0.637	0.78 (0.88)	0.378	0.80 (0.87)	0.361
Age 26 own SEP – IIINM			5.85 (0.80)	<0.001	4.16 (0.84)	<0.001	2.27 (0.68)	0.001	2.21 (0.68)	0.001
Age 26 own SEP - I and II			12.10 (0.89)	<0.001	9.89 (0.97)	<0.001	4.78 (0.79)	<0.001	4.87 (0.79)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					4.26 (0.70)	<0.001	1.62 (0.57)	0.005	1.56 (0.58)	0.007
Standardised age 8 cognitive score							6.65 (0.28)	<0.001	6.66 (0.28)	<0.001
Standardised age 8 cognitive score squared									-0.39 (0.26)	0.130
Constant	30.48 (1.40)	<0.001	25.92 (1.37)	<0.001	24.54 (1.37)	<0.001	29.18 (1.03)	<0.001	29.56 (1.08)	<0.001

Table A10.23: NSHD MI model development for women, with childhood household amenities own occupational SEP at age 26 and head of household occupational SEP at age 43, and outcome NART

WOMEN MI (N=2547)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.024		0.041		0.682		0.741
Childhood - lacking 2 amenities	1.27 (1.46)	0.387	1.21 (1.34)	0.366	1.25 (1.33)	0.351	0.10 (1.03)	0.926	0.05 (1.04)	0.959
Childhood - lacking 1 amenity	3.94 (1.49)	0.008	2.31 (1.34)	0.085	2.25 (1.34)	0.094	0.53 (1.03)	0.609	0.44 (1.04)	0.672
Childhood - lacking no amenities	4.79 (1.47)	0.001	2.81 (1.32)	0.035	2.67 (1.32)	0.045	0.64 (1.04)	0.540	0.55 (1.04)	0.597
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP - IIIM			2.31 (0.98)	0.019	1.86 (0.98)	0.059	0.94 (0.73)	0.198	0.89 (0.72)	0.219
Age 26 own SEP - IIINM			7.62 (0.65)	<0.001	6.62 (0.71)	<0.001	3.04 (0.54)	<0.001	2.88 (0.53)	<0.001
Age 26 own SEP - I and II			13.46 (0.75)	<0.001	11.74 (0.85)	<0.001	4.96 (0.65)	<0.001	5.00 (0.66)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					3.52 (0.71)	<0.001	1.89 (0.51)	<0.001	1.93 (0.51)	<0.001
Standardised age 8 cognitive score							7.02 (0.26)	<0.001	7.00 (0.26)	<0.001
Standardised age 8 cognitive score squared									-0.59 (0.25)	0.021
Constant	28.38 (1.38)	<0.001	23.38 (1.32)	<0.001	22.30 (1.37)	<0.001	28.33 (1.12)	<0.001	28.92 (1.18)	<0.001

Table A10.24: NSHD MID model development for women, with childhood household amenities own occupational SEP at age 26 and head of household occupational SEP at age 43, and outcome NART

WOMEN MID (N=1455)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		<0.001		0.087		0.169		0.559		0.606
Childhood - lacking 2 amenities	0.47 (1.49)	0.753	0.65 (1.34)	0.626	0.73 (1.34)	0.586	-0.03 (1.03)	0.974	-0.06 (1.03)	0.952
Childhood - lacking 1 amenity	2.64 (1.57)	0.092	1.31 (1.41)	0.351	1.27 (1.40)	0.366	0.30 (1.08)	0.785	0.22 (1.08)	0.552
Childhood - lacking no amenities	3.69 (1.48)	0.013	2.16 (1.34)	0.106	2.00 (1.34)	0.134	0.68 (1.03)	0.509	0.61 (1.03)	0.455
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP - IIIM			1.01 (1.22)	0.408	0.42 (1.21)	0.726	0.63 (0.88)	0.470	0.64 (0.86)	0.455
Age 26 own SEP - IIINM			5.85 (0.80)	<0.001	4.92 (0.80)	<0.001	2.41 (0.65)	<0.001	2.32 (0.65)	<0.001
Age 26 own SEP - I and II			12.10 (0.89)	<0.001	10.36 (0.91)	<0.001	4.68 (0.75)	<0.001	4.75 (0.76)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					3.59 (0.62)	<0.001	1.90 (0.50)	<0.001	1.92 (0.50)	<0.001
Standardised age 8 cognitive score							6.66 (0.28)	<0.001	6.67 (0.28)	<0.001
Standardised age 8 cognitive score squared									-0.45 (0.26)	0.083
Constant	30.48 (1.40)	<0.001	25.92 (1.37)	<0.001	24.71 (1.38)	<0.001	29.07 (1.03)	<0.001	29.47 (1.07)	<0.001

Appendix 11: NSHD Heckman selection model development – stricter conditions, Aim 1

Table A11.1: NSHD Heckman selection model development for men, with father's occupational SEP, educational qualifications and own occupational SEP at age 43, and outcome NART

MEN (N=1,086)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	2.75 (0.57)	<0.001	1.55 (0.60)	0.010	1.22 (0.61)	0.045	0.40 (0.60)	0.507	0.45 (0.60)	0.456
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			3.04 (0.89)	0.001	2.64 (0.95)	0.005	1.25 (0.88)	0.157	1.08 (0.88)	0.220
Education - GCE 'A'-Level			5.11 (0.84)	<0.001	4.49 (0.88)	<0.001	3.05 (0.86)	<0.001	3.03 (0.86)	<0.001
Education - Degree			8.68 (0.93)	<0.001	7.70 (1.00)	<0.001	4.78 (1.02)	<0.001	4.97 (0.99)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					1.99 (0.73)	0.006	1.80 (0.70)	0.010	1.78 (0.70)	0.011
Standardised age 8 cognitive score							3.77 (0.36)	<0.001	3.58 (0.34)	<0.001
Standardised age 8 cognitive score squared									-0.42 (0.22)	0.054
Constant	37.42 (0.51)	<0.001	34.24 (0.74)	<0.001	33.49 (0.77)	<0.001	34.69 (0.71)	<0.001	35.09 (0.74)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	182.89	<0.001	106.56	<0.001	100.03	<0.001	51.00	<0.001	49.73	<0.001

Table A11.2: NSHD Heckman selection model development for men, with father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

MEN (N=1,086)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	2.75 (0.57)	<0.001	1.61 (0.62)	0.009	1.44 (0.61)	0.018	0.56 (0.61)	0.360	0.59 (0.60)	0.328
Baseline: Age 26 own SEP - IV and V				<0.001		0.005		0.317		0.223
Age 26 own SEP - IIIM			-0.27 (1.00)	0.784	0.28 (0.98)	0.774	-0.70 (0.90)	0.433	-0.93 (0.90)	0.302
Age 26 own SEP - IIINM			2.90 (1.16)	0.012	2.09 (1.14)	0.067	0.30 (1.06)	0.779	0.20 (1.06)	0.853
Age 26 own SEP - I and II			3.75 (1.07)	<0.001	2.95 (1.06)	0.006	0.77 (1.00)	0.437	0.72 (0.99)	0.464
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.56 (0.74)	0.001	2.31 (0.72)	0.001	2.23 (0.72)	0.002
Standardised age 8 cognitive score							4.14 (0.36)	<0.001	3.97 (0.34)	<0.001
Standardised age 8 cognitive score squared									-0.40 (0.22)	0.066
Constant	37.42 (0.51)	<0.001	36.11 (1.00)	<0.001	34.72 (1.06)	<0.001	36.21 (0.92)	<0.001	36.73 (0.96)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	182.89	<0.001	134.04	<0.001	130.34	<0.001	58.27	<0.001	55.91	<0.001

Table A11.3: NSHD Heckman selection model development for men, with childhood household amenities, educational qualifications and own occupational SEP at age 43, and outcome NART

MEN (N=1,086)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.092		0.668		0.814		0.890		0.963
Childhood - lacking 2 amenities	1.57 (1.45)	0.277	0.84 (1.63)	0.604	1.04 (1.55)	0.504	1.10 (1.47)	0.453	0.64 (1.49)	0.666
Childhood - lacking 1 amenity	2.24 (1.44)	0.119	1.59 (1.65)	0.337	1.45 (1.60)	0.364	1.09 (1.54)	0.481	0.60 (1.56)	0.701
Childhood - lacking no amenities	2.87 (1.40)	0.040	1.40 (1.61)	0.384	1.31 (1.54)	0.397	0.90 (1.44)	0.532	0.41 (1.46)	0.780
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			3.17 (0.89)	<0.001	2.72 (0.94)	0.004	1.30 (0.88)	0.143	1.13 (0.88)	0.199
Education - GCE 'A'-Level			5.23 (0.84)	<0.001	4.54 (0.88)	<0.001	3.06 (0.86)	<0.001	3.06 (0.86)	<0.001
Education – Degree			9.08 (0.95)	<0.001	7.97 (1.01)	<0.001	4.90 (1.02)	<0.001	5.12 (1.00)	<0.001
Baseline: Age 43 own SEP – manual										
Age 43 own SEP - non-manual					2.12 (0.73)	0.003	1.87 (0.70)	0.007	1.86 (0.69)	0.007
Standardised age 8 cognitive score							3.79 (0.36)	<0.001	3.62 (0.34)	<0.001
Standardised age 8 cognitive score squared									-0.40 (0.22)	0.075
Constant	36.00 (1.41)	<0.001	33.38 (1.73)	<0.001	32.51 (1.65)	<0.001	33.78 (1.54)	<0.001	34.61 (1.61)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	215.89	<0.001	105.01	<0.001	100.41	<0.001	51.31	<0.001	50.10	<0.001

Table A11.4: NSHD Heckman selection model development for men, with childhood household amenities, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

MEN (N=1,086)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.092		0.544		0.596		0.719		0.859
Childhood - lacking 2 amenities	1.57 (1.45)	0.277	1.65 (1.64)	0.315	1.73 (1.57)	0.272	1.58 (1.41)	0.261	1.22 (1.42)	0.392
Childhood - lacking 1 amenity	2.24 (1.44)	0.119	2.12 (1.68)	0.206	1.92 (1.61)	0.234	1.48 (1.49)	0.320	1.06 (1.50)	0.479
Childhood - lacking no amenities	2.87 (1.40)	0.040	2.17 (1.61)	0.179	2.09 (1.54)	0.175	1.54 (1.37)	0.260	1.14 (1.37)	0.404
Baseline: Age 26 own SEP - IV and V				<0.001		0.001		0.203		0.139
Age 26 own SEP - IIIM			-0.47 (1.01)	0.641	0.09 (0.99)	0.931	-0.81 (0.91)	0.374	-0.99 (0.91)	0.275
Age 26 own SEP - IIINM			3.08 (1.18)	0.009	2.19 (1.16)	0.059	0.28 (1.06)	0.794	0.21 (1.06)	0.843
Age 26 own SEP - I and II			3.94 (1.08)	<0.001	3.08 (1.08)	0.004	0.83 (1.00)	0.407	0.80 (0.99)	0.416
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.63 (0.75)	<0.001	2.34 (0.71)	0.001	2.28 (0.72)	0.002
Standardised age 8 cognitive score							4.16 (0.36)	<0.001	4.02 (0.34)	<0.001
Standardised age 8 cognitive score squared									-0.35 (0.22)	0.112
Constant	36.00 (1.41)	<0.001	34.66 (1.85)	<0.001	33.25 (1.84)	<0.001	34.91 (1.61)	<0.001	35.72 (1.67)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	215.89	<0.001	138.73	<0.001	135.26	<0.001	58.46	<0.001	56.43	<0.001

Table A11.5: NSHD Heckman selection model development for women, with father's occupational SEP, educational qualifications and own occupational SEP at age 43, and outcome NART

WOMEN (N=1,095)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	3.46 (0.58)	<0.001	1.63 (0.58)	0.005	1.57 (0.57)	0.006	1.02 (0.50)	0.042	1.05 (0.50)	0.035
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			5.12 (0.73)	<0.001	4.71 (0.71)	<0.001	3.21 (0.66)	<0.001	3.18 (0.66)	<0.001
Education - GCE 'A'-Level			8.33 (0.88)	<0.001	7.69 (0.83)	<0.001	5.50 (0.72)	<0.001	5.59 (0.73)	<0.001
Education - Degree			11.94 (1.32)	<0.001	11.21 (1.28)	<0.001	6.01 (1.24)	<0.001	6.71 (1.24)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					1.99 (0.70)	0.004	0.81 (0.64)	0.205	0.75 (0.65)	0.245
Standardised age 8 cognitive score							4.25 (0.35)	<0.001	4.13 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.40 (0.20)	0.042
Constant	35.32 (0.50)	<0.001	31.87 (0.73)	<0.001	30.68 (0.86)	<0.001	32.49 (0.70)	<0.001	32.83 (0.73)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	210.22	<0.001	75.45	<0.001	95.58	<0.001	38.02	<0.001	35.93	<0.001

Table A11.6: NSHD Heckman selection model development for women, with father's occupational SEP, educational qualifications and head of household occupational SEP at age 43, and outcome NART

WOMEN (N=1,095)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	3.46 (0.58)	<0.001	1.63 (0.58)	0.005	1.67 (0.59)	0.004	1.06 (0.51)	0.039	1.09 (0.51)	0.033
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			5.12 (0.73)	<0.001	4.85 (0.71)	<0.001	3.15 (0.66)	<0.001	3.10 (0.66)	<0.001
Education - GCE 'A'-Level			8.33 (0.88)	<0.001	7.69 (0.87)	<0.001	5.35 (0.71)	<0.001	5.41 (0.71)	<0.001
Education - Degree			11.94 (1.32)	<0.001	11.02 (1.33)	<0.001	5.66 (1.25)	<0.001	6.38 (1.24)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					1.70 (0.61)	0.005	1.24 (0.57)	0.030	1.26 (0.57)	0.027
Standardised age 8 cognitive score							4.28 (0.35)	<0.001	4.14 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.43 (0.19)	0.026
Constant	35.32 (0.50)	<0.001	31.87 (0.73)	<0.001	31.08 (0.81)	<0.001	32.36 (0.68)	<0.001	32.68 (0.70)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	210.22	<0.001	75.45	<0.001	72.63	<0.001	34.27	<0.001	32.56	<0.001

Table A11.7: NSHD Heckman selection model development for women, with father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

WOMEN (N=1,095)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	3.46 (0.58)	<0.001	2.60 (0.58)	<0.001	2.48 (0.56)	<0.001	1.65 (0.55)	0.002	1.68 (0.54)	0.002
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP - IIIM			0.26 (1.08)	0.812	0.16 (1.11)	0.884	0.75 (1.06)	0.480	0.69 (1.04)	0.510
Age 26 own SEP - IIINM			2.73 (0.81)	0.001	1.87 (0.86)	0.031	1.61 (0.84)	0.054	1.58 (0.83)	0.057
Age 26 own SEP - I and II			6.79 (1.00)	<0.001	5.77 (1.01)	<0.001	3.64 (0.96)	<0.001	3.78 (0.95)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.43 (0.75)	0.001	1.20 (0.67)	0.075	1.13 (0.68)	0.094
Standardised age 8 cognitive score							4.68 (0.36)	<0.001	4.57 (0.34)	<0.001
Standardised age 8 cognitive score squared									-0.38 (0.19)	0.047
Constant	35.32 (0.50)	<0.001	32.69 (0.88)	<0.001	31.54 (1.04)	<0.001	32.82 (0.87)	<0.001	33.18 (0.89)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	210.22	<0.001	98.44	<0.001	122.01	<0.001	44.05	<0.001	42.56	<0.001

