Socioeconomic inequalities in young children's weight status in the UK

Nichola Louise Shackleton

Department of Quantitative Social Science

Institute of Education, University of London

Thesis for the degree of Doctor of Philosophy

October 2014

Abstract

The high prevalence rates of child overweight and obesity within the UK is a serious problem, and one that has received a lot of attention from policy makers, researchers and the media. There is some evidence for socioeconomic inequalities in child overweight and obesity, with children in less advantaged socioeconomic groups at an increased risk of being overweight or obese. The nature of these inequalities is not well understood. Within this thesis I consider different aspects of socioeconomic status and their relationship with child overweight and obesity. There are three distinct strands of the investigation. Firstly, I consider whether socioeconomic inequalities in child overweight and obesity have changed over time. This is followed by two separate analyses of the relationship between obesity and overweight with parental income and education.

I find evidence that socioeconomic inequalities in child overweight and obesity have widened over time, but only because of the relatively low increases in child overweight and obesity amongst children from the most advantaged families. I investigate whether there is an association between income and child overweight, and find that other parental characteristics, namely parental education, can explain correlations between familial income and child overweight and obesity. I find that father's education has a stronger association with child overweight and obesity than does mother's education, and that this is not solely because father's education is a better predictor of the family's financial and economic resources.

Contents

List of Tables

Appendix

List of Figures

Chapter 5

Appendix

Declaration of Authorship

I, Nichola Shackleton, declare that the thesis entitled "Socioeconomic inequalities in young children's weight status in the UK" is my own work, and have been generated by me as the result of my own original research. I confirm that:

- This work was done wholly or mainly in candidature for a research degree at this university.
- Where any part of the thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed.

Signed: N. Shackleton

Date: 13/07/2014

Acknowledgements

Firstly, I would like to thank all members of my supervisory team (Dick Wiggins, John Jerrim & Ingrid Schoon) for all their help, hard work and advice. I must also thank Vernon Gayle for providing detailed comments and suggestions relating to chapter 5 of this thesis. I would also like to thank fellow students Tammy Campbell and Annabel Mullin for their friendship and support. Feedback received from other Institute of Education staff and students during my registration has been critical to the development of my own skills and each of the chapters included in this work. A special thanks goes to my loving partner Jared, who sat through and critiqued numerous practice presentations, listened to all my research ideas and never let me get too stressed out. This thesis would not have possible without his support.

Word Count

Word count including references is 50, 836.

Chapter 1: An introduction to the topic of child overweight and obesity

The World Health organisation has recognised the seriousness of child obesity stating that it is one of the most serious public health challenges for the 21st century (World Health Organisation, 2013). There are severe health, social and psychological consequences to being overweight and obese in childhood. Within the UK much has been done to try and reduce child obesity rates. The government has focussed on increasing habitual physical activity levels, increasing sports participation, increasing the quality of school based physical education, and encouraging a healthy, balanced diet (Department of Health, 2012). For example, most readers will be familiar with the 'change 4 life' public health campaign (NHS, 2008), which encourages families to move more, eat well and live longer. There have also been other initiatives such as the school fruit and vegetable scheme, which provides children with a piece of fresh fruit daily (Department of Health, 2012). However, despite these efforts child obesity rates have remained high over the past decade. Evidence indicates that, not only does childhood adiposity (fat) have a socioeconomic gradient, with children in more advantaged families less likely to be overweight or obese (El-Sayed, Scarborough and Galea, 2012a; Observatory, 2012; Shrewsbury and Wardle, 2008), but these social inequalities in childhood adiposity may have widened over time (Stamatakis *et al.*, 2005; Stamatakis, Wardle and Cole, 2010).

The inequalities in child overweight are worthy of further consideration. If children from lower socioeconomic groups do have a higher prevalence of overweight and obesity, it is possible that the high prevalence rates of child obesity in the UK could be reduced by reducing these inequalities. A reduction in these inequalities could have indirect economic advantages through a reduction in the costs of treatment to the NHS. In addition it could be considered unjust that social circumstances dictate child obesity risk and that it is a matter of social justice that these inequalities are addressed. In order to effectively tackle the socioeconomic inequalities in child overweight/obesity we have to establish what is driving the inequalities. But whilst these inequalities have been established, there are gaps in the literature concerning the long term trends in these socioeconomic inequalities, as well as disentangling which specific aspects of social position (i.e income, parental education, social class) are driving these inequalities.

This introduction provides an overview of child obesity in the UK. Following the introduction there are seven main chapters. The second chapter explains how child obesity is measured. The third chapter provides detailed information on the main data set used for analysis. The next three chapters form the empirical core of the thesis. Chapter four considers how socioeconomic inequalities in child weight status have changed over time as prevalence rates of child obesity have increased. Chapter five considers whether there is an association between low familial income and child weight status. Chapter six considers the role of parental education on child weight status. The final chapter, chapter seven, presents a summary of the results and some conclusions.

In the remainder of this introductory chapter I explain why child overweight and obesity is problematic, referring specifically to the increased prevalence of child overweight and obesity over time and the consequences of child overweight/obesity to the child themselves as well as to society as a whole. I introduce explanations regarding the increases in child overweight and obesity, and explain the extent of knowledge on socioeconomic inequalities in child obesity and how this thesis provides a unique contribution. The aim of this chapter is to give the reader an overview of childhood obesity in the UK, how problematic it is, and a brief overview of what the current state of knowledge is.

Defining obesity

Obesity and overweight are terms used to describe a state of having an excessive amount of body fat which poses a risk to health (National Obesity Observatory, 2009). Obesity represents a larger proportion of body fat and higher risks to health than does overweight. Whilst these terms refer to the amount of fat in the body, the only way to truly measure this is through dissection.

Overweight and obesity are most commonly measured through Body Mass Index (BMI) (National Obesity Observatory, 2009). BMI represents a ratio of the persons weight for their height and is calculated in the following way: BMI= (Weight (kg)/ ((Height (m)^2)). Whilst this is a measure of excessive weight, rather than excessive fatness, and is influenced by both fat mass and lean (muscle) mass in the body, decades of research show that it does provide a good proxy measure for body fat and is associated with concurrent health risks (Lindsay *et al.*, 2001; National Obesity Observatory, 2009). This issue is discussed further in chapter two.

The prevalence of child obesity in the UK and changes over time.

The increases in obesity over time have been described as an epidemic[1] (James *et al.*, 2001). Current prevalence estimates of childhood overweight and obesity in the UK suggest that around one quarter to one third of children are overweight or obese, with the prevalence of overweight and obesity generally being higher for older children (Health and Social Care Information Centre, 2013a; Health and Social Care Information Centre, 2013b). International comparisons show that the UK also has high levels of child obesity, particularly when compared to other European countries (OECD, 2012; OECD, 2013).

Figure 1.1 shows the longer term trends in child overweight and obesity in the UK for boys and girls. As can be seen in figure 1.1, there have been large increases in child overweight/obesity over time, and these increases accelerated in the 1990's (Chinn and Rona, 2001; Jackson-Leach and Lobstein, 2006; Stamatakis *et al.*, 2005). This is shown by the change in the gradient of the lines for overweight and obesity between 1974 and 1994, and 1994 and 2002. In 1994 an estimated 13% of girls and 9% of boys were classified as overweight. Of these 3% of girls and 2% of boys were classified as obese. By 2002 an estimated 23% of girls and 16% of boys were classified as overweight, with 7% of girls and 5% of boys classified as obese.

The shorter terms trends in the prevalence of child overweight are shown in figure 1.2[2]. The shorter term trends suggest that from the mid 2000's onwards there has been

[1] In epidemiology an epidemic is the occurrence of more cases of a disease than would normally be expected in a specific place or group of people over a given period of time. The oxford dictionary also defines an epidemic as a sudden, widespread occurrence of an undesirable phenomenon.
[2] The criteria of classifying children as overweight is different in figures 1.1 and 1.2. In figure 1.1 the International Obesity Task Force criteria are used, in figure 1.2 the British 1990 reference criteria are used. The IOTF criteria are more conservative and always produce lower estimates than the UK90 criteria. This is discussed further in chapter 2.

relatively stability in the prevalence of child overweight (Health and Social Care Information Centre, 2013a; Health and Social Care Information Centre, 2013b).

Figure 1.1. The percentage of children aged 5-10 classified as overweight over time.

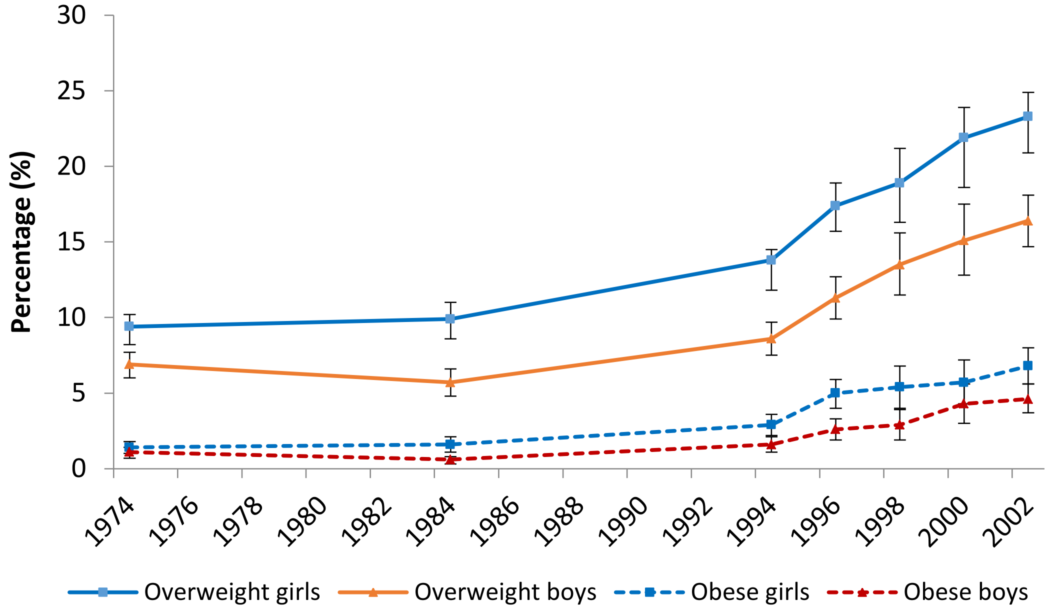


Figure notes: Adapted from 'Overweight and obesity trends from 1974 to 2003 in English children: what is the role of socioeconomic factors?' (Stamatakis et al., 2005).Error bars show 95% confidence intervals. Overweight and obesity measured using IOTF criteria. Data between 1974 & 1994 from National Study of Health growth, data from 1996 onwards from Health survey for England.

Figure 1.2. The proportion of children classified as overweight (including obese) from 1995-2011 for children of specific age groups.

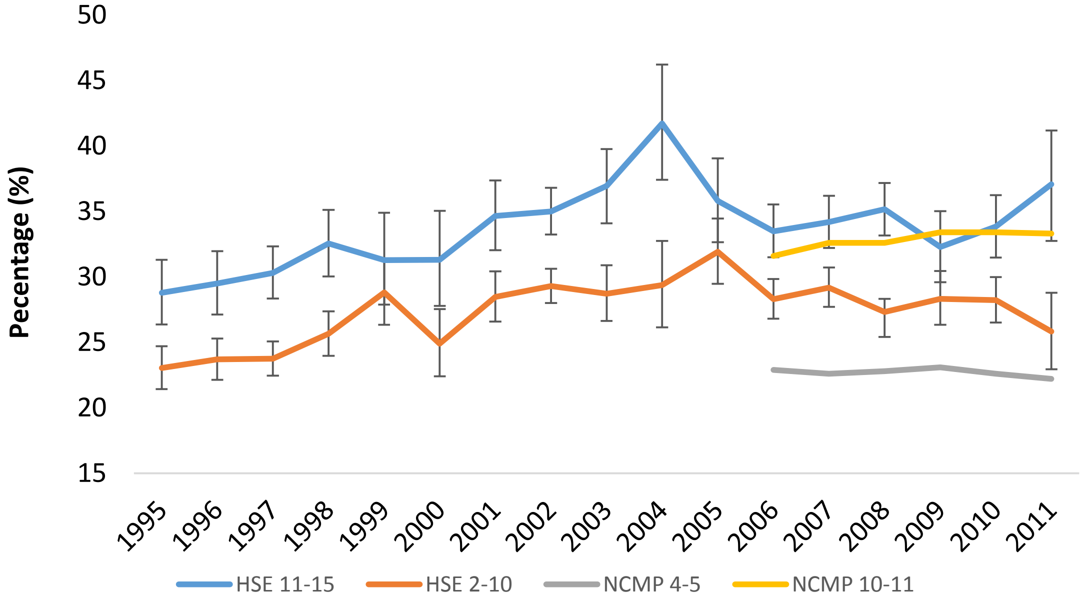


Figure notes: Error bars show 95% confidence intervals. Overweight and obesity measured using Figure notes: UK90 criteria. Data taken from Health Survey for England (HSE) trend tables (available from www.ic.nhs.uk) and from the National Child Measurement Programme (NCMP) available from www.hscic.gov.uk.