Table A11.8: NSHD Heckman selection model development for women, with father's occupational SEP, own occupational SEP at age 26 and head of household occupational SEP at age 43, and outcome NART

WOMEN (N=1,095)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	3.46 (0.58)	<0.001	2.60 (0.58)	<0.001	2.58 (0.59)	<0.001	1.67 (0.56)	0.003	1.70 (0.55)	0.002
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP - IIIM			0.26 (1.08)	0.812	0.25 (1.14)	0.829	0.68 (1.07)	0.529	0.60 (1.06)	0.573
Age 26 own SEP - IIINM			2.73 (0.81)	0.001	2.31 (0.82)	0.005	1.63 (0.80)	0.043	1.55 (0.80)	0.052
Age 26 own SEP - I and II			6.79 (1.00)	<0.001	6.00 (0.98)	<0.001	3.51 (0.92)	<0.001	3.62 (0.91)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					1.95 (0.67)	0.004	1.64 (0.58)	0.004	1.67 (0.58)	0.004
Standardised age 8 cognitive score							4.73 (0.35)	<0.001	4.59 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.42 (0.19)	0.022
Constant	35.32 (0.50)	<0.001	32.69 (0.88)	<0.001	31.85 (1.03)	<0.001	32.69 (0.87)	<0.001	33.05 (0.89)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	210.22	<0.001	98.44	<0.001	97.01	<0.001	39.02	<0.001	37.30	<0.001

Table A11.9: NSHD Heckman selection model development for women, with childhood household amenities, educational qualifications and own occupational SEP at age 43, and outcome NART

WOMEN (N=1,095)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.136		0.661		0.595		0.661		0.643
Childhood - lacking 2 amenities	1.94 (1.11)	0.080	0.83 (1.28)	0.513	0.88 (1.24)	0.475	0.68 (1.26)	0.591	0.55 (1.25)	0.659
Childhood - lacking 1 amenity	1.86 (1.14)	0.103	-0.05 (1.35)	0.968	-0.09 (1.30)	0.944	-0.14 (1.29)	0.911	-0.30 (1.28)	0.816
Childhood - lacking no amenities	2.40 (1.04)	0.020	0.24 (1.27)	0.850	0.21 (1.23)	0.864	0.04 (1.23)	0.972	-0.10 (1.23)	0.937
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			5.33 (0.71)	<0.001	4.88 (0.70)	<0.001	3.34 (0.65)	<0.001	3.32 (0.65)	<0.001
Education - GCE 'A'-Level			8.83 (0.89)	<0.001	8.14 (0.86)	<0.001	5.89 (0.71)	<0.001	6.00 (0.71)	<0.001
Education - Degree			12.76 (1.32)	<0.001	11.96 (1.30)	<0.001	6.52 (1.24)	<0.001	7.22 (1.24)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.04 (0.68)	0.003	0.83 (0.64)	0.196	0.77 (0.64)	0.232
Standardised age 8 cognitive score							4.27 (0.35)	<0.001	4.15 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.40 (0.20)	0.045
Constant	34.27 (1.06)	<0.001	31.77 (1.42)	<0.001	30.58 (1.40)	<0.001	32.44 (1.24)	<0.001	32.91 (1.24)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	254.71	<0.001	69.68	<0.001	90.14	<0.001	45.54	<0.001	43.05	<0.001

Table A11.10: NSHD Heckman selection model development for women, with childhood household amenities, educational qualifications and head of household occupational SEP at age 43, and outcome NART

WOMEN (N=1,095)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.136		0.661		0.557		0.672		0.652
Childhood - lacking 2 amenities	1.94 (1.11)	0.080	0.83 (1.28)	0.513	1.01 (1.24)	0.415	0.75 (1.26)	0.551	0.62 (1.25)	0.620
Childhood - lacking 1 amenity	1.86 (1.14)	0.103	-0.05 (1.35)	0.968	0.03 (1.31)	0.982	-0.01 (1.28)	0.991	-0.19 (1.28)	0.884
Childhood - lacking no amenities	2.40 (1.04)	0.020	0.24 (1.27)	0.850	0.35 (1.24)	0.779	0.10 (1.24)	0.936	-0.05 (1.23)	0.966
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			5.33 (0.71)	<0.001	5.05 (0.69)	<0.001	3.29 (0.65)	<0.001	3.25 (0.65)	<0.001
Education - GCE 'A'-Level			8.83 (0.89)	<0.001	8.19 (0.89)	<0.001	5.76 (0.70)	<0.001	5.84 (0.70)	<0.001
Education – Degree			12.76 (1.32)	<0.001	11.86 (1.32)	<0.001	6.21 (1.25)	<0.001	6.93 (1.24)	<0.001
Baseline: Age 43 HoH SEP – manual										
Age 43 HoH SEP - non-manual					1.70 (0.60)	0.004	1.22 (0.57)	0.031	1.24 (0.56)	0.028
Standardised age 8 cognitive score							4.31 (0.34)	<0.001	4.18 (0.32)	<0.001
Standardised age 8 cognitive score squared									-0.43 (0.19)	0.029
Constant	34.27 (1.06)	<0.001	31.77 (1.42)	<0.001	30.87 (1.42)	<0.001	32.26 (1.27)	<0.001	32.72 (1.27)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	254.71	<0.001	69.68	<0.001	66.79	<0.001	41.03	<0.001	39.03	<0.001

Table A11.11: NSHD Heckman selection model development for women, with childhood household amenities, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

WOMEN (N=1,095)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.136		0.174		0.207		0.611		0.621
Childhood - lacking 2 amenities	1.94 (1.11)	0.080	2.16 (1.05)	0.040	2.00 (1.06)	0.059	1.10 (1.20)	0.359	0.98 (1.19)	0.412
Childhood - lacking 1 amenity	1.86 (1.14)	0.103	1.26 (1.11)	0.255	1.02 (1.10)	0.355	0.26 (1.24)	0.833	0.11 (1.23)	0.926
Childhood - lacking no amenities	2.40 (1.04)	0.020	1.81 (1.00)	0.069	1.68 (1.01)	0.096	0.78 (1.19)	0.511	0.65 (1.18)	0.580
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP – IIIM			0.30 (1.05)	0.772	0.13 (1.07)	0.901	0.76 (1.05)	0.471	0.69 (1.04)	0.503
Age 26 own SEP – IIINM			3.08 (0.78)	<0.001	2.17 (0.83)	0.009	1.79 (0.83)	0.031	1.77 (0.83)	0.032
Age 26 own SEP - I and II			7.41 (0.98)	<0.001	6.29 (1.01)	<0.001	4.13 (0.94)	<0.001	4.27 (0.94)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.48 (0.72)	0.001	1.22 (0.68)	0.072	1.16 (0.68)	0.088
Standardised age 8 cognitive score							4.76 (0.35)	<0.001	4.66 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.36 (0.19)	0.061
Constant	34.27 (1.06)	<0.001	31.37 (1.24)	<0.001	30.39 (1.32)	<0.001	32.33 (1.27)	<0.001	32.79 (1.29)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	254.71	<0.001	95.23	<0.001	94.20	<0.001	54.84	<0.001	52.78	<0.001

Table A11.12: NSHD Heckman selection model development for women, with childhood household amenities, own occupational SEP at age 26 and head of household occupational SEP at age 43, and outcome NART

WOMEN (N=1,095)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.136		0.174		0.228		0.655		0.668
Childhood - lacking 2 amenities	1.94 (1.11)	0.080	2.16 (1.05)	0.040	2.26 (1.14)	0.059	1.18 (1.21)	0.330	1.04 (1.21)	0.387
Childhood - lacking 1 amenity	1.86 (1.14)	0.103	1.26 (1.11)	0.255	1.22 (1.20)	0.310	0.42 (1.25)	0.734	0.25 (1.25)	0.841
Childhood - lacking no amenities	2.40 (1.04)	0.020	1.81 (1.00)	0.069	1.76 (1.10)	0.109	0.81 (1.20)	0.500	0.66 (1.19)	0.579
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP – IIIM			0.30 (1.05)	0.772	0.29 (1.10)	0.793	0.70 (1.07)	0.510	0.62 (1.05)	0.553
Age 26 own SEP – IIINM			3.08 (0.78)	<0.001	2.65 (0.78)	0.001	1.83 (0.80)	0.023	1.77 (0.80)	0.026
Age 26 own SEP - I and II			7.41 (0.98)	<0.001	6.60 (0.97)	<0.001	4.02 (0.90)	<0.001	4.13 (0.890	<0.001
Baseline: Age 43 HoH SEP – manual										
Age 43 HoH SEP - non-manual					1.97 (0.65)	0.002	1.64 (0.58)	0.005	1.67 (0.58)	0.004
Standardised age 8 cognitive score							4.82 (0.34)	<0.001	4.69 (0.32)	<0.001
Standardised age 8 cognitive score squared									-0.40 (0.19)	0.031
Constant	34.27 (1.06)	<0.001	31.37 (1.24)	<0.001	30.58 (1.41)	<0.001	32.12 (1.31)	<0.001	32.61 (1.33)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	254.71	<0.001	95.23	<0.001	80.42	<0.001	48.73	<0.001	46.22	<0.001

Appendix 12: Whitehall II Heckman selection model development – stricter conditions, Aim 1

Table A12.1: Whitehall II Heckman selection model development for men, with childhood material deprivation, educational qualifications and occupational grade, and outcome Mill Hill test

Men (N=4,824)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e)	p-value
Childhood material deprivation	0.33 (0.10)	0.001	0.16 (0.10)	0.105	0.16 (0.10)	0.107
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			1.92 (0.38)	<0.001	1.69 (0.39)	<0.001
Education: O-Level/A-Level/National diploma/Certificate			2.37 (0.27)	<0.001	2.06 (0.27)	<0.001
Education: University degree			3.89 (0.28)	<0.001	3.22 (0.29)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					1.65 (0.36)	<0.001
Last occ. grade - Unified Grades 1-6					2.80 (0.38)	<0.001
Age (phase 9)	0.07 (0.02)	<0.001	0.09 (0.01)	<0.001	0.07 (0.01)	<0.001
No. of times taken cognitive tests	0.28 (0.08)	0.001	0.19 (0.08)	0.013	0.16 (0.08)	0.049
Constant	23.49 (1.07)	<0.001	19.77 (1.06)	<0.001	18.83 (1.06)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	256.56	<0.001	209.56	<0.001	89.11	<0.001

Table A12.2: Whitehall II Heckman selection model development for men, with father's occupational SEP, educational qualifications and occupational grade, and outcome Mill Hill test

Men (N=4,824)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e.)	p-value
Baseline: Father's occupational SEP in childhood - manual						
Father's occ. SEP - non-manual	0.86 (0.13)	<0.001	0.52 (0.13)	<0.001	0.46 (0.13)	<0.001
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			1.88 (0.38)	<0.001	1.65 (0.39)	<0.001
Education: O-Level/A-Level/National diploma/Certificate			2.30 (0.26)	<0.001	2.00 (0.27)	<0.001
Education: University degree			3.75 (0.28)	<0.001	3.12 (0.29)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					1.62 (0.36)	<0.001
Last occ. grade - Unified Grades 1-6					2.74 (0.37)	<0.001
Age (phase 9)	0.07 (0.01)	<0.001	0.08 (0.01)	<0.001	0.07 (0.01)	<0.001
No. of times taken cognitive tests	0.28 (0.08)	0.001	0.19 (0.08)	0.013	0.16 (0.08)	0.046
Constant	23.62 (1.03)	<0.001	19.90 (1.04)	<0.001	19.00 (1.03)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	267.31	<0.001	217.11	<0.001	94.76	<0.001

Table A12.3: Whitehall II Heckman selection model development for women, with childhood material deprivation, educational qualifications and occupational grade, and outcome Mill Hill test

Women (N=1,099)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e)	p-value
Childhood material deprivation	0.41 (0.20)	0.042	0.17 (0.20)	0.390	-0.07 (0.20)	0.717
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			0.44 (0.54)	0.415	0.13 (0.53)	0.811
Education: O-Level/A-Level/National diploma/Certificate			2.13 (0.36)	<0.001	1.37 (0.36)	<0.001
Education: University degree			4.36 (0.45)	<0.001	2.88 (0.48)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					1.76 (0.32)	<0.001
Last occ. grade - Unified Grades 1-6					3.25 (0.46)	<0.001
Age (phase 9)	0.06 (0.03)	0.075	0.13 (0.03)	<0.001	0.13 (0.03)	<0.001
No. of times taken cognitive tests	-0.45 (0.24)	0.064	-0.00 (0.23)	0.995	-0.00 (0.22)	0.997
Constant	23.88 (2.35)	<0.001	15.76 (2.35)	<0.001	14.84 (2.23)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	259.53	<0.001	187.95	<0.001	143.32	<0.001

Table A12.4: Whitehall II Heckman selection model development for women, with father's occupational SEP, educational qualifications and occupational grade, and outcome Mill Hill test

Women (N=1,099)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e.)	p-value
Baseline: Father's occupational SEP in childhood - manual						
Father's occ. SEP - non-manual	1.63 (0.25)	<0.001	0.86 (0.27)	0.001	0.63 (0.27)	0.018
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			0.27 (0.53)	0.608	-0.06 (0.52)	0.911
Education: O-Level/A-Level/National diploma/Certificate			1.84 (0.36)	<0.001	1.17 (0.37)	0.001
Education: University degree			3.84 (0.47)	<0.001	2.50 (0.49)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					1.68 (0.31)	<0.001
Last occ. grade - Unified Grades 1-6					3.06 (0.45)	<0.001
Age (phase 9)	0.06 (0.03)	0.062	0.12 (0.03)	<0.001	0.13 (0.03)	<0.001
No. of times taken cognitive tests	-0.33 (0.23)	0.164	0.01 (0.23)	0.953	-0.01 (0.22)	0.946
Constant	23.15 (2.24)	<0.001	16.22 (2.28)	<0.001	14.79 (2.18)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	265.44	<0.001	195.81	<0.001	149.40	<0.001

Appendix 13: NSHD Heckman selection model development – less strict conditions, Aim 1

Table A13.1: NSHD Heckman selection model development for men, with father's occupational SEP, educational qualifications and own occupational SEP at age 43, and outcome NART