The consequences of child obesity

Health consequences

Overweight and Obesity are well-established risk factors for numerous chronic and life threatening conditions including strokes, Type II diabetes, several forms of cancer, cardiovascular disease, asthma, gall bladder disease, osteoarthritis and chronic back pain (Guh *et al.*, 2009). Whilst many obese children do not appear to have any immediate health consequences, they are more likely to have risk factors which make children susceptible to chronic diseases later on in life such as raised blood pressure, raised cholesterol and hormonal changes (Daniels, 2006; Public Health England, 2013). However, some obesity related conditions do develop during childhood such as type II diabetes, asthma & musculoskeletal problems such as Blount's disease (Daniels, 2006; Hannon, Rao and Arslanian, 2005; Wills, 2004). These are life-long health conditions, requiring continuous treatment throughout the life course, regardless of weight status in adulthood.

Childhood overweight and obesity have also been linked to adult morbidities (Dietz, 1998; Reilly and Kelly, 2011), although this may be explained by the tracking of obesity from childhood to adulthood (Park *et al.*, 2012; Reilly and Kelly, 2011). Obese or overweight children are at increased risk of being obese in adulthood, particularly during young adulthood (Herman *et al.*, 2009; Serdula *et al.*, 1993; Singh *et al.*, 2008). Furthermore a dose-response relationship has been suggested between the length of time a person remains obese and the risks for cardiovascular disease, cancer, and all-cause mortality (Abdullah *et al.*, 2011). Therefore, overweight children are not only at risk of immediate health consequences, they are also at increased risk of long term health consequences due to the physiological changes induced in the body, and because overweight and obesity can persist into adulthood.

Psychological & Social consequences

There are not just physical consequences to childhood adiposity, but psychological and social consequences also (Puhl and Heuer, 2009; Rees *et al.*, 2009; Scott *et al.*, 2007). Evidence suggests that obesity has become increasingly stigmatised, despite the increasing frequency of the condition (Andreyeva, Puhl and Brownell, 2008; Latner and

Stunkard, 2003). Obese children and young people may experience negative attitudes from teachers, other parents, health professionals and peers (Neumark-Sztainer, Story and Harris, 1999; Peterson, Puhl and Luedicke, 2012; Puhl and Heuer, 2009; Puhl and Latner, 2007; Schroer, 1985; Zeller, Reiter-Purtill and Ramey, 2008). Indeed, implicit association experiments, whereby both adults and children have been asked to make judgements about images of 'normal healthy' children, children with differing disabilities, and obese children, have consistently demonstrated that obese children are attributed with the most negative and fewest positive attributes (Latner and Stunkard, 2003; Maddox, Back and Liederman, 1968; Musher-Eizenman *et al.*, 2004; Richardson *et al.*, 1961; Schwartz *et al.*, 2006).

The negative attitudes held towards child obesity could have consequences for children's life chances. If there is any bias based on the child's weight status in the way in which teachers, principals, health professional, or any other adult interacts with children, then children who are obese or overweight are not being given an equal opportunity to fulfil their potential. However, whilst the stigma towards obesity and child obesity is well established in the literature (Puhl, Heuer and Brownell, 2010), the actual impact that these negative stereotypes has on childhood outcomes has received less attention. Although I have investigated this, and I find little evidence that teacher judgements of pupils ability are indeed influenced by children's body shape (Shackleton and Campbell, 2013). Suggesting that whilst these negative stereotypes may exist, they do not necessarily influence children's chances of achieving highly in the first few years of school.

In a systematic review on children's views about obesity, the social consequences of obesity were of the most importance to children. Popularity, friendships and bullying were children's main concerns regarding obesity and weight gain (Rees *et al.*, 2009). The concerns of the children expressed in this review may reflect social realities. An analysis of children's social networks by Strauss & Pollack (2003), demonstrated that overweight children were less likely to be nominated as friends by other children, and tended to be on the periphery of social circles (Strauss and Pollack, 2003). Furthermore obese children may be more likely to be the victims of bullying or experience 'weight related teasing' (Haines *et al.*, 2013; Janssen *et al.*, 2004). Weight related teasing has been linked to low self-esteem, depressive symptoms, body dissatisfaction and weight control behaviours (Greenleaf, Petrie and Martin, 2014; Lampard *et al.*, 2014; Olvera *et al.*, 2013).

Shafer and Ferraro (2011) argue that the impact of this stigmatisation and social marginalisation could have far reaching consequences, including long term impacts on quality of life, psychological well-being & life chances. The psychological impact of such stigmatisation can also impact upon immediate and long term physical well-being (Shafer and Ferraro, 2011). Equally physical wellbeing can influence psychosocial wellbeing (Biddle and Asare, 2011). Therefore the effect of obesity on physical and psychosocial wellbeing is not mutually exclusive.

Economic consequences

In this section I consider the economic consequences of obesity to the individual and to society as a whole. I consider both the economic consequences of adult overweight and obesity as well as the economic consequences of child overweight and obesity. This is because overweight and obese children are more likely to become overweight or obese adults.

Obese children tend to attain at lower levels on average than non-obese children academically (Caird *et al.*, 2011). This, as well as potential hiring discrimination against the obese (Agerstrom and Rooth, 2011), has knock on effects to the opportunities that are open to obese people in the labour market. Furthermore, evidence suggests that even when accounting for the type of occupation and qualifications obtained, obese people, particularly obese females receive lower wages and are less likely to be offered jobs (Agerstrom and Rooth, 2011; Caliendo and Lee, 2013; Harper, 2000). Longitudinal research, which can disentangle the onset and duration of obesity and potential consequences of obesity, suggests that becoming obese is associated with a drop in wages relative to the non-obese rather than a drop in wages causing obesity (Baum and Ford, 2004). The obesity wage penalty, which more clearly exists for female's, is somewhat ironic given that many occupations require many hours of sedentary activity in order to perform the job (Church *et al.*, 2011).

The wider economic costs of overweight and obesity are very large. In 2007 the Foresight report showed that the cost of overweight and obesity in childhood and adulthood to the UK economy was £15.8 billion, including £4.2 billion in costs to the NHS (McPherson, Marsh and Brown, 2007). Although other estimates suggest an even higher cost of overweight and obesity to the NHS in 2006/2007 (Scarborough *et al.*, 2011). The wider costs of overweight and obesity included factors such as obesity related health conditions, missed days of work due to obesity related health problems, lower productivity, early mortality, unemployment and incapacity benefits. The Projections in the Foresight report suggested that by 2015 the costs of both overweight and obesity to the UK economy would be £27 billion and the direct costs to the NHS would be as high as £6.4 billion pounds. Other projections based on data from the Health survey for England[3] (HSE) (using data from 1993 -2008) suggest that the cost of treating overweight and obesity and related conditions would rise by up £2billion per year from 2008 onwards (Wang *et al.*, 2011).

Whilst these estimates are the best available, they are several years out of date and are based upon estimated obesity levels where no action is taken to address the obesity problem. These estimates are also based on a scenario, where in 2015 48% of the adult population is either overweight or obese. However, estimates from the HSE based on 2012 data suggest that already more than 48% of the adult population in the UK is classified as overweight or obese. Indeed an estimated 62% of the adult population (those above 16 years of age) are now either overweight or obese (Health and Social Care Information Centre, 2013b). Even if the costs of obesity have not increased from 2007 onwards, which is unlikely, the current coalition government called for massive spending cuts from 2010 onwards to repay the UK's deficit, and consequently the budget for the NHS has been tightly squeezed (Vize, 2011). Therefore the proportion of the budget spent on overweight and obesity is likely to have increased, even if actual spending on overweight and obesity has not.

Summary of consequences

[3] *The HSE is an annual cross sectional survey, commissioned by the Health and Social Care Information Centre, which has been undertaken since 1991 and included children aged 2-15 from 1995 onwards.*

In summary, there are serious consequences of overweight and obesity in childhood. Some of these consequences will persist into adulthood and potentially throughout the whole life-course. For example an obese child who develops type II diabetes, or Blount's disease has to live with this health condition for the rest of their life. The negative perceptions held by many about the obese most likely influence children's life chances, and once in adulthood these negative perceptions result in poorer employment chances and lower wages. At the societal level, the high prevalence rates of child obesity are an increasing economic burden, with obesity related health conditions and other indirect costs of obesity adding up to a vast amount. Given the extent and severity of the consequences of child obesity it could be seen as unfair that children from disadvantaged backgrounds are further disadvantaged by being more likely to be overweight or obese than children from more advantaged backgrounds.

In the following section I provide a summary of why increases in child obesity have occurred. I follow this by outlining the evidence for socioeconomic inequalities in child obesity and the gaps in the literature.

Causes of child obesity and increases in obesity

It is largely agreed that child obesity is essentially caused by an energy imbalance, whereby calories ingested are greater than those utilised (Foresight, 2007; Garrow, 1978; World Health Organisation, 2014). There is a hypothesised energy balance system, which seeks to regulate body weight (Garrow, 1978; Hyde and Miselis, 1983; Nisbett, 1972; Van den Pol, 1982). A combination of biological, psychological, social and ecological mechanisms can impact upon this energy balance system through changes in the amount of calories ingested and amount of energy expended, either through physical activity, the basal metabolism or thermogenesis (heat production) (Oswal and Yeo, 2007; Walley, Asher and Froguel, 2009).

The Foresight report (2007b), a report published by the UK government's science based futures think tank, currently provides the most comprehensive and detailed account of the many potential contributors to obesity, and how these contributors to obesity interact (Foresight, 2007). From literature reviews and empirical work they present a ‘system map’ of the contributors to obesity. Within the system map the biological energy balance system is the central driver in obesity. There are over 100 other variables included in the system map which impact upon the energy balance system. These other variables can be clustered into the following:

- **Physiology cluster.** This cluster includes all biological aspects of body weight management including metabolic, genetic, epigenetic, endocrinal and neurological. As well as including aspects of how well the energy balance system performs, it includes appetite control, the inheritance of body composition, and the influence of maternal feeding practises and growth during childhood. These variables are all connected such that, for example, maternal feeding practises influence growth during childhood and this subsequently influences appetite control and the performance of the energy balance system. It also contains a reinforcing loop that endeavours to maintain the appropriate body composition from one generation to another.
- **Individual activity.** This cluster consists of contributors such as level of recreational, transport activity and level of fitness. There is a reciprocal loop between level of fitness and physical activity levels, as one may need a certain level of fitness to engage in certain physical activities, but engaging in activities will increase physical fitness.
- **The physical environment cluster.** This includes aspects of the physical environment that may hinder or encourage physical activity. These can be structural, such as a lack of pavement to walk on, material such as the cost of exercise, or complex such as the perceived safety of the neighbourhood in which one lives, or cultural or attitudinal obstacles to physical activity.
- **Food consumption cluster.** This cluster includes characteristics of the ‘food market’ and the health characteristics of the food products available within this market. Therefore it includes items such as the variety of food available, the convenience of different foods, the nutritional quality of the food available and the portion sizes of foods.
- **Food production cluster**. This cluster consists of the drivers of the food industry such as the desire to be profitable and the cost of ingredients of foods, but it also reflects the wider social pressures in the UK to consume.