MEN (N=1,088)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	2.64 (0.57)	<0.001	1.47 (0.60)	0.014	1.14 (0.61)	0.062	0.39 (0.60)	0.514	0.44 (0.59)	0.462
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			3.17 (0.85)	<0.001	2.74 (0.90)	0.002	1.42 (0.86)	0.100	1.25 (0.86)	0.148
Education - GCE 'A'-Level			4.93 (0.80)	<0.001	4.26 (0.83)	<0.001	2.76 (0.84)	<0.001	2.73 (0.83)	0.001
Education - Degree			8.55 (0.89)	<0.001	7.49 (0.96)	<0.001	4.37 (0.99)	<0.001	4.58 (0.96)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.10 (0.70)	0.003	1.81 (0.69)	0.008	1.79 (0.68)	0.009
Standardised age 8 cognitive score							3.79 (0.35)	<0.001	3.59 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.46 (0.24)	0.059
Constant	37.39 (0.51)	<0.001	34.32 (0.72)	<0.001	33.56 (0.75)	<0.001	34.80 (0.69)	<0.001	35.24 (0.74)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	183.60	<0.001	114.15	<0.001	102.96	<0.001	61.07	<0.001	61.13	<0.001

Table A13.2: NSHD Heckman selection model development for men, with father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

MEN (N=1,088)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	2.64 (0.57)	<0.001	1.52 (0.61)	0.013	1.35 (0.61)	0.026	0.53 (0.61)	0.381	0.56 (0.60)	0.347
Baseline: Age 26 own SEP - IV and V				<0.001		0.005		0.407		0.278
Age 26 own SEP - IIIM			-0.22 (1.01)	0.829	0.38 (0.98)	0.702	-0.48 (0.88)	0.587	-0.73 (0.88)	0.347
Age 26 own SEP - IIINM			3.18 (1.15)	0.006	2.33 (1.13)	0.040	0.69 (1.03)	0.501	0.59 (1.02)	0.562
Age 26 own SEP - I and II			3.79 (1.08)	<0.001	3.00 (1.06)	0.005	0.79 (0.97)	0.412	0.75 (0.95)	0.431
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.61 (0.74)	0.001	2.22 (0.68)	0.001	2.12 (0.69)	0.002
Standardised age 8 cognitive score							4.14 (0.35)	<0.001	3.94 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.45 (0.24)	0.060
Constant	37.39 (0.51)	<0.001	36.02 (1.02)	<0.001	34.60 (1.06)	<0.001	36.13 (0.88)	<0.001	36.71 (0.91)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	183.60	<0.001	143.31	<0.001	138.88	<0.001	58.27	<0.001	68.07	<0.001

Table A13.3: NSHD Heckman selection model development for men, with childhood household amenities, educational qualifications and own occupational SEP at age 43, and outcome NART

MEN (N=1,088)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.100		0.668		0.864		0.949		0.994
Childhood - lacking 2 amenities	1.05 (1.40)	0.453	0.46 (1.60)	0.775	0.65 (1.53)	0.668	0.70 (1.50)	0.642	0.16 (1.54)	0.916
Childhood - lacking 1 amenity	1.84 (1.41)	0.194	1.23 (1.65)	0.456	1.08 (1.59)	0.494	0.93 (1.57)	0.554	0.37 (1.59)	0.814
Childhood - lacking no amenities	2.52 (1.36)	0.064	1.16 (1.58)	0.462	1.05 (1.52)	0.490	0.76 (1.46)	0.603	0.20 (1.48)	0.892
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			3.25 (0.85)	<0.001	2.80 (0.89)	0.002	1.45 (0.86)	0.094	1.27 (0.86)	0.140
Education - GCE 'A'-Level			5.01 (0.79)	<0.001	4.30 (0.82)	<0.001	2.76 (0.83)	0.001	2.75 (0.83)	0.001
Education – Degree			8.87 (0.91)	<0.001	7.72 (0.95)	<0.001	4.45 (1.00)	<0.001	4.70 (0.97)	<0.001
Baseline: Age 43 own SEP – manual										
Age 43 own SEP - non-manual					2.20 (0.70)	0.002	1.84 (0.68)	0.006	1.83 (0.67)	0.006
Standardised age 8 cognitive score							3.81 (0.35)	<0.001	3.62 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.44 (0.24)	0.064
Constant	36.33 (1.37)	<0.001	33.77 (1.73)	<0.001	32.89 (1.64)	<0.001	34.14 (1.58)	<0.001	35.08 (1.65)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	212.02	<0.001	112.10	<0.001	103.19	<0.001	58.52	<0.001	58.08	<0.001

Table A13.4: NSHD Heckman selection model development for men, with childhood household amenities, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

MEN (N=1,088)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.100		0.526		0.659		0.822		0.936
Childhood - lacking 2 amenities	1.05 (1.40)	0.453	1.29 (1.61)	0.421	1.34 (1.56)	0.392	1.06 (1.47)	0.472	0.61 (1.49)	0.680
Childhood - lacking 1 amenity	1.84 (1.41)	0.194	1.94 (1.65)	0.241	1.66 (1.61)	0.302	1.30 (1.54)	0.398	0.81 (1.54)	0.598
Childhood - lacking no amenities	2.52 (1.36)	0.064	2.00 (1.58)	0.205	1.83 (1.54)	0.232	1.31 (1.43)	0.360	0.84 (1.43)	0.556
Baseline: Age 26 own SEP - IV and V				<0.001		0.002		0.290		0.191
Age 26 own SEP - IIIM			-0.38 (1.01)	0.709	0.20 (0.99)	0.838	-0.57 (0.89)	0.522	-0.77 (0.88)	0.380
Age 26 own SEP - IIINM			3.33 (1.16)	0.004	2.43 (1.14)	0.034	0.70 (1.02)	0.493	0.63 (1.01)	0.531
Age 26 own SEP - I and II			3.96 (1.08)	<0.001	3.13 (1.07)	0.003	0.85 (0.97)	0.380	0.83 (0.95)	0.383
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.63 (0.74)	<0.001	2.20 (0.68)	0.001	2.13 (0.68)	0.002
Standardised age 8 cognitive score							4.15 (0.35)	<0.001	3.98 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.42 (0.24)	0.081
Constant	36.33 (1.37)	<0.001	34.76 (1.81)	<0.001	33.41 (1.83)	<0.001	35.13 (1.65)	<0.001	36.09 (1.68)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	212.02	<0.001	144.61	<0.001	141.63	<0.001	67.16	<0.001	65.68	<0.001

Table A13.5: NSHD Heckman selection model development for women, with father's occupational SEP, educational qualifications and own occupational SEP at age 43, and outcome NART

WOMEN (N=1,131)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	3.61 (0.56)	<0.001	1.70 (0.56)	0.003	1.64 (0.55)	0.003	1.01 (0.50)	0.042	1.04 (0.49)	0.035
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			5.05 (0.74)	<0.001	4.65 (0.72)	<0.001	3.13 (0.65)	<0.001	3.11 (0.65)	<0.001
Education - GCE 'A'-Level			8.30 (0.85)	<0.001	7.69 (0.82)	<0.001	5.57 (0.72)	<0.001	5.65 (0.72)	<0.001
Education - Degree			11.85 (1.28)	<0.001	11.16 (1.26)	<0.001	6.05 (1.22)	<0.001	6.63 (1.23)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					1.80 (0.69)	0.009	0.64 (0.64)	0.315	0.58 (0.64)	0.366
Standardised age 8 cognitive score							4.19 (0.34)	<0.001	4.08 (0.32)	<0.001
Standardised age 8 cognitive score squared									-0.32 (0.20)	0.101
Constant	35.29 (0.48)	<0.001	31.91 (0.70)	<0.001	30.87 (0.84)	<0.001	32.68 (0.68)	<0.001	32.95 (0.71)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	239.50	<0.001	93.68	<0.001	109.98	<0.001	45.20	<0.001	42.33	<0.001

Table A13.6: NSHD Heckman selection model development for women, with father's occupational SEP, educational qualifications and head of household occupational SEP at age 43, and outcome NART

WOMEN (N=1,131)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	3.61 (0.56)	<0.001	1.70 (0.56)	0.003	1.75 (0.57)	0.002	1.04 (0.51)	0.039	1.08 (0.50)	0.032
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			5.05 (0.74)	<0.001	4.82 (0.73)	<0.001	3.06 (0.65)	<0.001	3.02 (0.65)	<0.001
Education - GCE 'A'-Level			8.30 (0.85)	<0.001	7.73 (0.84)	<0.001	5.39 (0.70)	<0.001	5.45 (0.70)	<0.001
Education - Degree			11.85 (1.28)	<0.001	11.02 (1.29)	<0.001	5.70 (1.23)	<0.001	6.29 (1.23)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					1.60 (0.61)	0.008	1.16 (0.57)	0.040	1.17 (0.57)	0.038
Standardised age 8 cognitive score							4.21 (0.34)	<0.001	4.08 (0.32)	<0.001
Standardised age 8 cognitive score squared									-0.34 (0.19)	0.078
Constant	35.29 (0.48)	<0.001	31.91 (0.70)	<0.001	31.14 (0.78)	<0.001	32.48 (0.67)	<0.001	32.74 (0.68)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	239.50	<0.001	93.68	<0.001	90.20	<0.001	40.13	<0.001	37.73	<0.001

Table A13.7: NSHD Heckman selection model development for women, with father's occupational SEP, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

WOMEN (N=1,131)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	3.61 (0.56)	<0.001	2.81 (0.57)	<0.001	2.63 (0.54)	<0.001	1.66 (0.54)	0.002	1.69 (0.54)	0.002
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP - IIIM			0.48 (1.09)	0.656	0.31 (1.09)	0.773	0.61 (1.04)	0.559	0.55 (1.04)	0.599
Age 26 own SEP - IIINM			2.83 (0.83)	0.001	1.95 (0.85)	0.022	1.51 (0.82)	0.065	1.47 (0.82)	0.072
Age 26 own SEP - I and II			6.90 (0.98)	<0.001	5.80 (0.99)	<0.001	3.62 (0.94)	<0.001	3.71 (0.94)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.38 (0.73)	0.001	1.05 (0.66)	0.113	1.00 (0.67)	0.134
Standardised age 8 cognitive score							4.59 (0.36)	<0.001	4.49 (0.34)	<0.001
Standardised age 8 cognitive score squared									-0.28 (0.19)	0.147
Constant	35.29 (0.48)	<0.001	32.51 (0.84)	<0.001	31.49 (0.98)	<0.001	33.05 (0.84)	<0.001	33.32 (0.87)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	239.50	<0.001	160.11	<0.001	171.53	<0.001	48.44	<0.001	46.55	<0.001

Table A13.8: NSHD Heckman selection model development for women, with father's occupational SEP, own occupational SEP at age 26 and head of household occupational SEP at age 43, and outcome NART

WOMEN (N=1,131)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood SEP - manual										
Childhood SEP - non-manual	3.61 (0.56)	<0.001	2.81 (0.57)	<0.001	2.78 (0.57)	<0.001	1.67 (0.55)	0.002	1.71 (0.54)	0.002
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP - IIIM			0.48 (1.09)	0.656	0.44 (1.14)	0.697	0.55 (1.06)	0.602	0.47 (1.05)	0.653
Age 26 own SEP - IIINM			2.83 (0.83)	0.001	2.43 (0.85)	0.004	1.50 (0.79)	0.057	1.43 (0.79)	0.070
Age 26 own SEP - I and II			6.90 (0.98)	<0.001	6.13 (0.99)	<0.001	3.46 (0.90)	<0.001	3.54 (0.90)	<0.001
Baseline: Age 43 HoH SEP - manual										
Age 43 HoH SEP - non-manual					1.92 (0.64)	0.003	1.56 (0.57)	0.006	1.58 (0.57)	0.006
Standardised age 8 cognitive score							4.63 (0.35)	<0.001	4.52 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.32 (0.19)	0.095
Constant	35.29 (0.48)	<0.001	32.51 (0.84)	<0.001	31.68 (0.97)	<0.001	32.87 (0.86)	<0.001	33.16 (0.88)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	239.50	<0.001	160.11	<0.001	135.03	<0.001	42.69	<0.001	40.75	<0.001

Table A13.9: NSHD Heckman selection model development for women, with childhood household amenities, educational qualifications and own occupational SEP at age 43, and outcome NART

WOMEN (N=1,131)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.122		0.602		0.509		0.523		0.498
Childhood - lacking 2 amenities	1.88 (1.15)	0.101	0.91 (1.29)	0.479	0.97 (1.25)	0.434	0.75 (1.25)	0.548	0.67 (1.25)	0.591
Childhood - lacking 1 amenity	1.87 (1.18)	0.112	-0.05 (1.35)	0.970	-0.11 (1.30)	0.933	-0.25 (1.29)	0.845	-0.37 (1.28)	0.772
Childhood - lacking no amenities	2.58 (1.11)	0.019	0.29 (1.28)	0.821	0.26 (1.23)	0.836	0.07 (1.23)	0.953	-0.03 (1.23)	0.978
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			5.31 (0.71)	<0.001	4.87 (0.70)	<0.001	3.27 (0.64)	<0.001	3.27 (0.64)	<0.001
Education - GCE 'A'-Level			8.87 (0.84)	<0.001	8.20 (0.83)	<0.001	5.97 (0.70)	<0.001	6.07 (0.71)	<0.001
Education - Degree			12.72 (1.27)	<0.001	11.96 (1.26)	<0.001	6.58 (1.22)	<0.001	7.15 (1.22)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					1.87 (0.67)	0.005	0.68 (0.64)	0.287	0.62 (0.64)	0.331
Standardised age 8 cognitive score							4.22 (0.34)	<0.001	4.11 (0.32)	<0.001
Standardised age 8 cognitive score squared									-0.32 (0.19)	0.103
Constant	34.21 (1.10)	<0.001	31.75 (1.40)	<0.001	30.69 (1.39)	<0.001	32.59 (1.24)	<0.001	32.96 (1.24)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	322.38	<0.001	92.34	<0.001	106.80	<0.001	53.99	<0.001	50.48	<0.001

Table A13.10: NSHD Heckman selection model development for women, with childhood household amenities, educational qualifications and head of household occupational SEP at age 43, and outcome NART

WOMEN (N=1,131)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.122		0.602		0.502		0.537		0.508
Childhood - lacking 2 amenities	1.88 (1.15)	0.101	0.91 (1.29)	0.479	1.09 (1.27)	0.391	0.82 (1.26)	0.513	0.74 (1.25)	0.556
Childhood - lacking 1 amenity	1.87 (1.18)	0.112	-0.05 (1.35)	0.970	0.04 (1.32)	0.977	-0.14 (1.28)	0.913	-0.27 (1.28)	0.831
Childhood - lacking no amenities	2.58 (1.11)	0.019	0.29 (1.28)	0.821	0.41 (1.26)	0.748	0.12 (1.23)	0.921	0.01 (1.23)	0.995
Baseline: no qualifications/proficiency only				<0.001		<0.001		<0.001		<0.001
Education - up to GCE 'O'-Level			5.31 (0.71)	<0.001	5.09 (0.71)	<0.001	3.22 (0.64)	<0.001	3.19 (0.64)	<0.001
Education - GCE 'A'-Level			8.87 (0.84)	<0.001	8.31 (0.84)	<0.001	5.82 (0.69)	<0.001	5.90 (0.69)	<0.001
Education – Degree			12.72 (1.27)	<0.001	11.91 (1.28)	<0.001	6.26 (1.23)	<0.001	6.85 (1.22)	<0.001
Baseline: Age 43 HoH SEP – manual										
Age 43 HoH SEP - non-manual					1.60 (0.60)	0.007	1.15 (0.56)	0.041	1.16 (0.56)	0.039
Standardised age 8 cognitive score							4.24 (0.33)	<0.001	4.12 (0.32)	<0.001
Standardised age 8 cognitive score squared									-0.34 (0.19)	0.079
Constant	34.21 (1.10)	<0.001	31.75 (1.40)	<0.001	30.87 (1.42)	<0.001	32.36 (1.26)	<0.001	32.72 (1.26)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	322.38	<0.001	92.34	<0.001	88.84	<0.001	48.05	<0.001	45.12	<0.001

Table A13.11: NSHD Heckman selection model development for women, with childhood household amenities, own occupational SEP at age 26 and own occupational SEP at age 43, and outcome NART

WOMEN (N=1,131)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.122		0.209		0.213		0.478		0.478
Childhood - lacking 2 amenities	1.88 (1.15)	0.101	2.10 (1.13)	0.064	1.99 (1.11)	0.074	1.15 (1.19)	0.333	1.08 (1.19)	0.363
Childhood - lacking 1 amenity	1.87 (1.18)	0.112	1.27 (1.19)	0.285	1.01 (1.15)	0.381	0.17 (1.23)	0.889	0.07 (1.23)	0.953
Childhood - lacking no amenities	2.58 (1.11)	0.019	1.98 (1.10)	0.071	1.78 (1.07)	0.097	0.85 (1.18)	0.474	0.76 (1.18)	0.519
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP – IIIM			0.54 (1.09)	0.620	0.36 (1.07)	0.734	0.63 (1.04)	0.542	0.57 (1.03)	0.580
Age 26 own SEP – IIINM			3.17 (0.81)	<0.001	2.24 (0.82)	0.006	1.69 (0.81)	0.037	1.66 (0.81)	0.040
Age 26 own SEP - I and II			7.57 (0.96)	<0.001	6.36 (0.97)	<0.001	4.10 (0.92)	<0.001	4.19 (0.92)	<0.001
Baseline: Age 43 own SEP - manual										
Age 43 own SEP - non-manual					2.52 (0.69)	0.001	1.10 (0.67)	0.100	1.06 (0.687)	0.116
Standardised age 8 cognitive score							4.38 (0.34)	<0.001	4.59 (0.33)	<0.001
Standardised age 8 cognitive score squared									-0.26 (0.19)	0.186
Constant	34.21 (1.10)	<0.001	31.17 (1.31)	<0.001	30.25 (1.34)	<0.001	32.51 (1.26)	<0.001	32.85 (1.29)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	322.38	<0.001	170.73	<0.001	175.20	<0.001	60.43	<0.001	57.98	<0.001

Table A13.12: NSHD Heckman selection model development for women, with childhood household amenities, own occupational SEP at age 26 and head of household occupational SEP at age 43, and outcome NART

WOMEN (N=1,131)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
Baseline: Childhood - lacking 3 amenities		0.122		0.209		0.244		0.539		0.538
Childhood - lacking 2 amenities	1.88 (1.15)	0.101	2.10 (1.13)	0.064	2.15 (1.21)	0.075	1.23 (1.21)	0.309	1.14 (1.20)	0.344
Childhood - lacking 1 amenity	1.87 (1.18)	0.112	1.27 (1.19)	0.285	1.28 (1.25)	0.310	0.33 (1.24)	0.794	0.20 (1.24)	0.871
Childhood - lacking no amenities	2.58 (1.11)	0.019	1.98 (1.10)	0.071	1.94 (1.17)	0.098	0.87 (1.19)	0.466	0.77 (1.19)	0.519
Baseline: Age 26 own SEP - IV and V				<0.001		<0.001		<0.001		<0.001
Age 26 own SEP – IIIM			0.54 (1.09)	0.620	0.52 (1.13)	0.648	0.59 (1.06)	0.574	0.52 (1.05)	0.622
Age 26 own SEP – IIINM			3.17 (0.81)	<0.001	2.77 (0.82)	0.001	1.71 (0.78)	0.029	1.65 (0.78)	0.035
Age 26 own SEP - I and II			7.57 (0.96)	<0.001	6.79 (0.98)	<0.001	3.97 (0.88)	<0.001	4.05 (0.88)	<0.001
Baseline: Age 43 HoH SEP – manual										
Age 43 HoH SEP - non-manual					1.97 (0.63)	0.002	1.57 (0.57)	0.006	1.58 (0.57)	0.005
Standardised age 8 cognitive score							4.73 (0.33)	<0.001	4.63 (0.32)	<0.001
Standardised age 8 cognitive score squared									-0.29 (0.19)	0.124
Constant	34.21 (1.10)	<0.001	31.17 (1.31)	<0.001	30.32 (1.44)	<0.001	32.29 (1.30)	<0.001	32.65 (1.32)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	322.38	<0.001	170.73	<0.001	138.07	<0.001	53.57	<0.001	50.91	<0.001

Appendix 14: Whitehall II Heckman selection model development – less strict conditions, Aim 1

Table A14.1: Whitehall II Heckman selection model development for men, with childhood material deprivation, educational qualifications and occupational grade, and outcome Mill Hill test

Men (N=3,337)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e)	p-value
Childhood material deprivation	0.29 (0.10)	0.005	0.13 (0.10)	0.215	0.12 (0.10)	0.231
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			2.08 (0.38)	<0.001	2.02 (0.38)	<0.001
Education: O-Level/A-Level/National diploma/Certificate			2.55 (0.26)	<0.001	2.43 (0.27)	<0.001
Education: University degree			4.00 (0.28)	<0.001	3.66 (0.28)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					0.32 (0.29)	<0.001
Last occ. grade - Unified Grades 1-6					1.43 (0.29)	<0.001
Age (phase 9)	0.07 (0.01)	<0.001	0.09 (0.01)	<0.001	0.08 (0.01)	<0.001
No. of times taken cognitive tests	0.24 (0.08)	0.002	0.18 (0.08)	0.017	0.11 (0.08)	0.154
Constant	21.45 (1.01)	<0.001	17.56 (1.02)	<0.001	17.87 (1.00)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	299.09	<0.001	227.70	<0.001	132.40	<0.001

Table A14.2: Whitehall II Heckman selection model development for men, with father's occupational SEP, educational qualifications and occupational grade, and outcome Mill Hill test

Men (N=3,337)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e.)	p-value
Baseline: Father's occupational SEP in childhood - manual						
Father's occ. SEP - non-manual	0.86 (0.13)	<0.001	0.57 (0.13)	<0.001	0.53 (0.13)	<0.001
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			2.02 (0.37)	<0.001	1.97 (0.38)	<0.001
Education: O-Level/A-Level/National diploma/Certificate			2.47 (0.26)	<0.001	2.37 (0.27)	<0.001
Education: University degree			3.86 (0.27)	<0.001	3.54 (0.28)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					0.29 (0.28)	0.311
Last occ. grade - Unified Grades 1-6					1.37 (0.29)	<0.001
Age (phase 9)	0.07 (0.01)	<0.001	0.09 (0.01)	<0.001	0.08 (0.01)	<0.001
No. of times taken cognitive tests	0.24 (0.08)	0.003	0.18 (0.08)	0.020	0.11 (0.08)	0.162
Constant	21.46 (0.97)	<0.001	17.52 (0.99)	<0.001	17.84 (0.97)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	307.51	<0.001	236.07	<0.001	140.58	<0.001

Table A14.3: Whitehall II Heckman selection model development for women, with childhood material deprivation, educational qualifications and occupational grade, and outcome Mill Hill test

Women (N=1,083)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e)	p-value
Childhood material deprivation	0.41 (0.20)	0.046	0.06 (0.20)	0.754	-0.13 (0.20)	0.525
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			0.71 (0.75)	0.342	0.41 (0.72)	0.573
Education: O-Level/A-Level/National diploma/Certificate			2.66 (0.47)	<0.001	1.92 (0.47)	<0.001
Education: University degree			6.10 (0.54)	<0.001	4.77 (0.56)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					1.78 (0.33)	<0.001
Last occ. grade - Unified Grades 1-6					2.70 (0.44)	<0.001
Age (phase 9)	0.04 (0.03)	0.211	0.13 (0.03)	<0.001	0.13 (0.03)	<0.001
No. of times taken cognitive tests	-0.12 (0.24)	0.609	0.15 (0.22)	0.488	0.23 (0.22)	0.306
Constant	23.58 (2.35)	<0.001	14.03 (2.37)	<0.001	12.95 (2.29)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	242.67	<0.001	175.79	<0.001	140.52	<0.001

Table A14.4: Whitehall II Heckman selection model development for women, with father's occupational SEP, educational qualifications and occupational grade, and outcome Mill Hill test

Women (N=1,083)	Model 1		Model 2		Model 3	
	Coef (s.e.)	p-value	Coef (s.e.)	p-value	Coef (s.e.)	p-value
Baseline: Father's occupational SEP in childhood - manual						
Father's occ. SEP - non-manual	1.52 (0.27)	<0.001	0.78 (0.27)	0.001	0.64 (0.27)	0.017
Baseline: no educational qualifications				<0.001		<0.001
Education: School certificate/matriculation			0.59 (0.75)	0.608	0.28 (0.73)	0.703
Education: O-Level/A-Level/National diploma/Certificate			2.41 (0.48)	<0.001	1.72 (0.48)	0.001
Education: University degree			5.65 (0.56)	<0.001	4.41 (0.58)	<0.001
Baseline: Last recorded occupational grade (phase 7) - Clerical						<0.001
Last occ. grade - Senior and Higher Executive Officers					1.71 (0.32)	<0.001
Last occ. grade - Unified Grades 1-6					2.53 (0.43)	<0.001
Age (phase 9)	0.04 (0.03)	0.248	0.13 (0.03)	<0.001	0.14 (0.03)	<0.001
No. of times taken cognitive tests	0.03 (0.23)	0.898	0.19 (0.22)	0.397	0.23 (0.22)	0.300
Constant	23.14 (2.24)	<0.001	14.09 (2.31)	<0.001	12.62 (2.24)	<0.001
Model fit	Chi sq	p-value	Chi sq	p-value	Chi sq	p-value
Wald test rho=0	239.77	<0.001	182.45	<0.001	146.56	<0.001

Appendix 15: NSHD backwards selection models (life course analyses, method 1), Aim 1

Table A15.1: NSHD backwards selection for men from the saturated model (S_1: father's occupational SEP, S_{2b}: own occupational SEP at age 26, S_3: own occupational SEP at age 43), with outcome NART

Men	Model 1		Model 2		Model 3		Model 4		Model 5	
N=893	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_1	1.27 (2.07)	0.540	1.85 (1.84)	0.313	1.79 (1.50)	0.232	1.77 (1.50)	0.236	2.43 (0.70)	0.001
S_{2b}	3.87 (2.06)	0.061	4.44 (1.82)	0.015	4.47 (1.71)	0.009	4.93 (0.96)	<0.001	5.16 (0.86)	<0.001
S_3	4.26 (1.17)	<0.001	4.44 (1.13)	<0.001	4.42 (1.08)	<0.001	4.59 (0.90)	<0.001	4.58 (0.90)	<0.001
S_1S_{2b}	3.87 (3.17)	0.223	0.99 (1.85)	0.59	0.92 (1.66)	0.582	0.98 (1.66)	0.555		
$S_{2b}S_3$	1.46 (2.34)	0.534	0.67 (1.97)	0.734	0.65 (1.91)	0.732				
S_1S_3	1.26 (2.96)	0.669	-0.15 (2.09)	0.943						
$S_1S_{2b}S_3$	-3.89 (3.89)	0.316								
Constant	28.04 (0.70)	<0.001	27.98 (0.69)	<0.001	27.99 (0.68)	<0.001	27.94 (0.65)	<0.001	27.87 (0.63)	<0.001
R^2	0.2517		0.2509		0.2509		0.2508		0.2504	
BIC	6368.93		6363.08		6356.29		6349.66		6343.33	

Table A15.2: NSHD backwards selection for men from the saturated model (S_{1b}: childhood household amenities, S_2: educational qualifications, S_3: own occupational SEP at age 43), with outcome NART

Men	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
N=893	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_{1b}	1.04 (1.46)	0.478	1.25 (1.32)	0.345	1.42 (0.99)	0.153	1.18 (0.67)	0.080	1.24 (0.67)	0.068		
S_2	3.45 (2.26)	0.128	3.82 (1.78)	0.033	3.74 (1.69)	0.027	3.54 (1.55)	0.022	5.32 (0.75)	<0.001	5.57 (0.73)	<0.001
S_3	4.92 (1.24)	<0.001	5.08 (1.16)	<0.001	5.21 (0.96)	<0.001	5.22 (0.96)	<0.001	6.07 (0.80)	<0.001	6.19 (0.80)	<0.001
$S_{1b}S_2$	0.26 (2.99)	0.931	-0.68 (1.42)	0.632	-0.55 (1.31)	0.675						
S_2S_3	3.27 (2.54)	0.198	2.68 (1.74)	0.122	2.73 (1.70)	0.109	2.66 (1.71)	0.120				
$S_{1b}S_3$	0.85 (1.94)	0.662	0.37 (1.58)	0.813								
$S_{1b}S_2S_3$	-1.39 (3.36)	0.678										
Constant	27.64 (0.83)	<0.001	27.58 (0.81)	<0.001	27.53 (0.76)	<0.001	27.60 (0.72)	<0.001	27.22 (0.67)	<0.001	27.58 (0.63)	<0.001
R^2	0.2580		0.2577		0.2576		0.2575		0.2539		0.2501	
BIC	6361.46		6354.94		6348.23		6341.66		6339.09		6336.89	

Table A15.3: NSHD backwards selection for men from the saturated model (S_{1b}: childhood household amenities, S_{2b}: own occupational SEP at age 26, S_3: own occupational SEP at age 43), with outcome NART

Men	Model 1	
N=893	Coef. (s.e.)	p-value
S_{1b}	0.53 (1.34)	0.689
S_{2b}	2.42 (2.12)	0.255
S_3	3.73 (1.41)	0.008
$S_{1b}S_{2b}$	6.86 (2.77)	0.013
$S_{2b}S_3$	4.03 (2.49)	0.106
$S_{1b}S_3$	2.24 (2.09)	0.284
$S_{1b}S_{2b}S_3$	-8.49 (3.30)	0.010
Constant	27.98 (0.85)	<0.001
R^2	0.2518	
BIC	6368.77	

Table A15.4: NSHD backwards selection for women from the saturated model (S_1: father's occupational SEP, S_{2b}: own occupational SEP at age 26, S_3: own occupational SEP at age 43), with outcome NART