- **Individual psychology cluster.** This cluster contains psychological characteristics such as self-esteem and stress as well as level of parental control and level of children's control of diet.
- **Social psychology cluster.** This cluster includes variables which operate at the societal level such as the social acceptability of fatness and the importance of body-size, education and the media.

A diagram of this proposed system, the 'system map' is freely available to download from www.gov.uk. There is an interactive version of this system map at http://www.shiftn.com/obesity/Full-Map.html which clearly demonstrates the proposed pathways which belong to each cluster. This system map shows that the contributors to obesity are clearly numerous and the way in which these different contributors interact is complex. Readers are referred to the Foresight report (2007) for a full discussion of these different contributors to obesity.

Increases in obesity prevalence were too quick to be caused by genetics alone, therefore explanations for the increases in obesity include changes in environmental influences and the interaction between genes and environment (Hill and Peters, 1998). The obesogenic environment theory put forward by Egger and Swinburn in the late 1990's (Egger and Swinburn, 1997; Swinburn, Egger and Raza, 1999) focusses on how changes in the macro[4] environment interact with changes in micro[5] environments and with individual behaviours. Egger and Swinburn state that the macro environment has become obesogenic, i.e it encourages obesity. This is due to changes in the food system which have increased the availability and affordability of energy dense foods. There have also been reductions in the need to engage in physical activity due to things such as increased use of personal transport, and the invention of more sedentary leisure activities. They note that these environment influences interact differently with individuals based on their own biological energy balance system. There are variations in how effectively people expend and store calories, and whilst in the past there may have been an evolutionary advantage to being able to store energy effectively, this is certainly not an advantage in an environment which encourages high calorie consumption and minimum energy expenditure.

Socioeconomic inequalities in child obesity

All human societies have some inequality, with the distribution of power[6] and prestige[7] being unequal between individuals and social groups. Social stratification refers to the ranking of social groups based on factors such as wealth and prestige. Those who belong to a particular stratum will have a similar lifestyle which to some extent distinguishes them from members of other strata (Haralambos and Holborn, 2008). Socioeconomic status (SES) is an umbrella term which has been used in the description and measurement of individual's or group's position within the hierarchical structure of society. Therefore SES has been described as reflecting access to desired resources such as leisure time, power, material goods etc. (Oakes and Rossi, 2003).

[4] Macro environments refer to a broad set of social and economic conditions that are external and uncontrollable for example the economy, government policy-making, technology, social conditions, and nature
[5] Micro environments refer to all of the immediate physical, social and economic conditions that are close to the individual. For example the local economy, family and friends, local amenities.
[6] Power is the degree to which individuals or groups can impose their will upon others.
[7] The amount of honour associated with social positions

Socioeconomic status is usually measured through income, social class, status and education (Adler and Newman, 2002; Mackenbach and Kunst, 1997; Shavers, 2007). Area level measures of deprivation, such as the Index of Multiple Deprivation (IMD) or the Income Deprivation affecting Children Index (IDACI) have also been classified by some as indicators of socioeconomic status (El-Sayed, Scarborough and Galea, 2012a; Janssen *et al.*, 2006; Observatory, 2012). These indices include aspects such as area level income, employment rate, crime rate, health and disability, education and training opportunities and housing services.

The association between SES and health more generally is well documented (Bambra *et al.*, 2010; Mackenbach *et al.*, 2008a; Marmot and Bell, 2012; Marmot *et al.*, 1997). A plethora of research conducted since the influential Black report (1980) has shown that those in lower socioeconomic positions die younger, have higher prevalence rates for diseases, higher incidence of mental health problems, are more likely to be disabled and to rate their own wellbeing lower than those in more advantaged positions in society (Fryers, Melzer and Jenkins, 2003; Marmot and Bell, 2012; Marmot *et al.*, 1997).

In the adult population in developed countries those with lower socioeconomic status, are more likely to be obese (El-Sayed, Scarborough and Galea, 2012b). Evidence indicates that a similar relationship exists between parental SES and child overweight/obesity. Children in higher socioeconomic groups are less likely to be overweight or obese than children in lower socioeconomic groups (El-Sayed, Scarborough and Galea, 2012a; Observatory, 2012; Shrewsbury and Wardle, 2008; Stamatakis *et al.*, 2005; Stamatakis, Wardle and Cole, 2010; White *et al.*, 2007).

Whilst the majority of the available evidence indicates that socioeconomic inequalities exist in child obesity, there are several problems with the literature in its current state. Firstly, there is no uniformity in the measurement of SES in the health literature and there are many different measures of SES in use (Bartley, 2004). This is well illustrated in a systematic review of studies investigating social inequalities in child obesity (under 18) in the UK from 1980 – 2010. There were twenty different measures of socioeconomic status applied across twenty three different studies (El-Sayed, Scarborough and Galea, 2012a). Although similar indicators for SES have been used, it is common for either only one indicator of SES to be used i.e just income or just social class, or for composite/latent measures of SES to be created, each using different indicators and methods in their creation (Adler and Newman, 2002; Mackenbach and Kunst, 1997; Shavers, 2007). Where just one measure of SES is used, it is not possible to determine whether an association exists between that indicator of SES and child overweight and obesity, or whether the association is the result of a high correlation between the measure of SES included in the analysis and other aspects of SES that are not included in the analysis.

Where more than one measure of SES has been utilised there is no consistency in which measures are utilised in combination, making it difficult to determine whether relationships between different indicators of SES and child overweight and obesity are consistent. Where income and social class have been jointly considered, only social class had a statistically significant association with child overweight (Jebb, Rennie and Cole, 2004). However, the measurement of income was not ideal with income divided into only two groups (above £20,000 per annum/ not above £20,000 per annum). Therefore, the lack of association between income and child obesity could be due to the poor measurement of income.

Where income and maternal education have been considered, there was evidence for an association between maternal education and child overweight, but not income and overweight (Matijasevich *et al.*, 2009). Where maternal education and social class have been jointly considered, one study found evidence of an association between maternal education and child obesity, but not social class and child obesity conditional upon maternal education (Connelly, 2011). The other suggested that neither social class nor maternal education were important predictors of child weight status (Rona and Chinn, 1982). The former study was based on a contemporary cohort of children, whereas the latter is based on data that are considerably older. It may be that the relationship between parental education, social class and child obesity have changed over time. Where income, social class and maternal education have been jointly considered, only maternal education was found to have a statistically significant association with child overweight (Durantauleria, Rona and Chinn, 1995).

The lack of uniformity in the measurement of SES suggests that It has been assumed, even if only implicitly, that the different indicators of socioeconomic status are interchangeable (Bukodi and Goldthorpe, 2012). Therefore the assumption is that the indicator used will make little difference in determining the extent of inequalities in child obesity. The assumption that different indicators of SES are interchangeable is flawed, because:

i) The logical pathways through which parental income, social class and education influence child obesity are most likely different (Adler and Newman, 2002). We may expect, for example income to operate through mechanisms such as the price of healthy and unhealthy foods, with more affluent people being able to afford a better quality diet. Whereas we might expect education to operate through equipping parents with the skills to gain better access to health information and to critically evaluate this information (Adler and Newman, 2002). Social class may operate through mechanisms such as the values and norms associated with food choice and exercise behaviour and how these link with social class identity (Bourdieu, 1984). There is scope for overlap as it may well be that highly educated people on low incomes are able to make better use of services and their own resources, than people with lower levels of education on low incomes.

ii) Evidence indicates that the strength of the relationship between SES and child obesity varies by the measurement of SES chosen (El-Sayed, Scarborough and Galea, 2012a; Shrewsbury and Wardle, 2008). El-Sayed, Scarborough and Galea (2012) reviewed all studies looking at socioeconomic inequalities in child obesity in the UK between 1980 and 2010. Seven out of eleven studies found an association between social class and child overweight or obesity, with higher social class resulting in a lower risk of overweight or obesity. Four out of seven studies found that higher levels of maternal education were associated with a lower risk of child obesity, and only one out three studies found that higher levels of parental income were associated with lower risks of child obesity. Shrewsbury and Wardle (2008), reviewed studies from all Western developed countries between 1990 and 2005 which considered the relationship between SES and child obesity. Out of twenty studies considering the relationship between parental education and child obesity, fifteen showed that higher SES was associated with lower obesity risk. Thirteen studies considered the role of parental social class, but only five found a significant association with child obesity. Four out of eleven studies found that increases in familial income reduced child obesity risk.

iii) The different indicators of SES have been shown to have independent effects on a variety of other childhood outcomes (Bukodi and Goldthorpe, 2012; Geyer *et al.*, 2006; Leinonen, Martikainen and Lahelma, 2012). Bukodi and Goldthorpe (2012) considered the association between parental social class, status, education and children's attainment. They find significant independent associations between these different aspects of SES and children's attainment. They use these findings to argue that different aspects of SES should be included in analytic models, especially where researchers are interested in socioeconomic inequalities. They state that focussing attention on only one aspect of parental SES, for example social class, will result in an overestimation of the effect of social class, because the association will be confounded by the other aspects of SES, which are not included in the analytic model. They also argue that neglecting some aspects of parental SES, such as education, will result in an underestimation of the total inequalities associated with parental SES.

However, because many different measures of SES are used in practise, it is not possible to determine which aspects of SES have a relationship with child obesity and how strong that relationship is. This means it is also not possible to create causal narratives, or explain the mechanisms through which SES is influencing child obesity rates (Bartley, 2004; Lahelma *et al.*, 2004). This is problematic for designing policies to reduce inequalities in child obesity and for designing targeted interventions to reduce child obesity. Many interventions are already aimed at low-SES families (Hollar *et al.*, 2010; Horodynski *et al.*, 2011; Jurkowski *et al.*, 2014; Lakerveld *et al.*, 2013; Lubans, Morgan and Callister, 2012; Tyler and Horner, 2008), but these interventions may be even more effective with more detailed knowledge about the relationship between SES and child obesity.

Overview

Child obesity is a serious problem. In the UK there are large numbers of children classified as overweight and obese. Excessive fat in childhood is associated not only with immediate health risks, but health risks later on life. On top of these health risks are social and economic consequences also. Therefore, children who are overweight and obese may not have the same life chances as non-overweight children. Children from lower socioeconomic backgrounds, who are already disadvantaged in many other ways are also at an increased risk of child obesity. The aim of this thesis is to look specifically at the different aspects of SES and their relationship with child obesity. This thesis seeks to answer three main questions regarding the relationship between parental SES and child obesity. 1) How has the relationship between parental SES and child obesity changed over time, and what can this tell us about inequalities in child obesity presently? 2) Is parental income associated with child obesity, and if so what is the size of this relationship? 3) What role do mother's and father's education play in child obesity risk? Individual literature reviews are provided at the start of each chapter which highlight why each particular piece of research was necessary and how it contributes to the literature.

Chapter 2: How child overweight and obesity are measured

As stated in the introductory chapter, obesity and overweight are primarily measured by BMI. BMI is a measure of excessive weight for height, but obesity is a condition of excessive body fat. Body fat can be measured in various other ways, some of which are more 'direct[8]' than others. Commonly used measures of body fat include Bioelectrical Impedance Analysis (BIA)[9] (Lukaski *et al.*, 1985), skinfold thickness (Durnin and Rahaman, 1967), waist circumference measurements (Lean, Han and Morrison, 1995), hydrostatic weighing[10] (Goldman and Buskirk, 1961), and what is now considered the gold standard of body fat measurement - Dual Energy X-ray Absorptiometry[11] (DEXA) (Mazess *et al.*, 1990). BMI tends to correlate very highly with these more direct measures of body fat amongst children and is similarly predictive of health risks as the more direct measures of body fat (Flegal *et al.*, 2010; Freedman and Sherry, 2009; Kennedy, Shea and Sun, 2009; Lindsay *et al.*, 2001; Steinberger *et al.*, 2005; Wickramasinghe *et al.*, 2009).