Women	Model 1		Model 2		Model 3		Model 4		Model 5	
N=955	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_1	3.00 (2.60)	0.249	4.51 (1.90)	0.018	4.23 (1.85)	0.022	3.24 (1.47)	0.028	4.81 (0.63)	<0.001
S_{2b}	3.62 (1.40)	0.010	4.01 (1.30)	0.002	4.76 (0.92)	<0.001	4.47 (0.79)	<0.001	4.39 (0.80)	<0.001
S_3	3.22 (1.46)	0.028	3.58 (1.37)	0.009	4.28 (0.91)	<0.001	4.41 (0.90)	<0.001	4.82 (0.77)	<0.001
S_1S_{2b}	0.75 (3.16)	0.811	-1.70 (1.69)	0.315	-1.62 (1.66)	0.332				
$S_{2b}S_3$	1.96 (1.85)	0.290	1.29 (1.63)	0.428						
S_1S_3	4.93 (3.13)	0.116	2.15 (1.65)	0.193	2.42 (1.65)	0.141	1.99 (1.62)	0.219		
$S_1S_{2b}S_3$	-3.89 (3.69)	0.292								
Constant	25.38 (0.99)	<0.001	25.24 (0.96)	<0.001	24.97 (0.85)	<0.001	25.08 (0.82)	<0.001	24.85 (0.78)	<0.001
R^2	0.2233		0.2223		0.2215		0.2208		0.2194	
BIC	6831.95		6826.27		6820.32		6814.34		6809.18	

TableA15.5: NSHD backwards selection for women from the saturated model (S_1: father's occupational SEP, S_2: educational qualifications, S_{3b}: head of household occupational SEP at age 43), with outcome NART

Women	Model 1		Model 2		Model 3		Model 4	
N=955	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_1	5.22 (1.35)	<0.001	5.45 (1.24)	<0.001	5.33 (1.21)	<0.001	4.43 (0.83)	<0.001
S_2	7.38 (1.72)	<0.001	8.21 (1.38)	<0.001	9.08 (0.85)	<0.001	9.21 (0.84)	<0.001
S_{3b}	4.29 (0.88)	<0.001	4.37 (0.86)	<0.001	4.44 (0.83)	<0.001	4.09 (0.69)	<0.001
S_1S_2	-0.82 (2.43)	0.735	-2.69 (1.20)	0.025	-2.72 (1.20)	0.024	-3.22 (1.17)	0.006
S_2S_{3b}	2.05 (1.96)	0.295	1.05 (1.43)	0.460				
S_1S_{3b}	-1.41 (1.70)	0.408	-1.80 (1.46)	0.217	-1.61 (1.41)	0.252		
$S_1S_2S_{3b}$	-2.29 (2.78)	0.410						
Constant	27.35 (0.68)	<0.001	27.32 (0.67)	<0.001	27.28 (0.65)	<0.001	27.45 (0.60)	<0.001
R^2	0.2752		0.2749		0.2747		0.2736	
BIC	6765.82		6759.34		6752.81		6747.36	

Table A15.6: NSHD backwards selection for women from the saturated model (S_1: father's occupational SEP, S_{2b}: own occupational SEP at age 26, S_{3b}: head of household occupational SEP at age 43), with outcome NART

Women	Model 1		Model 2		Model 3		Model 4		Model 5	
N=955	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_1	5.18 (2.38)	0.029	5.60 (1.78)	0.002	5.62 (1.77)	0.002	5.54 (1.59)	0.001	4.58 (0.64)	<0.001
S_{2b}	5.15 (1.27)	<0.001	5.23 (1.20)	<0.001	5.18 (0.88)	<0.001	5.19 (0.87)	<0.001	5.00 (0.77)	<0.001
S_{3b}	4.59 (1.40)	0.001	4.70 (1.30)	<0.001	4.64 (0.82)	<0.001	4.60 (0.68)	<0.001	4.61 (0.68)	<0.001
S_1S_{2b}	-0.53 (2.75)	0.847	-1.11 (1.76)	0.529	-1.12 (1.76)	0.525	-1.16 (1.74)	0.504		
$S_{2b}S_{3b}$	0.08 (1.73)	0.964	-0.10 (1.53)	0.947						
S_1S_{3b}	0.68 (3.18)	0.831	-0.16 (1.43)	0.911	-0.18 (1.40)	0.900				
$S_1S_{2b}S_{3b}$	-1.06 (3.56)	0.766								
Constant	25.09 (0.97)	<0.001	25.05 (0.94)	<0.001	25.07 (0.82)	<0.001	25.08 (0.79)	<0.001	25.20 (0.74)	<0.001
R^2	0.2273		0.2272		0.2272		0.2272		0.2268	
BIC	6827.02		6820.25		6813.40		6806.55		6800.19	

Table A15.7: NSHD backwards selection for women from the saturated model (S_{1b}: childhood household amenities, S_2: educational qualifications, S_3: own occupational SEP at age 43), with outcome NART

Women	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
N=955	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_{1b}	-0.66 (1.40)	0.639	-0.46 (1.37)	0.737	-0.49 (1.36)	0.722	0.61 (0.64)	0.342				
S_2	2.36 (3.22)	0.463	6.81 (2.24)	0.002	6.23 (1.96)	0.002	5.90 (1.85)	0.001	6.08 (1.87)	0.001	8.88 (0.61)	<0.001
S_3	4.03 (1.07)	<0.001	4.15 (1.05)	<0.001	4.25 (1.02)	<0.001	4.84 (0.83)	<0.001	4.86 (0.82)	<0.001	5.00 (0.79)	<0.001
$S_{1b}S_2$	5.87 (3.85)	0.127	-0.89 (1.23)	0.472								
S_2S_3	7.19 (3.35)	0.032	2.54 (2.15)	0.237	2.65 (2.06)	0.200	3.05 (1.95)	0.117	2.97 (1.98)	0.133		
$S_{1b}S_3$	2.09 (1.68)	0.214	1.78 (1.62)	0.272	1.55 (1.54)	0.314						
$S_{1b}S_2S_3$	-7.15 (4.06)	0.078										
Constant	27.49 (0.86)	<0.001	27.42 (0.85)	<0.001	27.43 (0.85)	<0.001	27.03 (0.72)	<0.001	27.25 (0.68)	<0.001	27.17 (0.66)	<0.001
R^2	0.2585		0.2574		0.2571		0.2558		0.2548		0.2540	
BIC	6787.59		6782.10		6775.65		6770.47		6764.86		6759.05	

Table A15.8: NSHD backwards selection for women from the saturated model (S_{1b}: childhood household amenities, S_{2b}: own occupational SEP at age 26, S_3: own occupational SEP at age 43), with outcome NART

Women	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
N=955	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_{1b}	-0.83 (2.02)	0.682	-0.26 (1.66)	0.876	-0.35 (1.65)	0.833	-0.98 (1.34)	0.464	1.15 (0.67)	0.088		
S_{2b}	4.52 (1.59)	0.004	5.04 (1.40)	<0.001	5.66 (1.02)	<0.001	5.09 (0.81)	<0.001	5.01 (0.81)	<0.001	5.14 (0.81)	<0.001
S_3	2.57 (1.62)	0.112	3.05 (1.42)	0.033	3.62 (1.02)	<0.001	3.87 (1.01)	<0.001	5.11 (0.79)	<0.001	5.15 (0.78)	<0.001
$S_{1b}S_{2b}$	-0.14 (2.63)	0.957	-1.43 (1.66)	0.389	-1.43 (1.66)	0.389						
$S_{2b}S_3$	1.95 (2.07)	0.348	1.07 (1.64)	0.517								
$S_{1b}S_3$	4.78 (2.83)	0.091	3.44 (1.59)	0.031	3.56 (1.58)	0.024	3.00 (1.54)	0.051				
$S_{1b}S_{2b}S_3$	-2.21 (3.39)	0.515										
Constant	25.92 (1.11)	<0.001	25.74 (1.05)	<0.001	25.52 (0.95)	<0.001	25.72 (0.90)	<0.001	24.94 (0.80)	<0.001	25.33 (0.78)	<0.001
R^2	0.1837		0.1832		0.1827		0.1817		0.1768		0.1732	
BIC	6879.33		6873.10		6866.85		6861.08		6860.01		6857.24	

Table A15.9: NSHD backwards selection for men from the saturated model (S_{1b}: childhood household amenities, S_2: educational qualifications, S_{3b}: head of household occupational SEP at age 43), with outcome NART

Women	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
N=955	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_{1b}	0.82 (1.30)	0.53	0.59 (1.23)	0.634	0.58 (1.22)	0.637	0.62 (0.80)	0.439	0.50 (0.66)	0.443		
S_2	10.52 (1.75)	<0.001	8.91 (1.44)	<0.001	9.14 (0.90)	<0.001	9.12 (0.89)	<0.001	8.83 (0.59)	<0.001	8.91 (0.58)	<0.001
S_{3b}	4.43 (1.01)	<0.001	4.26 (0.97)	<0.001	4.28 (0.93)	<0.001	4.32 (0.71)	<0.001	4.32 (0.72)	<0.001	4.35 (0.70)	<0.001
$S_{1b}S_2$	-3.78 (2.51)	0.133	-0.58 (1.19)	0.628	-0.57 (1.18)	0.631	-0.54 (1.16)	0.642				
S_2S_{3b}	-1.71 (2.03)	0.399	0.28 (1.49)	0.849								
$S_{1b}S_{3b}$	-0.38 (1.61)	0.812	0.06 (1.45)	0.965	0.08 (1.45)	0.956						
$S_{1b}S_2S_{3b}$	3.90 (2.84)	0.170										
Constant	27.86 (0.73)	<0.001	27.95 (0.72)	<0.001	27.93 (0.71)	<0.001	27.92 (0.63)	<0.001	27.96 (0.61)	<0.001	28.15 (0.58)	<0.001
R^2	0.2488		0.2479		0.2479		0.2479		0.2478		0.2471	
BIC	6799.99		6794.28		6787.44		6780.59		6773.89		6767.87	

Table A15.10: NSHD backwards selection for women from the saturated model (S_{1b}: childhood household amenities, S_{2b}: own occupational SEP at age 26, S_{3b}: head of household occupational SEP at age 43), with outcome NART

Women	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
N=955	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_{1b}	-1.86 (2.08)	0.37	-0.14 (1.65)	0.931	-0.16 (1.19)	0.893	-0.14 (1.17)	0.907	0.95 (0.67)	0.162		
S_{2b}	4.82 (1.39)	0.001	5.84 (1.26)	<0.001	5.83 (1.16)	<0.001	5.71 (0.78)	<0.001	5.61 (0.78)	<0.001	5.71 (0.78)	<0.001
S_{3b}	2.72 (1.59)	0.086	4.36 (1.39)	0.002	4.37 (1.36)	0.001	4.19 (0.92)	<0.001	4.98 (0.70)	<0.001	5.04 (0.69)	<0.001
$S_{1b}S_{2b}$	2.86 (2.50)	0.253	-0.03 (1.62)	0.985								
$S_{2b}S_{3b}$	2.18 (1.95)	0.263	-0.27 (1.56)	0.863	-0.27 (1.54)	0.859						
$S_{1b}S_{3b}$	6.10 (2.73)	0.026	1.86 (1.42)	0.192	1.85 (1.43)	0.196	1.82 (1.41)	0.197				
$S_{1b}S_{2b}S_{3b}$	-5.99 (3.19)	0.061										
Constant	26.10 (1.05)	<0.001	25.61 (1.00)	<0.001	25.62 (0.95)	<0.001	25.68 (0.83)	<0.001	25.32 (0.77)	<0.001	25.62 (0.75)	<0.001
R^2	0.1940		0.1895		0.1895		0.1895		0.1874		0.1850	
BIC	6867.27		6865.67		6858.81		6851.99		6847.65		6843.56	

Appendix 16: Whitehall II backwards selection models (life course analyses, method 1), Aim 1

Table A16.1: Whitehall II backwards selection for women from the saturated model (S_{1b}: childhood material deprivation, S_2: educational qualifications, S_3: occupational grade at phase 1), with outcome Mill Hill test

Men (N=2,440)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_{1b}	0.36 (0.23)	0.124	0.37 (0.22)	0.095	0.37 (0.22)	0.092	0.21 (0.18)	0.244	0.20 (0.18)	0.259
S_2	0.79 (0.45)	0.075	0.82 (0.33)	0.013	0.79 (0.26)	0.003	0.88 (0.25)	<0.001	0.87 (0.25)	0.001
S_3	2.29 (0.29)	<0.001	2.31 (0.26)	<0.001	2.29 (0.25)	<0.001	2.03 (0.14)	<0.001	2.04 (0.14)	<0.001
$S_{1b}S_2$	0.80 (0.53)	0.133	0.76 (0.32)	0.016	0.76 (0.32)	0.016	0.63 (0.30)	0.035	0.63 (0.30)	0.033
S_2S_3	-0.01 (0.55)	0.986	-0.05 (0.31)	0.868						
$S_{1b}S_3$	-0.37 (0.37)	0.314	-0.39 (0.30)	0.205	-0.39 (0.30)	0.200				
$S_{1b}S_2S_3$	-0.06 (0.66)	0.928								
Age (phase 9)	0.00 (0.01)	0.721	0.00 (0.01)	0.722	0.00 (0.01)	0.719	0.00 (0.01)	0.749		
No. of times taken cog. tests	0.27 (0.09)	0.002	0.27 (0.09)	0.002	0.27 (0.09)	0.002	0.28 (0.09)	0.001	0.28 (0.09)	0.002
Constant	22.61 (0.86)	<0.001	22.61 (0.86)	<0.001	22.61 (0.86)	<0.001	22.74 (0.86)	<0.001	22.99 (0.33)	<0.001
R^2	0.1610		0.1610		0.1609		0.1604		0.1603	
BIC	12818.41		12810.62		12802.85		12796.70		12789.00	

Table A16.2: Whitehall II backwards selection for women from the saturated model (S_{1b}: childhood material deprivation, S_2: educational qualifications, S_3: occupational grade at phase 1), with outcome Mill Hill test

Women (N=826)	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value	Coef. (s.e.)	p-value
S_{1b}	0.63 (0.38)	0.095	0.64 (0.38)	0.092	0.57 (0.37)	0.123	0.71 (0.36)	0.050	0.65 (0.36)	0.072
S_2	4.08 (0.85)	<0.001	4.06 (0.85)	<0.001	3.74 (0.76)	<0.001	3.27 (0.71)	<0.001	3.64 (0.68)	<0.001
S_3	2.10 (1.16)	0.070	2.13 (1.15)	0.066	1.52 (0.92)	0.099	2.75 (0.57)	<0.001	3.42 (0.41)	<0.001
$S_{1b}S_2$	-3.42 (1.03)	0.001	-3.42 (1.03)	0.001	-2.94 (0.87)	0.001	-2.29 (0.78)	0.003	-2.10 (0.77)	0.006
S_2S_3	0.13 (1.66)	0.938	0.13 (1.66)	0.937	1.39 (0.81)	0.087	1.39 (0.82)	0.089		
$S_{1b}S_3$	0.84 (1.32)	0.526	0.81 (1.32)	0.537	1.60 (0.95)	0.093				
$S_{1b}S_2S_3$	1.66 (1.90)	0.385	1.66 (1.90)	0.385						
Age (phase 9)	-0.07 (0.03)	0.009	-0.07 (0.03)	0.007	-0.07 (0.03)	0.006	-0.07 (0.03)	0.009	-0.06 (0.03)	0.014
No. of times taken cog. tests	0.17 (0.22)	0.439								
Constant	26.38 (2.02)	<0.001	27.07 (1.81)	<0.001	27.15 (1.81)	<0.001	26.85 (1.80)	<0.001	26.53 (1.79)	<0.001
R^2	0.2259		0.2253		0.2246		0.2219		0.2191	
BIC	4813.58		4807.47		4801.52		4797.66		4793.87	