Lindsay et al (2001) compared 985 children's BMI measures to DEXA derived percentage body fat and fat mass for children of three different age groups: ages 5-9 years, ages 10-14 years, and ages 15-19 years, with analyses separated by gender. Across all age groups and sexes correlations between BMI and percentage body fat were between 0.83 – 0.94, and correlations between BMI and fat mass were 0.96 – 0.98. They found the strongest correlations in the younger age group (age 5-9) with correlations between BMI and percentage body fat at 0.94 for males and 0.92 for females. For children aged 10-14 correlations between BMI and percentage body fat were 0.84 for males and 0.86 for females (Lindsay *et al.*, 2001). Steinberger et al (2005) report similar correlation coefficients when comparing BMI and percentage body fat, and fat mass determined by DEXA amongst 130 children aged 11-17. The correlation

[8] Direct refers to how closely they are actually measuring the amount of fat in the body.
[9] BIA involves measuring the flow of an electric current through the body and measuring the resistance. Current flows more easily through parts of the body composed of higher concentrations of water, such as muscle and less easily through parts of the body with lower concentrations of water, such as fat tissue. By combining the resistance the current encounters with information about the person's height, weight, age and gender, predictions of body fat percentage are made.
[10] Hydrostatic weighing requires being submerged in a specialized tank of water. Bone and muscle are denser than water, therefore a person with a larger percentage of fat free mass will weigh more in the water and have a lower percent body fat.
[11] In a body scanner low dose x-rays are used to separate body mass into body minerals, fat mass and fat free mass.

between BMI and percentage body fat were 0.85, and the correlation between BMI and fat mass was 0.95.

However, findings from a study of 1196 5-18 year olds suggested that correlations between BMI and DEXA measured body fat were higher when children have higher levels of body fat (Freedman *et al.*, 2005). When children in this study had high levels of BMI, i.e when they were classified as overweight on BMI based measures of overweight, correlations between BMI and DEXA determined fat mass were between 0.85-0.96 across all age and sex categories, however for children below the median BMI the correlations were between 0.22-0.65 across age and sex categories. For these children BMI was more strongly correlated with fat free mass, with correlations between 0.56 – 0.83. Therefore BMI may only be an appropriate measure of body fat in children with higher levels of body fat, and it may not be useful to use the full distribution of BMI, even if adjusting for age and sex (Freedman *et al.*, 2005).

The fact that BMI provides a measure of excessive weight rather than excessive body fat is seen as less problematic for adults, because changes in weight usually reflect changes in fat mass. Although, there are exceptions to this, for example gains in fat free mass, i.e muscle growth, will increase BMI without increasing health risks (Prentice and Jebb, 2001). However, during childhood and adolescence, the body is growing and developing so that changes in weight may not only result from changes in fat mass. Therefore when using BMI to measure weight status in childhood, children are always compared to a 'reference population' of children of the same age and sex. The BMI of children is calculated using the standard BMI formula, and this is compared to children's BMI of the same age and sex from the reference population.

There are currently several different BMI based measures for children's weight status with differing reference populations that could be used in the UK. These are the International Obesity Task Force (IOTF) criteria (Cole *et al.*, 2000), the British 1990 Growth reference (UK90) criteria (Cole, Freeman and Preece, 1995), and the World Health Organization child growth standards (WHO) (Onis *et al.*, 2007). In this thesis I use the IOTF criteria.

The IOTF criteria were created with the specific intention to compare prevalence rates of child obesity internationally (Cole *et al.*, 2000). The IOTF criteria are based on the largest reference population of all the BMI based measures, consisting of over 190 000 children aged 0-25 in 6 international data sets (survey data from Brazil, Great Britain, Hong Kong, the Netherlands, Singapore, and the United States). Unlike the UK90 and WHO reference criteria, the classification of children's weight status is linked to the adult overweight and obesity. Therefore the IOTF criteria are explicitly linked to 'poor health' outcomes in adulthood, and do not just reflect relative ranking of BMI. Furthermore, the reference population of the IOTF is the most diverse, and includes children from many different ethnic backgrounds, so it is less prone to bias in children from ethnic minority backgrounds (Gatineau and Mathrani, 2011).

As shown in figure 2.1, childhood BMI follows a very distinct pattern with age, which is reasonably consistent across different countries. For the average (median) child, BMI decreases from early infancy (about age 1) until young childhood (about the age of 6) and then increases in a linear fashion into adulthood. The point at which BMI begins to increase again during young childhood is often referred to as the period of 'adiposity rebound' (Cole, 2004; Rolland-Cachera *et al.*, 1984). This pattern of growth is very similar amongst all children, including the very thin and very fat. Adiposity rebound will occur between the ages of 3-8 for most children (Cole, 2004). Thin children tend to experience adiposity rebound at a later age, and children with larger amounts of body fat tend to experience adiposity rebound at a younger age (Rolland-Cachera *et al.*, 1984).

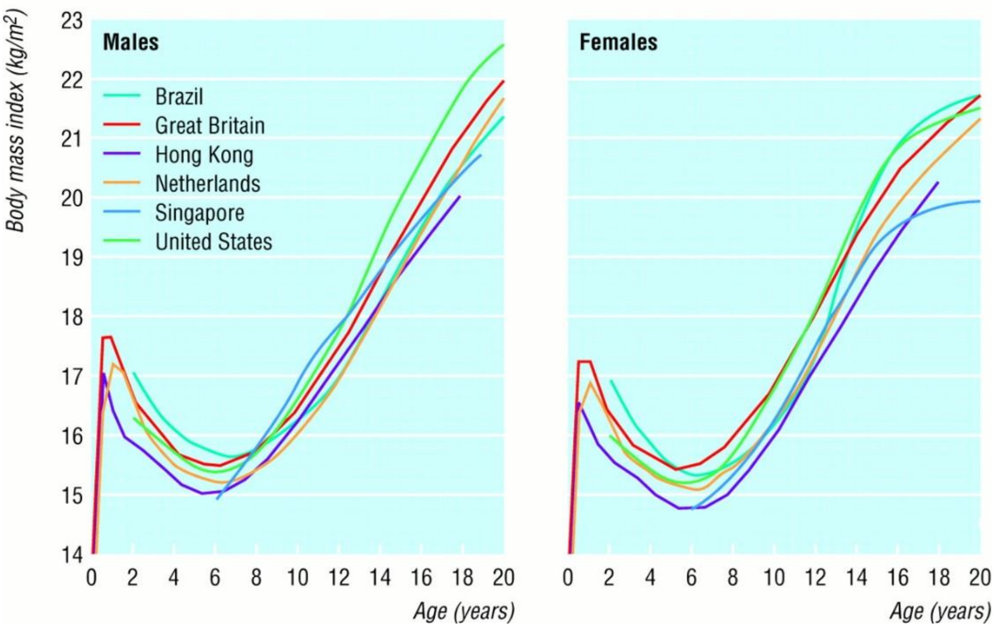
The BMI based measures of child obesity, including the IOTF criteria, take this pattern of growth into account by comparing children's measured BMI to children of the same age in the reference population. What is also depicted in figure 2.1 is the slight differences in children's BMI by sex. For example, on average up to the age of about 9 to 10 years old females BMI's tend to be slightly lower than males BMI's. These sex specific differences are accounted for in determining the BMI cut off values for overweight and obesity for a given age. The cut off criteria shown in figure 2.2 depict the measure of childhood weight status used throughout this thesis. These are overlaid onto the distribution of BMI by age and sex in the MCS cohort in figure 2.3 to provide the reader with a sense of how these criteria work, and how they relate to the distribution of BMI.

The creators of the IOTF criteria suggest that the overweight criteria are more useful than the obesity criteria because they are more sensitive[12] (Cole *et al.*, 2000). This has been shown using comparisons of the IOTF criteria for overweight and obesity to more direct measures of body fat (Reilly, Dorosty and Emmett, 2000). The threshold at which a child is classified as obese for the IOTF criteria is very high. Whilst the IOTF obese criteria are a very specific measure, meaning that all children who are classified as obese are also classified as obese based on more direct measures of body fat, the IOTF criteria for obesity does not identify all children who should be classified as obese. Therefore some children who are obese based on body fat measures are not classified as obese with the IOTF criteria. Furthermore the IOTF criteria for obesity are differentially sensitive in male and female children (Reilly, Dorosty and Emmett, 2000; Wickramasinghe *et al.*, 2009). This can create a significant difference in the prevalence of obesity for male and female’s that is driven purely by the differential sensitivity of the measure.

The IOTF criteria for overweight includes the obese category. The IOTF criteria for overweight represent more extreme levels of excessive weight than other BMI based measures of childhood overweight available (Cole *et al.*, 2000). They correspond to the $88^{th-}90^{th}$ percentile of BMI in the UK90 criteria, well above the overweight (85^{th} percentile) cut off used with the UK90 criteria for population studies (Cole *et al.*, 2000). Again, compared to more direct measures of body fat, the IOTF criteria for overweight are very specific, meaning they are very unlikely to classify somebody as overweight if they have healthy levels of body fat (Leal, da Costa and Altenburg de Assis, 2013; O'Neill *et al.*, 2007; Zimmermann *et al.*, 2004). But because the cut off value is lower, this measure is also more sensitive to selecting children with unhealthy levels of body fat. Therefore the IOTF criteria for overweight are utilised throughout this thesis.

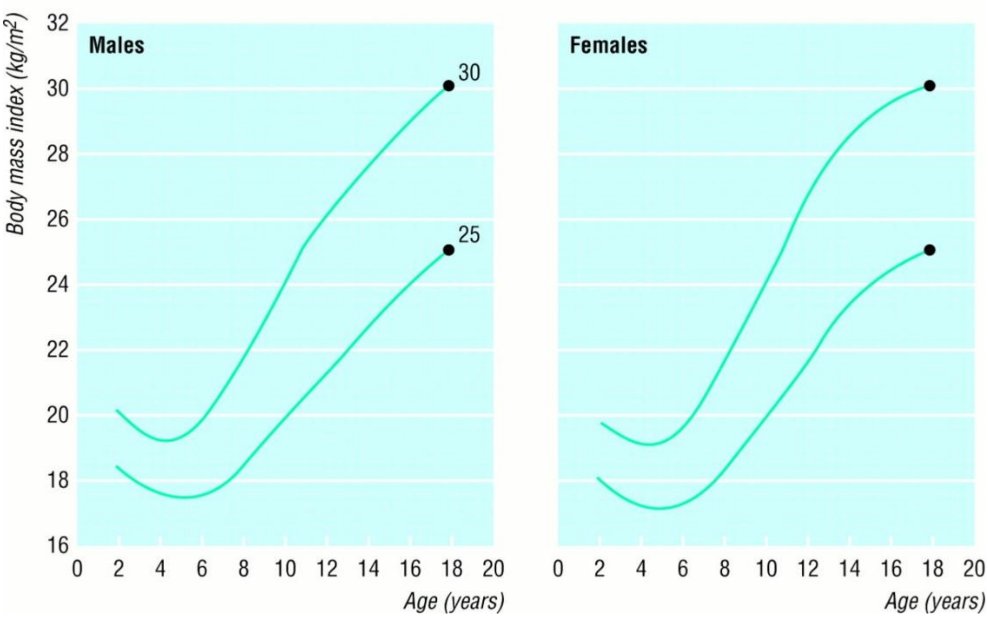
[12] Sensitivity refers to the proportion that are correctly identified as obese/overweight with the IOTF criteria, given that they are actually obese/overweight. Specificity refers to the proportion of negatives which are correctly identified, i.e children who are not obese/overweight are not categorised as obese/overweight by the IOTF measure (Parikh *et al.*, 2008).