Appendix 17: NSHD life course analyses (life course analyses, method 2), Aim 1

Table A7.1: NSHD tests of models for alternative life course hypotheses using childhood household amenities, educational qualifications and own occupational SEP at age 43), with outcome NART

	Women			Men			
Hypothesis (Equation number)	df	F statistic	P-value*	df	F statistic	P-value*	BIC
No effect	7,947	54.72	<0.0001	7,885	36.93	<0.0001	
Accumulation models							
Accumulation (6.3)	6,947	15.37	<0.0001	6,885	6.40	<0.0001	
Adult accumulation (6.5)	6,947	2.61	0.0162	6,885	1.00	**0.4223**	6330.48
Mutually adjusted accumulation (6.4)	4,947	1.40	**0.2319**	4,885	0.73	**0.5713**	6339.09
Critical period models							
Childhood (6.6)	6,947	60.38	<0.0001	6,885	37.49	<0.0001	
Early adulthood (6.6)	6,947	10.35	<0.0001	6,885	12.59	<0.0001	
Adulthood (6.6)	6,947	37.23	<0.0001	6,885	13.44	<0.0001	
Social mobility models							
Inter generational (6.8)	5,947	54.52	<0.0001	5,885	43.89	<0.0001	
Intra generational (6.10)	5,947	61.58	<0.0001	5,885	45.46	<0.0001	
Any mobility (6.12)	5,947	63.89	<0.0001	5,885	42.75	<0.0001	
Any mobility with 3-way interaction (6.14)	4,947	32.80	<0.0001	4,885	12.69	<0.0001	

* The p-values test whether the life course model is significantly different from the saturated model

Table A7.2: NSHD tests of models for alternative life course hypotheses using childhood household amenities, own occupational SEP at age 26 and own occupational SEP at age 43), with outcome NART

	Women				Men		
Hypothesis (Equation number)	df	F statistic	P-value*	BIC	df	F statistic	P-value*
No effect	7,947	20.15	<0.0001		7,885	32.6	<0.0001
Accumulation models							
Accumulation (6.3)	6,947	4.30	0.0003		6,885	4.63	0.0001
Adult accumulation (6.5)	6,947	1.53	**0.1661**	6850.38	6,885	2.78	0.0110
Mutually adjusted accumulation (6.4)	4,947	1.44	**0.2174**	6860.01	4,885	1.89	**0.1097**
Critical period models							
Childhood (6.6)	6,947	20.94	<0.0001		6,885	32.35	<0.0001
Early adulthood (6.6)	6,947	9.58	<0.0001		6,885	7.69	<0.0001
Adulthood (6.6)	6,947	8.89	<0.0001		6,885	14.15	<0.0001
Social mobility models							
Inter generational (6.8)	5,947	25.27	<0.0001		5,885	36.99	<0.0001
Intra generational (6.10)	5,947	21.67	<0.0001		5,885	42.97	<0.0001
Any mobility (6.12)	5,947	21.83	<0.0001		5,885	38.77	<0.0001
Any mobility with 3-way interaction (6.14)	4,947	8.71	<0.0001		4,885	17.90	<0.0001

* The p-values test whether the life course model is significantly different from the saturated model

Table A7.3: NSHD tests of models for alternative life course hypotheses using childhood household amenities, educational qualifications and head of household occupational SEP at age 43), with outcome NART

	Women		
Hypothesis (Equation number)	df	F statistic	P-value*
No effect	7,947	55.85	<0.0001
Accumulation models			
Accumulation (6.3)	6,947	16.24	<0.0001
Adult accumulation (6.5)	6,947	3.96	0.0006
Mutually adjusted accumulation (6.4)	4,947	0.65	**0.6257**
Critical period models			
Childhood (6.6)	6,947	59.47	<0.0001
Early adulthood (6.6)	6,947	8.72	<0.0001
Adulthood (6.6)	6,947	41.75	<0.0001
Social mobility models			
Inter generational (6.8)	5,947	57.58	<0.0001
Intra generational (6.10)	5,947	65.91	<0.0001
Any mobility (6.12)	5,947	68.66	<0.0001
Any mobility with 3-way interaction (6.14)	4,947	28.78	<0.0001

* The p-values test whether the life course model is significantly different from the saturated model

Table A7.4: NSHD tests of models for alternative life course hypotheses using childhood household amenities, own occupational SEP at age 26 and head of household occupational SEP at age 43), with outcome NART

	women			
Hypothesis (Equation number)	df	F statistic	P-value*	BIC
No effect	7,947	21.2	<0.0001	
Accumulation models				
Accumulation (6.3)	6,947	4.36	0.0002	
Adult accumulation (6.5)	6,947	1.48	**0.1823**	6837.17
Mutually adjusted accumulation (6.4)	4,947	1.28	**0.2745**	6847.65
Critical period models				
Childhood (6.6)	6,947	21.64	<0.0001	
Early adulthood (6.6)	6,947	10.65	<0.0001	
Adulthood (6.6)	6,947	11.25	<0.0001	
Social mobility models				
Inter generational (6.8)	5,947	27.26	<0.0001	
Intra generational (6.10)	5,947	24.84	<0.0001	
Any mobility (6.12)	5,947	24.97	<0.0001	
Any mobility with 3-way interaction (6.14)	4,947	13.97	<0.0001	

* The p-values test whether the life course model is significantly different from the saturated model

Table A17.5: Whitehall II tests of models for alternative life course using childhood material deprivation, educational qualifications and own occupational SEP, adjusted for age at phase 9 and the number of times the cognitive tests were taken, with outcome Mill Hill test

	Women			Men		
Hypothesis (Equation number)	df	F statistic	P-value*	df	F statistic	P-value*
No effect	7,816	28.40	<0.0001	7,2,430	63.59	<0.0001
Accumulation models						
Accumulation (6.3)	6,816	8.90	<0.0001	6,2,430	11.35	<0.0001
Adult accumulation (6.5)	6,816	2.94	0.0076	6,2,430	4.02	0.0005
Mutually adjusted accumulation (6.4)	4,816	3.49	0.0078	4,2,430	1.53	**0.1901**
Critical period models						
Childhood (6.6)	6,816	31.26	<0.0001	6,2,430	71.02	<0.0001
Early adulthood (6.6)	6,816	14.27	<0.0001	6,2,430	36.15	<0.0001
Adulthood (6.6)	6,816	7.59	<0.0001	6,2,430	16.45	<0.0001
Social mobility models						
Inter generational (6.8)	5,816	32.26	<0.0001	5,2,430	73.41	<0.0001
Intra generational (6.10)	5,816	38.00	<0.0001	5,2,430	86.42	<0.0001
Any mobility (6.12)	5,816	29.45	<0.0001	5,2,430	64.22	<0.0001
Any mobility with 3-way interaction (6.14)	4,816	3.36	0.0097	4,2,430	16.15	<0.0001

* The p-values test whether the life course model is significantly different from the saturated model

Table A17.6: NSHD: tests of models for alternative life course hypotheses (father's occupational SEP, educational qualifications at age 26, own occupational SEP at age 43), with outcome NART, adjusted for childhood cognitive function

	Women			Men			
Hypothesis	df	F statistic	P-value*	df	F statistic	P-value*	BIC
No effect	7,945	18.09	<0.0001	7,883	13.28	<0.0001	
Accumulation models							
Accumulation	6,945	2.86	0.0092	6,883	2.27	0.0350	
Adult accumulation	6,945	3.39	0.0026	6,883	0.85	**0.5305**	6080.87
Mutually adjusted accumulation	4,945	0.92	**0.4519**	4,883	0.68	**0.6072**	6090.72
Critical period models							
Childhood	6,945	16.16	<0.0001	6,883	12.69	<0.0001	
Early adulthood	6,945	3.69	0.0013	6,883	6.34	<0.0001	
Adulthood	6,945	15.67	<0.0001	6,883	5.58	<0.0001	
Social mobility models							
Inter generational	5,945	21.98	<0.0001	5,883	17.31	<0.0001	
Intra generational	5,945	19.70	<0.0001	5,883	15.75	<0.0001	
Any mobility	5,945	24.04	<0.0001	5,883	16.52	<0.0001	
Any mobility with 3-way interaction	4,945	12.66	<0.0001	4,883	6.78	<0.0001	

* The p-values test whether the life course model is significantly different from the saturated model

Table A17.7: NSHD: tests of models for alternative life course hypotheses (father's occupational SEP, own occupational SEP at age 26, own occupational SEP at age 43), with outcome NART, adjusted for childhood cognitive function

	Women				Men			
Hypothesis	df	F statistic	P-value*	BIC	df	F statistic	P-value*	BIC
No effect	7,945	7.68	<0.0001		7,883	10.33	<0.0001	
Accumulation models								
Accumulation	6,945	0.70	**0.6516**	6386.90	6,883	1.53	**0.1636**	6091.76
Adult accumulation	6,945	3.45	0.0023		6,883	1.06	**0.3867**	6085.26
Mutually adjusted accumulation	4,945	1.03	**0.3895**	6414.59	4,883	0.84	**0.5019**	6096.11
Critical period models								
Childhood	6,945	5.20	<0.0001		6,883	9.11	<0.0001	
Early adulthood	6,945	4.90	0.0001		6,883	2.92	<0.0001	
Adulthood	6,945	5.61	<0.0001		6,883	5.51	<0.0001	
Social mobility models								
Inter generational	5,945	10.35	<0.0001		5,883	13.43	<0.0001	
Intra generational	5,945	10.27	<0.0001		5,883	13.79	<0.0001	
Any mobility	5,945	9.94	<0.0001		5,883	14.14	<0.0001	
Any mobility with 3-way interaction	4,945	2.42	0.0472		4,883	7.71	<0.0001	

* The p-values test whether the life course model is significantly different from the saturated model

Table A17.8: NSHD: tests of models for alternative life course hypotheses (father's occupational SEP, educational qualifications at age 26, head of household occupational SEP at age 43), with outcome NART, adjusted for childhood cognitive function

	Women		
Hypothesis	df	F statistic	P-value*
No effect	7,945	17.44	<0.0001
Accumulation models			
Accumulation	6,945	2.97	0.0071
Adult accumulation	6,945	2.90	0.0084
Mutually adjusted accumulation	4,945	0.69	**0.5966**
Critical period models			
Childhood	6,945	15.93	<0.0001
Early adulthood	6,945	4.16	0.0004
Adulthood	6,945	14.84	<0.0001
Social mobility models			
Inter generational	5,945	21.26	<0.0001
Intra generational	5,945	21.74	<0.0001
Any mobility	5,945	24.27	<0.0001
Any mobility with 3-way interaction	4,945	11.46	<0.0001

* The p-values test whether the life course model is significantly different from the saturated model

Table A17.9: NSHD: tests of models for alternative life course hypotheses (father's occupational SEP, own occupational SEP at age 26, head of household occupational SEP at age 43), with outcome NART, adjusted for childhood cognitive function

	Women			
Hypothesis	df	F statistic	P-value*	BIC
No effect	7,945	8.34	<0.0001	
Accumulation models				
Accumulation	6,945	0.52	**0.7938**	6367.96
Adult accumulation	6,945	2.73	0.0124	
Mutually adjusted accumulation	4,945	0.62	**0.6512**	6381.01
Critical period models				
Childhood	6,945	5.89	<0.0001	
Early adulthood	6,945	6.37	<0.0001	
Adulthood	6,945	4.66	0.0001	
Social mobility models				
Inter generational	5,945	11.29	<0.0001	
Intra generational	5,945	11.01	<0.0001	
Any mobility	5,945	10.79	<0.0001	
Any mobility with 3-way interaction	4,945	4.86	0.0007	

* The p-values test whether the life course model is significantly different from the saturated model

Table A17.10: NSHD: tests of models for alternative life course hypotheses (childhood household amenities, educational qualifications at age 26, own occupational SEP at age 43), with outcome NART, adjusted for childhood cognitive function

	Women			Men			
Hypothesis	df	F statistic	P-value*	df	F statistic	P-value*	BIC
No effect	7,945	17.68	<0.0001	7,883	12.45	<0.0001	
Accumulation models							
Accumulation	6,945	6.42	<0.0001	6,883	3.24	0.0037	
Adult accumulation	6,945	3.67	0.0013	6,883	0.73	**0.6237**	6080.87
Mutually adjusted accumulation	4,945	1.83	**0.1218**	4,883	0.82	**0.5132**	6091.72
Critical period models							
Childhood	6,945	19.00	<0.0001	6,883	13.12	<0.0001	
Early adulthood	6,945	2.63	0.0157	6,883	6.07	<0.0001	
Adulthood	6,945	17.1	<0.0001	6,883	4.88	<0.0001	
Social mobility models							
Inter generational	5,945	18.71	<0.0001	5,883	15.41	<0.0001	
Intra generational	5,945	17.94	<0.0001	5,883	14.99	<0.0001	
Any mobility	5,945	20.45	<0.0001	5,883	14.80	<0.0001	
Any mobility with 3-way interaction	4,945	15.17	<0.0001	4,883	5.68	<0.0001	

* The p-values test whether the life course model is significantly different from the saturated model

Table A17.11: NSHD: tests of models for alternative life course hypotheses (childhood household amenities, own occupational SEP at age 26, own occupational SEP at age 43), with outcome NART, adjusted for childhood cognitive function

	Women				Men			
Hypothesis	df	F statistic	P-value*	BIC	df	F statistic	P-value*	BIC
No effect	7,945	4.11	0.0002		7,883	10.48	<0.0001	
Accumulation models								
Accumulation	6,945	1.06	**0.3868**	6407.17	6,883	2.65	0.0149	
Adult accumulation	6,945	0.62	**0.7106**	6402.84	6,883	1.57	**0.1538**	6085.26
Mutually adjusted accumulation	4,945	0.51	**0.7264**	6414.59	4,883	1.79	**0.1289**	6096.11
Critical period models								
Childhood	6,945	4.26	0.0003		6,883	10.88	<0.0001	
Early adulthood	6,945	1.47	**0.1843**	6411.36	6,883	5.01	<0.0001	
Adulthood	6,945	2.66	0.0147		6,883	3.55	0.0018	
Social mobility models								
Inter generational	5,945	4.67	0.0003		5,883	13.54	<0.0001	
Intra generational	5,945	5.38	0.0001		5,883	13.31	<0.0001	
Any mobility	5,945	5.15	0.0001		5,883	12.63	<0.0001	
Any mobility with 3-way interaction	4,945	1.94	**0.1018**	6439.33	4,883	6.79	<0.0001	

* The p-values test whether the life course model is significantly different from the saturated model