Figure 2.1. Median Body Mass index by age and sex in the six nationally representative data sets (Cole et al., 2000)



Source : Cole T J et al. BMJ 2000;320:1240

Figure 2.2. International cut off points for body mass index by sex for overweight and obesity, passing through body mass index 25 and 30 kg/m2 at age 18

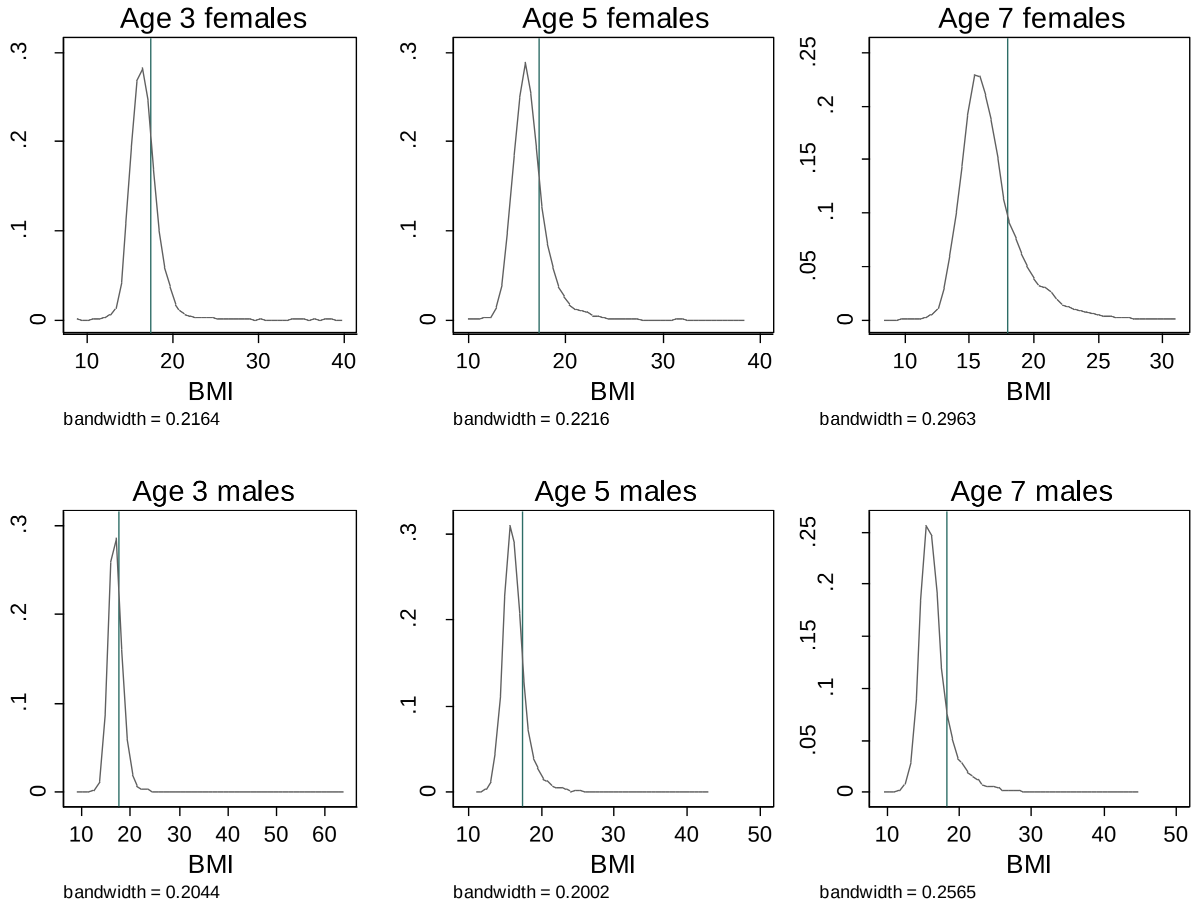


Source : Cole T J et al. BMJ 2000;320:1240

Age and sex specific BMI cut off values for overweight and obesity were provided for every 6 months of children's age from ages of 2 – 18, these are shown in table 2.1 (Cole *et al.*, 2000). The relevant age appropriate cut off values were calculated using linear interpolation of the IOTF half year BMI values, based on the sex and age of the children at measurement of height and weight. Linear interpolation involves calculating the values that constitute the slope of a straight line between the cut off values given at each six month period.

BMI is often the only viable measure of child adiposity available, as height and weight data are often routinely collected in large scale representative social surveys. BMI may not be a direct measure of body fat, but for younger children, and for children who have higher BMI's, it appears that BMI correlates very highly with body fat percentage and fat mass, as determined by DEXA.

Figure 2.3. How the IOTF criteria relate to the distribution of BMI in the MCS by approximate age and sex



Kernel density estimates estimated with the epanechnikov kernel

Table 2.1. The BMI cut of values for the IOTF criteria for overweight and obesity

	Overweight		Obese	
Age	Male	Female	Male	Female
2	18.4	18	20.1	20.1
2.5	18.1	17.8	19.8	19.5
3	17.9	17.6	19.6	19.4
3.5	17.7	17.4	19.4	19.2
4	17.6	17.3	19.3	19.1
4.5	17.5	17.2	19.3	19.1
5	17.4	17.1	19.3	19.2
5.5	17.5	17.2	19.5	19.3
6	17.6	17.3	19.8	19.7
6.5	17.7	17.5	20.2	20.1
7	17.9	17.8	20.6	20.5
7.5	18.2	18	21.1	21
8	18.4	18.3	21.6	21.6
8.5	18.8	18.7	22.2	22.2
9	19.1	19.1	22.8	22.8
9.5	19.5	19.5	23.4	23.5
10	19.8	19.9	24	24.1
10.5	20.2	20.3	24.6	24.8
11	20.6	20.7	25.1	25.4
11.5	20.9	21.2	25.6	26.1
12	21.2	21.7	26	26.7
12.5	21.6	22.1	26.4	27.2
13	21.9	22.6	26.8	27.8
13.5	22.3	23	27.2	28.2
14	22.6	23.3	27.6	28.6
14.5	23	23.7	28	28.9
15	23.3	23.9	28.3	29.1
15.5	23.6	24.2	28.6	29.3
16	23.9	24.4	28.9	29.4
16.5	24.2	24.5	29.1	29.6
17	24.5	24.7	29.4	29.7
17.5	24.7	24.8	29.7	29.8
18	25	25	30	30

Chapter 3: An introduction to the Millennium Cohort Study

The Millennium Cohort Study is used in all empirical chapters in this thesis. Therefore the sampling design will be described here in order to avoid repetition. I start by outlining the sample design of the MCS and follow this by explaining how the complex design of the MCS survey is taken into account in the analyses to follow.

The design of the Millennium Cohort Study

As previously mentioned, the main data source in this thesis is the Millennium Cohort Study (MCS), a longitudinal, large scale survey of nearly 19,000 babies born (mostly) in 2000-2001 and their families in all four UK countries. The sample was drawn from all live births in the UK over a specific time period, where children were alive and living in the UK at 9 months of age and were eligible to receive child benefit. In England and Wales children were included if their birthdays were between 1st September 2000 and 31st August 2001. In Scotland and Northern Ireland this period started on 23rd November 2000 and ended on the 11th January 2002 (Plewis, 2007a). At the time of writing, Data has been collected for each child at approximately age 9 months (sweep 1), 3 years (sweep 2), 5 years (sweep 3), 7 years (sweep 4) and 11 years old (sweep 5). Sweep 6 is scheduled for 2015 when children are approximately 14 years old. There were 18, 818 children with productive responses at sweep 1 of the MCS, 15, 808 at sweep 2[13], 15, 459 at sweep 3 and 14, 043 at sweep 4 (Hansen *et al.*, 2010). This thesis focusses mainly on sweep 4 of data collection.

The overall goal of the MCS sampling design was to represent the total population in the UK, as well as to have sufficient numbers of sub-groups including children from disadvantaged backgrounds, children from non-white ethnic backgrounds and children within the smaller countries of the UK. Probability methods of selection as well as stratification and clustering were used to achieve this goal. 398 electoral wards were randomly selected in the UK. These wards are the primary sampling units (PSU). The wards were stratified into non-disadvantaged, disadvantaged and ethnic wards in England and into non-disadvantaged and disadvantaged wards in Wales, Scotland and

[13] There were 692 'new families' introduced into the sample in the second sweep. These are families who were living at an eligible address at the first sweep, but were not identified in the Child benefit register until after the first sweep of data collection. This was due to a time lapse in the families moving into an eligible ward and the DWP updating their address list.

Northern Ireland. This provides 9 strata in total, three strata in England and two each from Wales, Scotland and Northern Ireland.

The wards were stratified because no individual level family disadvantage or ethnic background information was available for the population at the time of the survey design. In England the ethnic stratum included children living in wards in which at least 30% of their population were from minority ethnic groups. The disadvantaged strata included children living in the poorest 25% of wards according to the ward based Child Poverty Index for England and Wales. The wards were classified as disadvantaged only if they did not meet the criteria for the ethnic strata. The non-disadvantaged stratum included children living in wards that were not classified as ethnic or disadvantaged (Hansen *et al.*, 2010).

The MCS sample employed disproportionate sampling to ensure adequate representation of the smaller countries of the UK and ethnic minority groups. Children living in disadvantaged wards, ethnic wards and wards within Wales, Scotland and Northern Ireland were oversampled to allow for sufficient numbers of children in these sub-groups for statistical analyses.

Children were selected into the sample at 9 months of age. Eligible children were identified from the child benefit[14] register held at the Department for Work and Pensions (DWP). The DWP removed some ‘sensitive cases’ from the register. Of the cases identified by the DWP for inclusion in the sample, 89% were issued to the field. The remaining 11% were not included due to either opting out, moving out of an MCS ward prior to the child being 9 months, or because these were the ‘sensitive cases’ (Plewis, 2007a).

How to analyse the MCS

To take into account the complex sampling design, data are analysed using the ‘svy’ commands in Stata version 13 (StataCorp, 2013). The sample has non independent observations, i.e observations are clustered within electoral wards. Ignoring the sampling design would result in underestimating the standard errors making

[14] Child benefit is a universal provision which is taken up by an estimated 97% of families with children.

significance tests unreliable. The oversampling of certain subgroups means that the sample is not representative if treated as a simple random sample, because the sample contains a higher proportion of children from disadvantaged and ethnic backgrounds, as well as children from Wales, Scotland and Northern Ireland.

There are three key pieces of information needed to obtain correct point estimates and standard errors when analysing the MCS data: The PSU identifier, the strata identifier and the inverse probability weights. The PSU identifier, which identifies the electoral wards is contained in variable ‘sptn00’. The strata identifier is contained in variable ‘pttype2’. There are several probability weights provided with the data. The probability weight which adjusts observations to make them nationally representative of the UK is ‘weight2’. This corrects for the inverse probability of selection into the MCS sample. This accounts for the oversampling of subgroups by adjusting the influence given to each response and provides estimates that reflect a sample which is representative of the UK population. The probability weights are not only used in the calculation of standard errors but they are also used to obtain correct point estimates.

It is also important to account for non-response when analysing the MCS. There are four main types of non-response: i) Attrition, is a permanent loss of cohort members from the sample. ii) Sweep non-response is a temporary loss of cohort members from the sample, meaning that cohort members return to the sample at least once after non response in at least one sweep. iii) Unit non response is when cohort members do not participate in the study from the outset and iv) item non-response is when cohort members do not respond to all items on the survey. Non-response is a concern because people who do not respond are likely to have different characteristics from those that are included in the study. Therefore non-response is not random. This means that the sample no longer represents the population it is supposed to, i.e. the sample is biased.

To try and account for attrition and sweep non-response probability weights have also been created which adjust for the inverse probability of non-response based on the observable characteristics of the cohort members, as well as the probability of selection into the sample in the first place. To create these non-response weights a list of observable characteristics that explain the non-response in the specific sweep of interest are included as covariates in a regression which is used to predict non-response in that sweep (Plewis, 2007b). The inverse of this provides a probability for response. This sweep specific probability of response is then multiplied by the probability of being selected into the sample in the first instance ‘weight2’. For the UK representative weight at sweep 4, this information in contained in the variable ‘dovwt2’ (Jones and Ketende, 2010; Plewis, 2007b). All analyses presented in the empirical chapters use the ‘svy’ commands where possible.

The following chapter marks the start of the empirical chapters for this thesis. Within chapter four I consider whether socioeconomic inequalities in childhood overweight and obesity have widened over time. As well as using information from the MCS, I use information from two other British birth cohorts to compare socioeconomic disparities in child overweight and obesity over time.

Chapter 4: Overweight and obesity Trends across Three British Birth Cohorts: The Influence of Socioeconomic Factors.

This chapter is concerned with whether socioeconomic inequalities in child overweight and obesity have changed over time with the increasing prevalence of child overweight and obesity. Previous research points to a widening in socioeconomic inequalities over time, but, as will be discussed, this research is limited. There is little known about obesity trends for the most advantaged and disadvantaged socioeconomic groups. There is also little known about different indicators of SES and whether different indicators of SES tell a different story about change over time. This chapter contributes to the literature by addressing both of these issues.

Current literature and research questions

Whether or not the relationship between familial socioeconomic status and child obesity has varied historically over time is not well understood. Very few studies have focussed on how socioeconomic inequalities in child obesity have changed over time. The few studies which do exist suggest that socioeconomic inequalities in child obesity/ child overweight have widened with time, and that the socioeconomic groups became more disparate post 1999 (Stamatakis *et al.*, 2005; Stamatakis, Wardle and Cole, 2010; Stamatakis *et al.*, 2010). These studies suggest that the prevalence of child overweight and obesity has increased to a greater extent for those in more disadvantaged socioeconomic groups. However, there is not complete consensus on this finding, as one study found that whilst socioeconomic inequalities in the prevalence of child obesity were marked, they had not widened between 1997 and 2004 (White *et al.*, 2007).

The current evidence in the UK for whether or not inequalities in childhood adiposity have increased over time is limited in several ways. Firstly, there is barely any research considering the longer term (pre 1995) relationship between SES and child obesity (Stamatakis *et al.*, 2005). This is problematic because inequalities in child overweight prior to, or during the inception of, the child obesity epidemic cannot be identified. This information is potentially important for understanding the origin of socioeconomic inequalities in child obesity. Therefore, this paper considers the trends in child obesity spanning four decades.

Secondly, the measurement of socioeconomic status (SES) has been inadequate. SES has been measured using binary measures of social class or income (Stamatakis *et al.*, 2005; White *et al.*, 2007), or composite measures of income and class have been divided into only three groups due to sample size restrictions (Stamatakis, Wardle and Cole, 2010). This provides a blunt measure of SES which groups together families with very different socioeconomic circumstances. This is problematic because it makes it more difficult to detect socioeconomic inequalities and Impossible to detect distinct patterns in the most advantaged/disadvantaged social groups. This paper will therefore use measures of socioeconomic position with several categories. This means that trends in the most advantaged and disadvantaged groups can be considered. The measurement of SES has also been inadequate because, for the most part, the different aspects of socioeconomic status have not been considered. Trends are generally only considered for one aspect of SES. Yet, as discussed in chapter 1, the different aspects of SES may work through different mechanisms. Therefore we might expect differential changes over time.

Thirdly, the current evidence for widening inequalities in child obesity in the UK is based, at least in part, on analysis of the Health Survey for England (HSE) (Robinson, Craig and Bridges, 2013), with most of the evidence for the longer term trends derived from the National Study of Health and Growth (NSHG) (Rona, 1995). Whilst these are valuable resources, the sample sizes of children are small and so it is necessary to group children into wide age bands for analysis. There are other, larger, data sources covering similar time periods which have not been used. The current literature has shown similar findings within the same data sources, but this paper will contribute to the literature by considering whether the same findings hold with different data sources.

The aim of this paper is to provide further evidence on the temporal relationship between socioeconomic status and child obesity, by using information from different data sources and more fine grained measures of socioeconomic status. This will be, to the best of the author's knowledge, the first paper to compare child overweight and obesity prevalence rates across three British birth cohort studies. Firstly I present evidence on the overall long terms trends in child obesity, and then I present evidence for the socioeconomic inequalities within these trends using different measures of socioeconomic status and relative advantage.

Data Sources

This analysis uses data from the National Child Development Survey (NCDS) (Shepherd, 1995), the British Cohort Survey (BCS70) (Butler, Despotidou and Shepherd, 1980; Goodman and Butler, 1986), and the Millennium cohort survey (MCS) (Hansen *et al.*, 2010). The NCDS, BCS and MCS are continuing, multidisciplinary longitudinal studies providing information on individual's development in a range of areas throughout the life course. The birth cohort studies in the UK are unique in that they are the richest and longest running in the world.

The NCDS sample was derived by taking all births in England, Scotland and Wales during a week in spring in 1958 (17, 414 in total). During childhood, information was collected when children were approximately age 7, age 11 and age 16 (Power and Elliott, 2006). The BCS70 sample was derived by taking all births in England, Scotland, Wales and Northern Ireland during a week in spring in 1970 (17, 198 in total). During childhood, information was collected at approximately age 5, 10 and 16 (Plewis *et al.*, 2004). The MCS sample was described in chapter three. Currently the MCS has collected information on children at approximately 9months, 3 years, 5 years, 7 years and 11 years[15].

The large sample sizes, the richness of the data collected, and the quality of the data documentation make these data an ideal candidate for secondary analysis. For the descriptive analysis section, all available data from the NCDS, BCS70 and MCS will be utilised where height and weight data are available: NCDS approximately age 7 (13,296), age 11 (12, 499), and age 16 (11,040), BCS70 approximately age 10 (12160) and age 16 (5723), MCS approximately age 3 (14518), age 5 (15195), and age 7

[15] At the time of writing the age 11 data was not available to include in this analysis.

(13,813). However for the main results section only data from the NCDS and MCS at approximately age 7 and the BCS70 at approximately age 10 will be utilised. The characteristics of the sample are described in table 4.1[16]. As shown in table 4.1 the sample size is similar in three birth cohorts, as is the distribution of gender within these samples. The social class composition of the cohorts has changed over time as has the proportion of parents staying past compulsory education.

There are age effects in child overweight and obesity, whereby children of different ages have different probabilities of being overweight or obese (Chinn and Rona, 2001; Observatory, 2012), therefore the sample was restricted to one time point from the different data sets where children were of similar ages (age 7-10). This creates a more homogenous sample group so that the observed changes over time can be attributed to changes over time and not differences in biological age. Another advantage to using only responses from children between ages 7-10 is that attrition from the NCDS and BCS cohorts tends to be lower when children are younger[17]. Whilst comparing children of age 7 to those of age 10 is not ideal, BMI for children age 5 in the BCS cohort cannot be calculated due to no measures of weight being available. Also an age band of 7-10 is much smaller than age bands used in the majority of studies considering trends over time, whereby age grouping which span 5 years or more are commonly used (Chinn and Rona, 2001; Stamatakis *et al.*, 2005; Stamatakis, Wardle and Cole, 2010; Stamatakis *et al.*, 2010).

In chapter 2 I described in more detail how the relationship between children's BMI and fat mass changed with age. The IOTF criteria do try to account for differences in the relationship between BMI and fat mass as children age. This is because children's BMI measurements are compared to a large reference group of children of the same age and sex as themselves.

[16] Comparisons of the characteristics of responders and non-responders in the NCDS and BCS cohorts are shown in the appendix.
[17] *Furthermore there are problems with the BCS70 sample at age 16. Due to industrial action by teachers between 1984 - 1986, data collection was delayed and fewer cohort members took part in this sweep. Only 6143 cohort members had responses to the medical examination form (the instrument used to collect Height & Weight data). An analysis of response bias for the BCS70 Sixteen-year Follow-up found that males and those from more disadvantaged social backgrounds were underrepresented (Goodman and Butler, 1986: See appendix 5) Although this bias is small.*

Table 4.1. Characteristics of the sample for the main analysis.

	NCDS /1965	BCS70 /1980	MCS/ 2007
n	13296	12160	13813
Female (%)	48	49	50
Overweight (%)	9	8	20
Mean BMI	16	17	17
BMI Range	(9-29)	(10-31)	(10-45)
RGSC (%)			
(I) Professional occupations	5	6	9
(II) Managerial and technical occupations	15	24	39
(III)Skilled Occupations	56	53	38
(IV) Partly-skilled occupations	18	12	12
(V) Unskilled occupations	6	4	2
Missing[a] (n)	*509*	*1329*	*1302*
NS-SEC (%)			
(1) Higher managerial and professional occupations	8	11	19
(2) Lower managerial and professional occupations	12	15	34
(3) Intermediate occupations (clerical, sales, service)	9	9	11
(4) Small employers and own account workers	12	13	12
(5) Lower supervisory and technical occupations	16	17	6
(6) Semi-routine occupations	18	14	10
(7) Routine occupations	24	20	9
Missing (n)	*3914*	*1502*	*496*
Education (%)			
Mother stayed at school past minimum	25	31	57
Missing (n)	*453*	*2096*	*539*
Father stayed at school past minimum	24	31	48
Missing (n)	*512*	*2595*	*441*

[a] *As analyses are only bivariate, listwise deletion is applied. For the latent variable modelling missing data was handled using a maximum likelihood approach which is discussed in more detail later.*

Measuring Child weight status

Height and weight measures

In order to compare childhood weight status across the three birth cohorts, measures of height and weight (the key components of BMI) need to have been administered in a similar way within each cohort. All height and weight data were measured by professionals or trained interviewers. This is important as self-reported, or in this case parental report of children's height and weight could result in increased measurement error (Huybrechts *et al.*, 2006). Each of the three British birth cohorts used a standardised protocol for measuring height and weight, so that measurement error could be minimised within each cohort. These protocols differed between each of the birth cohorts and the measurements have become more precise with time.

In the NCDS and BCS, medical officers measured the height and weight of the children, whereas in the MCS trained interviewers measured the children's height and weight. For the NCDS, weight was measured in underclothes to the nearest pound at ages 7 & 11 (approx. 0.5 kilogram) and height was measured without shoes to the nearest centimetre or 0.25 inch. At age 16 in the NCDS weight was measured to the nearest 0.01 kg and height to the nearest cm. There isn't specific information given in the data documentation about the type of weighing scale or height measuring equipment used by the medical officers.

For the BCS70 weight was measured in underclothes to the nearest 0.1kg or 0.25 ounce and height was measured to the nearest 0.1 cm at age 10 and 0.5cm at age 16 (Butler, Despotidou and Shepherd, 1980; Goodman and Butler, 1986). Beam balance scales were used for most of the measurements of weight (70%). Imperial measures were transformed into metric values by members of the data coding team at, what was called, the Social Statistics Research Unit (SSRU), but is now called the Centre for Longitudinal Studies (CLS) (Butler, Despotidou and Shepherd, 1980). For the measurement of height the following instructions were given on the medical examinations forms at ages 10 and 16:

> *Please position the child upright against a flat wall or a door. Encourage him/her to stretch to their full height, keeping heels on the floor. Heels and buttock should be flush against wall or door. Place a hardboard book on the Childs head. Mark the position of the lower edge with a pencil and then measure the height from the ground with a wood or steal measuring rod or steel tape measures. Alternatively, use measuring device on back of weighing machine and observe precautions as above (Butler, 1980, P.10).*

For the MCS cohort, height was measured to the nearest millimetre with a Leicester stadiometer, using a standardised protocol whereby the interviewer stretched the child to their full height with the parent's/guardians assistance. With permission, weight was measured to the nearest 0.1kg whilst children were wearing light, indoor clothing using Tanita HD-305 scales at sweep 2 (GfK NOP) and Tanita BF-522W (body fat) scales at sweep 3 and 4 (Gray, Gatenby and Huang, 2010). Parental assistance was sought in measuring the children to ensure they were standing correctly during height measurements and stood still during weight measurements. Weighing scales were checked for accuracy prior to use by placing a 20kg concrete paving slab on them. Scales which displayed between 59.8 kilograms and 60.2 kilograms inclusive were deemed acceptable, otherwise they were sent for recalibration (Gray *et al.*, 2010).

In all three cohorts, weight was measured whilst the child was wearing light clothing and height was measured without shoes. However, there are differences in the precision of measurement. Height and weight are much less precisely measured in the earlier sweeps of the NCDS cohort and the distribution of height and weight are affected from the rounding up/down of measurements. The distribution of height, weight and BMI in the three cohorts are shown in appendix A**.** Whilst the precision of height and weight differ between the birth cohorts, the protocols for the measurements of height and weight are similar enough to warrant comparison.