Table A17.12: NSHD: tests of models for alternative life course hypotheses (childhood household amenities, educational qualifications at age 26, head of household occupational SEP at age 43), with outcome NART, adjusted for childhood cognitive function

	Women			
Hypothesis	df	F statistic	P-value*	BIC
No effect	7,945	16.89	<0.0001	
Accumulation models				
Accumulation	6,945	5.3	<0.0001	
Adult accumulation	6,945	1.78	**0.1003**	6350.35
Mutually adjusted accumulation	4,945	0.37	**0.8303**	6356.76
Critical period models				
Childhood	6,945	18.16	<0.0001	
Early adulthood	6,945	3.16	0.0045	
Adulthood	6,945	14.23	<0.0001	
Social mobility models				
Inter generational	5,945	18.24	<0.0001	
Intra generational	5,945	20.86	0.0436	
Any mobility	5,945	21.25	0.0057	
Any mobility with 3-way interaction	4,945	10.24	0.0040	

* The p-values test whether the life course model is significantly different from the saturated model

Table A17.13: NSHD: tests of models for alternative life course hypotheses (childhood household amenities, own occupational SEP at age 26, head of household occupational SEP at age 43), with outcome NART, adjusted for childhood cognitive function

	Women			
Hypothesis	df	F statistic	P-value*	BIC
No effect	7,945	6.13	<0.0001	
Accumulation models				
Accumulation	6,945	2.14	0.0463	
Adult accumulation	6,945	0.91	**0.4836**	6370.18
Mutually adjusted accumulation	4,945	1.16	**0.3273**	6381.01
Critical period models				
Childhood	6,945	6.66	<0.0001	
Early adulthood	6,945	4.30	0.0003	
Adulthood	6,945	2.84	0.0096	
Social mobility models				
Inter generational	5,945	7.70	<0.0001	
Intra generational	5,945	8.21	<0.0001	
Any mobility	5,945	7.19	<0.0001	
Any mobility with 3-way interaction	4,945	4.49	0.0013	

* The p-values test whether the life course model is significantly different from the saturated model

Appendix 18: Life course models, accounting for missing data

Table A18.1: NSHD: testing life course models using childhood material deprivation, educational qualifications and own occupational SEP at age 43, with outcome NART, under complete case, multiple imputation then deletion and Heckman selection

	Women			Men		
	p-value			p-value		
Hypothesis	CC	MID	Heckman	CC	MID	Heckman
No effect	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Accumulation models						
Accumulation	<0.0001	<0.0001	<0.0001	0.0037	0.0003	0.0042
Adult accumulation	0.0013	0.0113	0.0053	**0.6237**	**0.3858**	**0.9111**
Mutually adjusted accumulation	**0.1218**	**0.7970**	**0.0725**	**0.5132**	**0.4297**	**0.8515**
Critical period models						
Childhood	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Early adulthood	0.0157	0.0293	0.0319	<0.0001	<0.0001	0.0187
Adulthood	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0002
Social mobility models						
Inter generational	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Intra generational	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility with 3-way interaction	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0004

Table A18.2: NSHD: testing life course models using childhood material deprivation, own occupational SEP at age 26 and own occupational SEP at age 43, with outcome NART, under complete case, multiple imputation then deletion and Heckman selection

	Women			Men		
	p-value			p-value		
Hypothesis	CC	MID	Heckman	CC	MID	Heckman
No effect	0.0002	<0.0001	0.0041	<0.0001	<0.0001	<0.0001
Accumulation models						
Accumulation	**0.3868**	**0.3535**	**0.3514**	0.0149	0.0090	0.0388
Adult accumulation	**0.7106**	**0.5310**	**0.9048**	**0.1538**	**0.3334**	**0.2784**
Mutually adjusted accumulation	**0.7264**	**0.8706**	**0.7517**	**0.1289**	**0.3378**	**0.2024**
Critical period models						
Childhood	0.0003	<0.0001	0.0031	<0.0001	<0.0001	<0.0001
Early adulthood	**0.1843**	**0.0283**	**0.3407**	<0.0001	<0.0001	0.0009
Adulthood	0.0147	0.0003	**0.1366**	0.0018	0.0014	**0.1044**
Social mobility models						
Inter generational	0.0003	<0.0001	0.0096	<0.0001	<0.0001	<0.0001
Intra generational	0.0001	<0.0001	0.0014	<0.0001	<0.0001	<0.0001
Any mobility	0.0001	<0.0001	0.0051	<0.0001	<0.0001	<0.0001
Any mobility with 3-way interaction	**0.1018**	0.0050	**0.0659**	<0.0001	0.0002	0.0023

Table A18.3: NSHD: testing life course models using childhood material deprivation, educational qualifications and head of household occupational SEP at age 43, with outcome NART, under complete case, multiple imputation then deletion and Heckman selection

	Women		
	p-value		
Hypothesis	CC	MID	Heckman
No effect	<0.0001	<0.0001	<0.0001
Accumulation models			
Accumulation	<0.0001	<0.0001	0.0001
Adult accumulation	**0.1003**	0.0234	0.0498
Mutually adjusted accumulation	**0.8303**	**0.9097**	**0.5326**
Critical period models			
Childhood	<0.0001	<0.0001	<0.0001
Early adulthood	0.0045	0.0031	**0.0554**
Adulthood	<0.0001	<0.0001	<0.0001
Social mobility models			
Inter generational	<0.0001	<0.0001	<0.0001
Intra generational	0.0436	<0.0001	<0.0001
Any mobility	0.0057	<0.0001	<0.0001
Any mobility with 3-way interaction	0.0040	<0.0001	<0.0001

Table A18.4: NSHD: testing life course models using childhood material deprivation, own occupational SEP at age 26 and head of household occupational SEP at age 43, with outcome NART, under complete case, multiple imputation then deletion and Heckman selection

	Women		
	p-value		
Hypothesis	CC	MID	Heckman
No effect	<0.0001	<0.0001	0.0003
Accumulation models			
Accumulation	0.0463	0.0760	**0.1214**
Adult accumulation	**0.4836**	**0.0190**	**0.6435**
Mutually adjusted accumulation	**0.3273**	**0.5090**	**0.3859**
Critical period models			
Childhood	<0.0001	<0.0001	0.0002
Early adulthood	0.0003	0.0002	0.0221
Adulthood	0.0096	<0.0001	**0.0503**
Social mobility models			
Inter generational	<0.0001	<0.0001	0.0003
Intra generational	<0.0001	<0.0001	0.0001
Any mobility	<0.0001	<0.0001	0.0004
Any mobility with 3-way interaction	0.0013	<0.0001	0.0128

Table A18.5: Whitehall II: testing life course models using childhood material deprivation, educational qualifications and last recorded own occupational SEP at phase 7, with outcome Mill Hill test, under complete case, multiple imputation then deletion and Heckman selection

	Women			Men		
	p-value			p-value		
Hypothesis	CC	MID	Heckman	CC	MID	Heckman
No effect	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Accumulation models						
Accumulation	<0.0001	<0.0001	0.0021	<0.0001	<0.0001	<0.0001
Adult accumulation	0.0076	0.0041	**0.4009**	0.0005	0.0001	0.0051
Mutually adjusted accumulation	0.0078	**0.2548**	**0.3731**	**0.1901**	**0.0739**	0.0274
Critical period models						
Childhood	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Early adulthood	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Adulthood	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Social mobility models						
Inter generational	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Intra generational	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility with 3-way interaction	0.0097	<0.0001	0.0050	<0.0001	<0.0001	<0.0001

Appendix 19: Weighted life course analyses (life course analyses, method 2), Aim 1

Table A19.1: NSHD: testing the weighted life course hypotheses (father's occupational SEP, educational qualifications at age 26, own occupational SEP at age 43), with outcome NART

	Women				Men			
Hypothesis	df	F statistic	P-value*	BIC	df	F statistic	P-value*	BIC
No effect	7,945	18.09	<0.0001		7,883	13.28	<0.0001	
Accumulation models								
Accumulation	6,945	2.41	0.0259		6,883	1.83	**0.0913**	6083.28
Adult accumulation	6,945	3.39	0.0026		6,883	0.85	**0.5305**	6080.87
Mutually adjusted accumulation	4,945	0.92	**0.4519**	6362.36	4,883	0.68	**0.6072**	6090.72
Critical period models								
Childhood	6,945	16.16	<0.0001		6,883	12.69	<0.0001	
Early adulthood	6,945	3.69	0.0013		6,883	6.34	<0.0001	
Adulthood	6,945	15.67	<0.0001		6,883	5.58	<0.0001	
Social mobility models								
Inter generational	5,945	2.74	0.0182		5,883	6.81	<0.0001	
Intra generational	5,945	3.62	0.0030		5,883	4.36	0.0006	
Any mobility	5,945	2.11	**0.0615**	6363.10	5,883	2.80	0.0161	
Any mobility with 3-way interaction	4,945	2.21	**0.0655**	6368.67	4,883	1.57	**0.1803**	6097.24

* The p-values test whether the life course model is significantly different from the saturated model

Table A19.2: NSHD: testing the weighted life course hypotheses (childhood household amenities, educational qualifications at age 26, own occupational SEP at age 43), with outcome NART

	Women			Men			
Hypothesis	df	F statistic	P-value*	df	F statistic	P-value*	BIC
No effect	7,945	17.68	<0.0001	7,883	12.45	<0.0001	
Accumulation	6,945	5.68	<0.0001	6,883	2.70	0.0133	
Adult accumulation	6,945	3.67	0.0013	6,883	0.73	**0.6237**	6080.87
Mutually adjusted accumulation	4,945	1.83	**0.1218**	4,883	0.82	**0.5132**	6091.72
Critical Period							
Childhood	6,945	19.00	<0.0001	6,883	13.12	<0.0001	
Early adulthood	6,945	2.63	0.0157	6,883	6.07	<0.0001	
Adulthood	6,945	17.1	<0.0001	6,883	4.88	<0.0001	
Social Mobility							
Inter generational	5,945	4.80	0.0002	5,883	6.59	<0.0001	
Intra generational	5,945	2.95	0.0120	5,883	3.95	0.0015	
Any mobility	5,945	6.15	<0.0001	5,883	2.20	**0.0520**	6095.96
Any mobility with 3-way interaction	4,945	5.89	0.0001	4,883	1.37	**0.2415**	6096.67

* The p-values test whether the life course model is significantly different from the saturated model

Table A19.3: NSHD: testing the weighted life course hypotheses (father's occupational SEP, educational qualifications at age 26, head of household occupational SEP at age 43), with outcome NART

	Women				Men			
Hypothesis	df	F statistic	P-value*	BIC	df	F statistic	P-value*	BIC
No effect	7,945	7.68	<0.0001		7,883	10.33	<0.0001	
Accumulation	6,945	0.74	**0.6166**	6387.45	6,883	1.27	**0.2701**	6089.62
Adult accumulation	6,945	3.45	0.0023		6,883	1.06	**0.3867**	6085.26
Mutually adjusted accumulation	4,945	1.03	**0.3895**	6400.20	4,883	0.84	**0.5019**	6097.50
Critical Period								
Childhood	6,945	5.20	<0.0001		6,883	9.11	<0.0001	
Early adulthood	6,945	4.90	0.0001		6,883	2.92	<0.0001	
Adulthood	6,945	5.61	<0.0001		6,883	5.51	<0.0001	
Social Mobility								
Inter generational	5,945	3.17	0.0077		5,883	3.30	0.0058	
Intra generational	5,945	4.62	0.0004		5,883	1.66	**0.1419**	6101.26
Any mobility	5,945	2.33	0.0405		5,883	0.94	**0.4550**	6096.08
Any mobility with 3-way interaction	4,945	0.94	**0.4388**	6400.27	4,883	1.08	**0.3641**	6100.88

* The p-values test whether the life course model is significantly different from the saturated model

Table A19.4: NSHD: testing the weighted life course hypotheses (father's occupational SEP, own occupational SEP at age 26, own occupational SEP at age 43), with outcome NART

	Women				Men			
Hypothesis	df	F statistic	P-value*	BIC	df	F statistic	P-value*	BIC
No effect	7,945	4.11	0.0002		7,883	10.48	<0.0001	
Accumulation	6,945	0.88	**0.5063**	6405.53	6,883	2.27	0.0348	
Adult accumulation	6,945	0.62	**0.7106**	6402.84	6,883	1.57	**0.1538**	6085.26
Mutually adjusted accumulation	4,945	0.51	**0.7264**	6414.59	4,883	1.79	**0.1289**	6096.11
Critical Period								
Childhood	6,945	4.26	0.0003		6,883	10.88	<0.0001	
Early adulthood	6,945	1.47	**0.1843**	6411.36	6,883	5.01	<0.0001	
Adulthood	6,945	2.66	0.0147		6,883	3.55	0.0018	
Social Mobility								
Inter generational	5,945	2.14	**0.0586**	6420.10	5,883	5.44	0.0001	
Intra generational	5,945	0.92	**0.4674**	6410.46	5,883	3.24	0.0066	
Any mobility	5,945	0.73	**0.5996**	6409.71	5,883	2.30	0.0436	
Any mobility with 3-way interaction	4,945	0.35	**0.8445**	6413.60	4,883	2.64	0.0325	

* The p-values test whether the life course model is significantly different from the saturated model

Table A19.5: NSHD: testing the weighted life course hypotheses (father's occupational SEP, educational qualifications at age 26, head of household occupational SEP at age 43), with outcome NART

	Women			
Hypothesis	df	F statistic	P-value*	BIC
No effect	7,945	17.44	<0.0001	
Accumulation	6,945	2.50	0.0084	
Adult accumulation	6,945	2.90	0.0084	
Mutually adjusted accumulation	4,945	0.69	**0.5966**	6349.40
Critical Period				
Childhood	6,945	15.93	<0.0001	
Early adulthood	6,945	4.16	0.0004	
Adulthood	6,945	14.84	<0.0001	
Social Mobility				
Inter generational	5,945	3.65	0.0028	
Intra generational	5,945	2.94	0.0123	
Any mobility	5,945	1.62	**0.1522**	6345.26
Any mobility with 3-way interaction	4,945	1.53	**0.1913**	6351.14

* The p-values test whether the life course model is significantly different from the saturated model

Table A19.6: NSHD: testing the weighted life course hypotheses (childhood household amenities, educational qualifications at age 26, head of household occupational SEP at age 43), with outcome NART

	Women			
Hypothesis	df	F statistic	P-value*	BIC
No effect	7,945	16.89	<0.0001	
Accumulation	6,945	4.53	0.0002	
Adult accumulation	6,945	1.78	**0.1003**	6350.35
Mutually adjusted accumulation	4,945	0.37	**0.8303**	6356.76
Critical Period				
Childhood	6,945	18.16	<0.0001	
Early adulthood	6,945	3.16	0.0045	
Adulthood	6,945	14.23	<0.0001	
Social Mobility				
Inter generational	5,945	5.27	0.0001	
Intra generational	5,945	1.78	**0.1137**	6353.44
Any mobility	5,945	3.97	0.0014	
Any mobility with 3-way interaction	4,945	2.62	0.0337	

* The p-values test whether the life course model is significantly different from the saturated model