IOTF criteria for analysis of trends

The IOTF criteria for overweight, described in detail in chapter 2, are used as the measure of child weight status. The IOTF overweight category includes those that are obese also. There is little information available in the literature on the relative performance of different BMI based measures of child obesity when looking at long term trends in child weight status. However, children are becoming increasingly taller and heavier on average over time (Freedman *et al.*, 2000). BMI provides a measure of weight in kg per metre of height squared. But taller children (and adults) generally have higher BMI's. This may because the BMI calculation does not adequately adjust for height (Bonthuis *et al.*, 2013), or it may actually be because taller children really do have a higher proportion of body fat (Freedman *et al.*, 2004; Metcalf *et al.*, 2011).

The reference populations of the BMI based measures of child obesity consist of data from a specific period of time. The reference population for the IOTF criteria is based on data sets covering three decades from 1963 – 1993 (Cole *et al.*, 2000). But the IOTF criteria are not designed to take into account the increasing trends in height or weight over time, nor are any currently available BMI based measure of child weight status. This may be problematic for the more recently collected height and weight data i.e the measures taken from the MCS. The reference population does not represent a concurrent group of healthy children with which to make comparisons regarding BMI. Although it could be argued that the changing norm in weight of children reflects increases in fatness, and therefore it is appropriate to compare to the 'healthier' or 'slimmer' population of children prior to 1994, the same argument cannot be made for the changes in height. Indeed Freedman (2000) suggested that we must be cautious when using noncurrent reference data to compare growth trends in children (Freedman *et al.*, 2000). Nevertheless, the IOTF criteria have been shown to be highly specific when compared to other more direct measures of body fat on contemporary samples of children (Monasta *et al.*, 2011; O'Neill *et al.*, 2007; Reilly, Dorosty and Emmett, 2000; Wickramasinghe *et al.*, 2009; Zimmermann *et al.*, 2004)

Measuring Socioeconomic status

The measurement of socioeconomic status is an important component of this chapter. This particular sub-section includes the description of all measures of socioeconomic status that were considered and provides a detailed account of each measure. Measures of socioeconomic status are categorised as relational (social class and status) and attribution (education and income). Firstly measures of social class are discussed, followed by parental education, income and a latent measure of socioeconomic status that is based upon several variables. Full details of the latent variable modelling approach are provided including a description of the model fit criteria.

Socioeconomic status is a broad term which refers to one's social and economic position in relation to others. In other words it reflects one's place within a hierarchical social structure. Goldthrorpe (2012) highlights two broad approaches to measuring socioeconomic inequalities (Goldthorpe, 2012). The first is the stance generally taken by the sociologists, in which inequality arises from social relationships in which people are more or less advantaged. Social class and status are examples of measures which reflect the relational conceptualisation of socioeconomic status. The second approach is generally taken by economists and epidemiologists, who tend to measure socioeconomic inequalities in an attribution sense, whereby inequality is determined by having more or less of something, such as income, education or assets (Goldthorpe, 2012).

The birth cohort data sets have several different indicators of parental socioeconomic status based on both the relational conceptualisation of socioeconomic inequalities and the attribution conceptualisation of socioeconomic inequalities. The measures that are available, or can be constructed in all three cohorts using the data provided are described below. Firstly measures based on occupation are considered. The Registrar General's Social Class (RGSC) (Szreter, 1984) and the National Statistics Socioeconomic Classifications (NS-SEC) (2005; Pevalin and Rose, 2002) are available during childhood. Secondly attribution measures are considered and these measures are combined to create a latent measure of socioeconomic status. Different measures of SES are used with the aim of giving a more comprehensive account of the socioeconomic trends in child obesity over time.

Relational measures of social position

There are two measures of social position available in the birth cohorts which attempt to measure the relational conceptualisation of socioeconomic inequalities. The National Statistics Socioeconomic Classification (NS-SEC) and the Registrar Generals Social Classifications (RGSC). The NS-SEC is a measure of social class. The theoretical grounding for the NS-SEC measure is largely based on the work of John Goldthorpe and Robert Erikson (Erikson and Goldthorpe, 1992). The classifications are based on both theoretical and empirical work. Questions covering a wide range of criteria including job security, payment intervals, pay increases, control, and autonomy were asked to 60,000 citizens in the UK labour force survey 1997. The responses were used to help create the classifications of the NS-SEC (Rose, Pevalin and O'Reilly, 2005).

The analytic version of the NS-SEC has 8 classes, the measure I will be using in the birth cohorts has only 7 classes and does not include the 'never worked and long term unemployed category'. The NS-SEC classes are given below:

NS-SEC 8 category	NS-SEC 3 category
1. Higher managerial and professional occupations 2. Lower managerial and professional occupations	1. Higher managerial, administrative and professional occupations
3. Intermediate occupations (clerical, sales, service) 4. Small employers and own account workers	2. Intermediate occupations
5. Lower supervisory and technical occupations 6. Semi-routine occupations 7. Routine occupations	3. Routine and manual occupations
8. Never worked and long-term unemployed	

The 8 category version of the NS-SEC is not ordinal largely because of the emergence of own account workers, i.e a move from NS-SEC 4 to NS-SEC 3 does not necessarily relate to an "increase" in social class (Rose, Pevalin and O'Reilly, 2005). The 3 category NS-SEC can be conceptualised as ordinal, but using this three tier categorisation results in a loss of information as the self-employed are merged with those employed in intermediate occupations. The NS-SEC measure initially differentiates people by whether they are self-employed, an employer or an employee. Employees are then further segregated by their employment contracts. There are two main types of contracts: service contracts and labour contracts. With the service contract productivity is not (and potentially cannot be) directly monitored. Therefore rewards such as high levels of job security, salary increments, and a progressive career are offered as incentives for good and productive employees. In contrast, employees with a labour contract perform work that is more easily monitored, they have little autonomy, and tend to be more closely supervised and restricted in their patterns of work. There are contracts which combine both aspects of the service and labour contract, these are sometimes referred to as 'intermediate' contracts. The employment relations and conditions of occupations are believed to be central to showing the structure of socioeconomic positions in modern societies, and helping to explain variations in social behaviour and other social phenomena (Bartley, 2004; Rose, 2005; Rose, Pevalin and O'Reilly, 2005) .

The NS-SEC measure is available as a measure of parental social class in all three birth cohorts during the cohort member's childhood. Unfortunately, the NS-SEC is only available in the NCDS when children were aged 11 and 16, but not when they were aged 7. I am specifically interested in child outcomes in the NCDS at age 7. Therefore to use the NS-SEC an assumption has to be made that the vast majority of fathers did not change social classification between 1965 (when children were aged 7) and 1969. In order to explore whether this assumption can be made the RGSC measure was utilised. This measure is explained in detail in the next section. The RGSC was measured at each sweep, and whilst the RGSC and NS-SEC are very different measures conceptually, they are both derived from occupation, so it is reasonable to assume that stability in one measure would, to some extent, reflect stability in the other, as this would reflect occupational stability over time.

The correlation[18] between RGSC of the father in 1965 (when the children were aged 7) and 1969 (when the children were aged 11) was 0.74, with cross tabulations showing a reasonably high degree of stability in social classifications. Whilst there was a high degree of stability in the RGSC between 1965 and 1969, there is variation over time. This variation is, in most instances, only a movement of 1 RGSC defined social class position. Therefore by using the NS-SEC measured at age 11, I will be measuring the association between social class and child obesity with less precision as some fathers will have changed social class over time. An alternative option would be to use the RGSC measure as the main measure of social class.

The Registrar Generals Social Classification

The RGSC measure is provided in the NCDS and BCS cohorts for all ages during childhood. The RGSC was constructed from the NS-SEC in the MCS cohort using continuity guidelines provided by the Office for National Statistics (Rose, Pevalin and O'Reilly, 2005). The RGSC is a hierarchical measure. Between 1921 and 1971, the

[18] Polychoric correlation coefficients are a measure of association for ordinal variables which rests upon an assumption of an underlying joint continuous distribution.

classification of occupation by the RGSC was based on occupational status & prestige, but in 1980 the classification system was changed to greater reflect occupational skill (Brewer, 1986; Szreter, 1984). Although no empirical work was ever done to test the extent to which the RGSC reflected occupational prestige or occupational skill. Despite the change in the conceptual underpinning of the RGSC, comparisons using the 1970 classifications and the 1980 classifications yielded almost identical results (Brewer, 1986).

The measure was provided in the NCDS and BCS when children were age 7 and 10 respectively. It was not measured in the MCS, but was constructed from the full version of the NS-SEC when children were aged 7 (Rose, Pevalin and O'Reilly, 2005). The Registrar Generals Social classification has 5 major classes. Class III is sometimes sub-divided into skilled manual (IIIM) and non-manual (IIINM), but this is not the case in the present analysis. The RGSC classifications are given below:

I. Professional occupations
II. Managerial and technical occupations
III. Skilled occupations
IV. Partly-skilled occupations
V. Unskilled occupations

The RGSC originated at the end of the 19th century and therefore describes an industrial society and economy that is no longer relevant today. Although the RGSC was updated every census, the structure and beliefs imbedded in it, such as the manual/non manual divide no longer reflect social realities (Szreter, 1984). Equally there is no clear theoretical underpinning in the RGSC measure and there is very little empirical work supporting it (Jones and Cameron, 1984). However, numerous studies show that it does group people based on differential health risks, see for example (Donkin, Goldblatt and Lynch, 2002; Edwards *et al.*, 2006; Weightman *et al.*, 2012). Therefore even though it lacks a theoretical underpinning, and is potentially outdated, it continues to perform well (Bartley, 2004).

There are also problems specific to the use of RGSC to measure social position in different time periods. The decline in 'manual' occupations and the rise in managerial, professional & clerical occupations over time suggest that it is less appropriate to classify ones social position based on occupational skill level alone (Gallie, 2001), therefore it is likely that the RGSC is a less effective measure of social position with the most recent birth cohort, the MCS. The NS-SEC however, does not attempt to group people by occupational skill, rather it seeks to measure social relations within the labour market i.e whether people are employers, self-employed, or employees. Therefore the NS-SEC measure of social class is less likely to be biased by changes in occupations over time.

Whilst it would be preferable to use the NS-SEC as a measure of social position, the limitations regarding its retrospective measurement mean that I will also be using the RGSC. Given that social class is dependent upon labour market participation, and because labour market participation of women has changed so dramatically over the time period under study (1965-2007) (Goldin, 1991), and because the proportion of single parent families has increased over time (Finch, 2002), the main results are presented for the NS-SEC or RGSC of the father in the NCDS and BCS cohort, and the highest NS-SEC or RGSC of either parent in the MCS cohort.

Attribution measures of parental socioeconomic status

The datasets were searched for comparable measures that may reflect social advantage or disadvantage in some way, which were present in all three cohorts. This lead to several possible indicators of relative advantage/ disadvantage: parental education, measured by whether parents were at school past minimum leaving age, housing tenure, and a measure of house crowding created by dividing the number of people residing in the household divided by the number of rooms in the family home excluding kitchens and bathrooms. All these variables are measured at the same time as the height and weight measures were taken from the children. Table 4.2 describes the comparable variables available in each cohort.

Parental education

Parental education is measured in all three birth cohorts. In order to create a comparable measure of parental education, a binary variable for both mothers and father's education based on age of leaving school was created. This was the only measure of parental education available in the NCDS when the cohort members were children. More detailed measures of parental education are available in the BCS and MCS cohorts, but for comparison purposes the most similar measures have to be utilised. For the BCS cohort, parents were classified as staying past the minimum school leaving age based on if they reported any years of education past the age of 15. This was cross checked with the age at which they reported leaving school. For the MCS cohort, parents were classified as staying past the minimum age if they stayed in full time continuous education past the age of 16. Staying past the minimum school leaving age is most likely a better indicator of socioeconomic status in the NCDS and BCS cohorts, than for parents in the MCS cohort.