Table A19.7: NSHD: testing the weighted life course hypotheses (father's occupational SEP, own occupational SEP at age 26, head of household occupational SEP at age 43), with outcome NART

	Women			
Hypothesis	df	F statistic	P-value*	BIC
No effect	7,945	8.34	<0.0001	
Accumulation	6,945	0.49	**0.8166**	6367.70
Adult accumulation	6,945	2.73	0.0124	
Mutually adjusted accumulation	4,945	0.62	**0.6512**	6381.01
Critical Period				
Childhood	6,945	5.89	<0.0001	
Early adulthood	6,945	6.37	<0.0001	
Adulthood	6,945	4.66	0.0001	
Social Mobility				
Inter generational	5,945	4.78	0.0003	
Intra generational	5,945	3.5	0.0039	
Any mobility	5,945	3.01	0.0106	
Any mobility with 3-way interaction	4,945	1.86	**0.1157**	6387.62

* The p-values test whether the life course model is significantly different from the saturated model

Table A19.8: NSHD: testing the weighted life course hypotheses (childhood household amenities, own occupational SEP at age 26, head of household occupational SEP at age 43), with outcome NART

	Women			
Hypothesis	df	F statistic	P-value*	BIC
No effect	7,945	6.13	<0.0001	
Accumulation	6,945	1.85	**0.0863**	6389.93
Adult accumulation	6,945	0.91	**0.4836**	6381.57
Mutually adjusted accumulation	4,945	1.16	**0.3273**	6394.15
Critical Period				
Childhood	6,945	6.66	<0.0001	
Early adulthood	6,945	4.30	0.0003	
Adulthood	6,945	2.84	0.0096	
Social Mobility				
Inter generational	5,945	5.51	0.0001	
Intra generational	5,945	0.92	**0.4640**	6385.79
Any mobility	5,945	1.48	**0.1928**	6390.82
Any mobility with 3-way interaction	4,945	1.33	**0.2565**	6395.31

* The p-values test whether the life course model is significantly different from the saturated model

Table A19.9: Whitehall II weighted life course models using childhood material deprivation, educational qualifications and last recorded own occupational SEP at phase 7, with outcome Mill Hill test

Hypothesis	Women			Men		
	df	F-statistic	P-value*	df	F-statistic	P-value*
No effect	7,816	28.48	<0.0001	7,2,430	64.89	<0.0001
Accumulation models						
Accumulation	6,816	5.04	<0.0001	6,2,430	4.74	0.0001
Adult accumulation	6,816	2.13	0.0479	6,2,430	4.08	0.0004
Mutually adjusted accumulation	4,816	2.57	0.0367	4,2,430	1.51	**0.1969**
Critical period models						
Childhood	6,816	31.26	<0.0001	6,2,430	72.66	<0.0001
Early adulthood	6,816	14.29	<0.0001	6,2,430	37.33	<0.0001
Adulthood	6,816	7.00	<0.0001	6,2,430	16.76	<0.0001
Social mobility models						
Inter generational	5,816	15.31	<0.0001	5,2,430	46.02	<0.0001
Intra generational	5,816	5.48	0.0001	5,2,430	9.81	<0.0001
Any mobility	5,816	3.31	0.0058	5,2,430	6.85	<0.0001
Any mobility with 3-way interaction	4,816	2.96	0.0191	4,2,430	3.46	0.0079

* The p-values test whether the life course model is significantly different from the saturated model

Appendix 20: Developing the Whitehall II imputation model, Aim 2

Table A20.1: Step 1 of identifying variables for the imputation model: whether variables predicted missingness of variables in the model of interest

		Significant predictor of missing:				
Phase	Variable	Memory score at phase 9	Memory score at phase 7	Memory score at phase 5	Educational qualifications	Childhood material deprivation
1	Age finished full time education	✔	✔	✔	✔	✔
1	Accommodation type	✔	✔	✔	✔	✔
1	Age mother finished full time education	✔	x	✔	x	x
1	State of health in the last year	✔	✔	✔	✔	✔
1	Any longstanding illnesses?	✔	✔	✔	x	✔
1	Smoking status	✔	✔	✔	x	✔
1	Job satisfaction	x	✔	✔	x	x
1	Usually pressed for time	✔	✔	✔	✔	✔
1	Believe no one cares much about you	x	x	x	x	x
1	Believe it is safer to trust no one	x	x	x	x	x
1	Believe you are not easily angered	x	x	x	x	x
1	Isolation score	✔	✔	✔	✔	✔
1	Depression case from GHQ	x	✔	x	x	✔
3	Last recorded occupational grade	✔	✔	✔	✔	✔
3	Marital status	✔	✔	✔	✔	✔
3	AH4 score	✔	✔	✔	✔	✔
3	Mill Hill test score	✔	✔	✔	✔	✔
3	Verbal fluency - S words	✔	✔	✔	✔	✔
3	Memory score	✔	✔	✔	✔	✔

		Significant predictor of missing:				
Phase	Variable	Memory score at phase 9	Memory score at phase 7	Memory score at phase 5	Educational qualifications	Childhood material deprivation
3	Verbal fluency - animals	✔	✔	✔	✔	✔
3	Job involves travel away from home	✔	✔	✔	✔	✔
4	Ever told had depression	✔	✔	✔	✔	✔
4	Ever told had anxiety	✔	✔	✔	✔	✔
5	Childhood emotional deprivation	x	x	x	x	-
5	Last recorded occupational grade	✔	✔	-	-	-
5	Ever told high blood pressure	✔	✔	-	-	-
5	Ever diagnosed with cancer	✔	✔	-	-	-
5	Deprivation score	✔	✔	-	-	-
5	AH4 score	✔	✔	-	-	-
5	Mill Hill test score	✔	✔	-	-	-
5	Verbal fluency - S words	✔	✔	-	-	-
5	Verbal fluency - animals	✔	✔	-	-	-
5	How financially secure do you feel in next 10 years	✔	✔	-	-	-
5	To what extent do you feel you might as well give up because you can't make things better for yourself	✔	✔	-	-	-
7	Last recorded occupational grade	✔	-	-	-	-
7	General health	✔	-	-	-	-
7	Health stops you from doing what you want to do	✔	-	-	-	-
7	CASP score	✔	-	-	-	-
7	Clinic or home visit	✔	-	-	-	-

		Significant predictor of missing:
Phase	Variable	Memory score at phase 9
7	MMSE score	✔
7	Still at civil service	✔
7	Marital status	✔
7	AH4 score	✔
7	Mill Hill test score	✔
7	Verbal fluency - S words	✔
7	Verbal fluency - animals	✔

Table A20.2: Step 2 of identifying variables for the imputation model: whether variables in step 1 were associated with the variables in the model of interest

		Associated with:							
Phase	Variable	Memory score at phase 9	Memory score at phase 7	Memory score at phase 5	Educational qualifications	Childhood material deprivation	Memory score at phase 3	Father's occupational SEP in childhood	Occupational grade at phase 1
1	Age finished full time education	✓	✓	✓	✓	✓	✓	✓	✓
1	Accommodation type	✓	✓	✓	✓	✓	✓	✓	✓
1	Age mother finished full time education	✓	✓	✓	✓	✓	✓	model did not converge	✓
1	State of health in the last year	✓	✓	✓	✓	✓	✓	✓	✓
1	Any longstanding illnesses?	✓	✓	✓	x	✓	x	x	x
1	Smoking status	✓	✓	✓	✓	✓	x	✓	✓
1	Job satisfaction	x	x	x	x	✓	x	x	✓
1	Usually pressed for time	✓	✓	✓	✓	x	✓	✓	✓
1	Believe no one cares much about you	x	x	x	x	✓	x	x	✓
1	Isolation score	✓	✓	✓	✓	✓	✓	✓	✓
1	Depression case from GHQ	✓	✓	x	x	✓	x	x	✓
3	Last recorded occupational grade	✓	✓	✓	✓	✓	✓	✓	✓
3	Marital status	x	x	✓	✓	x	x	✓	✓
3	AH4 score	✓	✓	✓	✓	✓	✓	✓	✓
3	Mill Hill test score	✓	✓	✓	✓	✓	✓	✓	✓
3	Verbal fluency - S words	✓	✓	✓	✓	✓	✓	✓	✓
3	Memory score	✓	✓	✓	✓	✓	✓	✓	✓

Table A20.2 continued

		Associated with:							
Phase	Variable	Memory score at phase 9	Memory score at phase 7	Memory score at phase 5	Educational qualifications	Childhood material deprivation	Memory score at phase 3	Father's occupational SEP in childhood	Occupational grade at phase 1
3	Verbal fluency - animals	✓	✓	✓	✓	✓	✓	✓	✓
3	Job involves travel away from home	✓	✓	✓	✓	✓	✓	✓	✓
4	Ever told had depression	x	x	x	x	✓	x	x	✓
4	Ever told had anxiety	x	x	x	x	✓	x	x	✓
5	Last recorded occupational grade	✓	✓	✓	✓	✓	✓	✓	✓
5	Ever told high blood pressure	✓	✓	✓	✓	✓	✓	✓	✓
5	Ever diagnosed with cancer	x	x	✓	✓	x	x	x	x
5	Deprivation score	✓	x	x	✓	✓	x	x	✓
5	AH4 score	✓	✓	✓	✓	✓	✓	✓	✓
5	Mill Hill test score	✓	✓	✓	✓	✓	✓	✓	✓
5	Verbal fluency - S words	✓	✓	✓	✓	✓	✓	✓	✓
5	Verbal fluency - animals	✓	✓	✓	✓	✓	✓	✓	✓
5	How financially secure do you feel in next 10 years	x	✓	✓	✓	✓	x	✓	✓
5	To what extent do you feel you might as well give up because you can't make things better for yourself	x	x	✓	x	✓	x	x	✓
7	Last recorded occupational grade	✓	✓	✓	✓	✓	✓	✓	✓

Table A20.2 continued

		Associated with:							
Phase	Variable	Memory score at phase 9	Memory score at phase 7	Memory score at phase 5	Educational qualifications	Childhood material deprivation	Memory score at phase 3	Father's occupational SEP in childhood	Occupational grade at phase 1
7	General health	✓	✓	✓	✓	✓	✓	✓	✓
7	Health stops you from doing what you want to do	✓	✓	✓	✓	✓	✓	✓	✓
7	CASP score	x	✓	x	x	✓	x	x	✓
7	Clinic or home visit	✓	✓	✓	✓	✓	x	✓	✓
7	MMSE score	✓	✓	✓	✓	✓	✓	✓	✓
7	Still at civil service	✓	✓	✓	✓	✓	✓	x	✓
7	Marital status	✓	x	✓	✓	✓	x	✓	✓
7	AH4 score	✓	✓	✓	✓	✓	✓	✓	✓
7	Mill Hill test score	✓	✓	✓	✓	✓	✓	✓	✓
7	Verbal fluency - S words	✓	✓	✓	✓	✓	✓	✓	✓
7	Verbal fluency - animals	✓	✓	✓	✓	✓	✓	✓	✓

Table A20.3: Potential auxiliary variables for the imputation model

Potential auxiliary variable	Hypothesised to be associated with:	Was it associated?
AH4 score (phase 9)	Memory (phase 9)	Yes
Mill Hill test score (phase 9)	Memory (phase 9)	Yes
Verbal fluency - S words (phase 9)	Memory (phase 9)	Yes
Verbal fluency - animals (phase 9)	Memory (phase 9)	Yes
Last recorded occupational SEP (phase 9)	Occupational grade at phase 1	Yes

Appendix 21: Whitehall II life course analyses, Aim 2

Table A21.1: Testing the life course hypotheses for the intercept, using multiply imputed data, with childhood material deprivation (left) and father's occupational SEP (right) as the childhood SEP measure, and outcome memory score

	Partial F-test against saturated model (p-value)			
	Childhood material deprivation		Father's occupational SEP	
Hypothesis	MI	MID	MI	MID
No effect	<0.0001	<0.0001	<0.0001	<0.0001
Accumulation models				
Accumulation	<0.0001	<0.0001	<0.0001	<0.0001
Adult accumulation	<0.0001	<0.0001	0.0003	0.0013
Mutually adjusted accumulation	**0.8285**	**0.3559**	**0.7332**	**0.4453**
Critical period models				
Childhood	<0.0001	<0.0001	<0.0001	<0.0001
Early adulthood	<0.0001	<0.0001	<0.0001	<0.0001
Adulthood	<0.0001	<0.0001	0.0067	0.0016
Social mobility models				
Inter generational	<0.0001	<0.0001	<0.0001	<0.0001
Intra generational	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility with 3 way interaction	<0.0001	<0.0001	<0.0001	<0.0001

Table A21.2: Testing the life course hypotheses for the slope, using multiply imputed data, with childhood material deprivation (left) and father's occupational SEP (right) as the childhood SEP measure, and outcome memory score

	Partial F-test against saturated model (p-value)			
	Childhood material deprivation		Father's occupational SEP	
Hypothesis	MI	MID	MI	MID
No effect	**0.1877**	0.0234	**0.7739**	**0.3292**
Accumulation models				
Accumulation	**0.8384**	**0.1444**	**0.9256**	**0.4641**
Adult accumulation	**0.4985**	**0.0949**	**0.9961**	**0.6541**
Mutually adjusted accumulation	**0.6702**	**0.2405**	**0.9558**	**0.6304**
Critical period models				
Childhood	**0.3570**	0.0282	**0.4385**	**0.1341**
Early adulthood	**0.4725**	**0.2960**	**0.9921**	**0.8966**
Adulthood	**0.2100**	0.0191	**0.8595**	**0.3229**
Social mobility models				
Inter generational	**0.0881**	0.0418	**0.7061**	**0.3974**
Intra generational	**0.0739**	**0.0767**	**0.6517**	**0.5467**
Any mobility	**0.0880**	0.0077	**0.5898**	**0.2040**
Any mobility with 3 way interaction	**0.4495**	0.0285	**0.8746**	**0.2931**

Table A21.3: Testing the life course hypotheses for the intercept and slope, using multiply imputed data, with childhood material deprivation as the childhood SEP measure, and outcome memory score

	Partial F-test against saturated model (p-value)			
	Childhood material deprivation		Father's occupational SEP	
Hypothesis	MI	MID	MI	MID
No effect	<0.0001	<0.0001	<0.0001	<0.0001
Accumulation models				
Accumulation	<0.0001	<0.0001	<0.0001	<0.0001
Adult accumulation	<0.0001	<0.0001	0.0113	0.0066
Mutually adjusted accumulation	**0.8577**	**0.2283**	**0.9606**	**0.6956**
Critical period models				
Childhood	<0.0001	<0.0001	<0.0001	<0.0001
Early adulthood	<0.0001	<0.0001	<0.0001	<0.0001
Adulthood	0.0001	<0.0001	**0.0705**	0.0041
Social mobility models				
Inter generational	<0.0001	<0.0001	<0.0001	<0.0001
Intra generational	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility	<0.0001	<0.0001	<0.0001	<0.0001
Any mobility with 3 way interaction	<0.0001	<0.0001	<0.0001	<0.0001