There was a large increase in the number of people staying at school past the minimum school leaving age, driven by expansion in the higher education system and the increasing requirement for qualifications from employers (Greenaway and Haynes, 2003). Therefore those who choose to stay at school in the MCS cohort are likely a more heterogeneous group of people than those who stayed at school in the NCDS and BCS70 cohorts. Therefore Parental education was also considered for those who stayed in full time education past the age of 18 in the MCS cohort (about one quarter of the parents). As shown in table 4.1, roughly one quarter of parents in the NCDS cohort, just under a third of parents in the BCS cohort and half of the parents in the MCS cohort stayed at school past the minimum leaving age.

Income

The decision was made not to include income as an attributional measure of socioeconomic status. This is because the measures of income in the NCDS and BCS could be considered problematic. This issue is discussed in detail by Erikson and Goldthorpe (2010) and Blanden, Gregg and MacMillan (2010) (Blanden, Gregg and Macmillan, 2010; Erikson and Goldthorpe, 2010).

Table 4.2. Comparable attributional measures of socioeconomic status in the three birth cohorts.

	NCDS	BCS70	MCS
Parental Education	Did mother stay past minimum school leaving age? (n537 - sweep 0) Did father stay past minimum school leaving age? (n194 - sweep 1) If yes, at what age did father finish schooling (n195 - sweep1)	Age mother left school (E191 - Sweep 1) Age Father left school (E192 - Sweep 1) Mother's years of fulltime education after school/ after age 15 (E193/E195 - sweep1) Father's years of fulltime education after school/ after age 15 (E194/ E196 - sweep 1)	Age main respondent left fulltime continuous education(amlfte00 - sweep1) Age partner respondent left fulltime continuous education (aplfte00 - sweep 1)
Tenure	Owner, council rented, private rented, rent free, other, don't know, not applicable (n200 - sweep 1)	owned outright, being bought, council rented, private rented unfurnished, private rented furnished, tied to occupation, other (d2 - sweep 2)	own outright, own - mortgage/loan, part rent/part mortgage (shared equity), rent local authority, rent housing association, rent privately, live with parents, live rent free, other (dmroow00 - sweep 4)
House crowding	Number of room excluding bathroom, scullery or kitchen - unless used as a dining room (n201 - sweep 1)	Number of rooms excluding kitchens, bathrooms/toilets and rooms used solely for business or trade purposes (d5 - sweep2). How many of these rooms are bedrooms (d5_2 - sweep 2)	Number of rooms excluding kitchens, bathrooms, toilets and halls and garages (dmroma00 - sweep 4)

The measurement of income in the NCDS is particularly problematic for the comparisons over time. Familial income is not measured in the NCDS until the children are aged 16. This means that the measurement of income would be transitory. Transitory income is a measure of income from one point in time. It is well known that income can fluctuate a great deal within short periods of time, and that this increases the amount of measurement error providing a less precise estimate of actual income (Blanden, Gregg and Macmillan, 2010). Furthermore, the measure of income would be a transitory measure of income taken 9 years after the measurement of child weight status, and therefore will likely not be very informative of the economic resources available to the family at the time of the child's weight status measurement, or the long run economic resources of the family.

A further concern for the measurement of income in the NCDS is that some measurements of income in the NCDS were taken during the 'Three-day Week' when working hours were restricted due to a coal shortage, therefore it is possible that the reported income is that of the three-day week rather than usual weekly income (Blanden, Gregg and Macmillan, 2010). Additionally, there is a large amount of missingness in income in the NCDS, with only 8390 of father's and 6751 of mother's reporting earnings.

Additionally, in both the NCDS and BCS the income data reported by the self-employed has been deemed of low quality (Erikson and Goldthorpe, 2010). The differential construction of income in the three cohorts also makes it more difficult to make comparisons (Erikson and Goldthorpe, 2010). The relationship between income and child overweight and obesity is considered in detail in the next chapter, chapter five, using data form the Millennium Cohort Study only.

Whilst income is not considered here, housing tenure and house crowding provide proxy measures of wealth and poverty[19]. Changes in housing tenure trends over the 1980's suggests that the owner-occupier group is not homogenous across time. However within each cohort, if we compare people within the same time period, those

[19] Wealth here is defined as an abundance of material resources, assets and income. Poverty is defined by not having enough income or material resources to have a standard of living that is considered acceptable within the society in which they live (Bradshaw *et al.*, 2008).

who own their home are still likely to be wealthier than those who do not, even though the relative differences between the home-owners/ non home owners may be smaller in some time periods than others.

Latent measure of SES

Using latent measures of SES has been advocated by some researchers (Mackenbach and Kunst, 1997; Oakes and Rossi, 2003) indeed it has been suggested that SES is conceptualised as a latent construct in the UK (Mackenbach and Kunst, 1997). This is not a consensus view, and there are discipline specific approaches to conceptualising SES. Creating a single latent measure of socioeconomic status will result in a uni-dimensional SES construct, whereby children are ranked based on their parents responses to several different indicators. It could be argued that SES is a multidimensional construct and should not be measured in this way, but there are methodological advantages to doing so, including the specification of measurement error. With latent variable approaches it is possible to partition the variance in a set of variables so that the measurement error can be defined and separated out from the regression estimates.

Defining Latent variables

latent variables are theoretical constructs that cannot be directly observed or directly measured (Byrne, 2011). They are defined by the observable behaviour believed to represent them, under the assumption that the latent variable causes the observed behaviour. Therefore the unobservable latent variable can be indirectly observed through these behaviours (Byrne, 2011). Several observed variables are used to identify a latent variable. The latent variable is then defined by the shared covariance amongst the observed variables (indicators). The variance within the indicators is partitioned into the shared covariance, which defines the latent variable, and the unique variances and covariances among indicators, which become the residual terms on the indicators. The variance that is unique to the indicators can also be defined as the measurement error. This variance partitioning is achieved through a technique known as factor analysis.

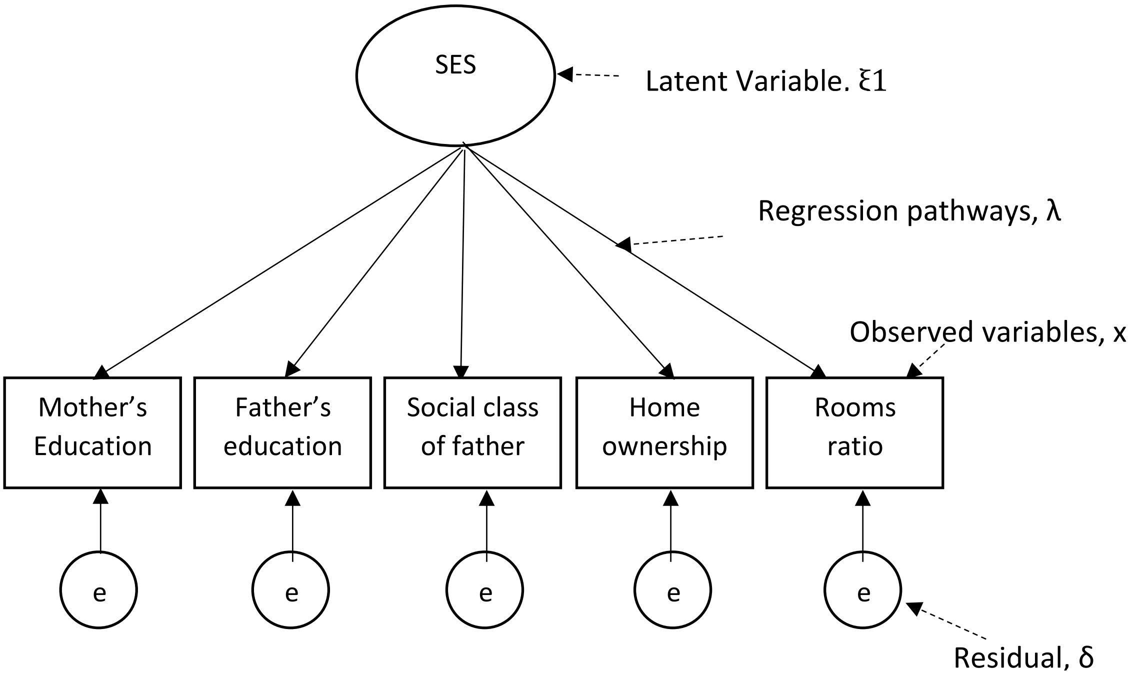
Factor analysis is primarily concerned with the extent to which observed variables are generated by the underlying latent construct (Hurley *et al.*, 1997). Regression paths between the latent variable and the observed indicators are called factor loadings and the strength of these loadings demonstrate the variance in the indicator that is explained by the latent variable (the factor) (Hurley *et al.*, 1997). Confirmatory Factor analysis, also known as 'measurement models' in the language of Structural Equation Modelling (SEM), is used to test hypothesised models to demonstrate how well the relationships we hypothesise between the indicators and latent variables is replicated in the data we have available (Byrne, 2011).

Creating a latent Measure of SES

Composite or latent measures of SES have been widely utilised in health research. In order to maximise the amount of information available to describe social position and to create a relative ranking of social position within each cohort, I tested a single latent variable measurement model to represent social position based upon several indicators: Whether mother stayed at school past the minimum school leaving age, whether father stayed at school past minimum school leaving age, social class, home ownership and the ratio of rooms per person in the household. The assumption behind this model is that socioeconomic status is an abstract concept that cannot be directly observed. We assume that an individual's socioeconomic status determines their scores on the five indicators listed above.

Through capturing the shared covariance of these variables a measure of socioeconomic status can be identified that reflects an individual's relative position in the social hierarchy and potentially their access to collectively desired resources. This latent measure therefore combines both relational and attributional aspects of SES. This model estimates a single latent variable which is defined by partitioning the variance in the observed variables into the shared covariance and the variance which is unique to each indicator (measurement error). The variance in the observed variables that does not co-vary with the variance in the other variables is placed in the residual also known as the error term. This measurement model is depicted in figure 4.1.

Figure 4.1. Measurement model for a single latent construct measuring socioeconomic status.



The measurement model depicted in figure 4.1, can also be specified as an equation. Equation 4.1 demonstrates the factor analysis model and postulates that observed variables, x1…x5, are linear functions of the latent variable, ξ1, plus a residual term, δ. The regression pathways, λ, provide an estimate of the relationship between the latent variable and the manifest variable.

Equation 4.1

$$\begin{pmatrix} x1 \\ x2 \\ x3 \\ x4 \\ x5 \end{pmatrix} = \begin{pmatrix} \lambda 11 \\ \lambda 21 \\ \lambda 31 \\ \lambda 41 \\ \lambda 51 \end{pmatrix} \times \xi 1 + \begin{pmatrix} \delta 1 \\ \delta 2 \\ \delta 3 \\ \delta 4 \\ \delta 5 \end{pmatrix}$$

In order to obtain the estimates for equation 4.1, we must estimate the model implied variance-covariance matrix. The model implied covariance matrix is defined in equation 4.2.

Equation 4.2

$$\Sigma = \Lambda \Psi \Lambda' + \Theta$$

Where,

$\sum$ is the model implied variance-covariance matrix
Λ is the matrix of loadings
Λ' is the transpose of the Λ matrix
Ψ is a matrix of variances and covariance's among the latent constructs, (just the variance in CFA)
Θ is the matrix of the residual variances and residual covariance's among the indicators

The greek letter sigma, $\sum$, refers to the variance-covariance matrix that is calculated based on the fixed and estimated parameters contained in the right hand side of the equation. The matrix lambda, Λ, contains the estimated loadings that link the indicators to the underlying construct. The matrix psi, Ψ, in this instance represents the variance of the latent construct $\xi 1$, and the matrix theta, Θ, contains the residual variances and covariances of the indicators.

As confirmatory factor analysis (CFA) is based on covariance, it can only be conducted with data that are continuous, ordinal or binary (Brown, 2006). Therefore housing tenure was recoded into a binary variable indicating home ownership (own outright or mortgage) or not. The NS-SEC was collapsed into three categories, which can be conceptualised as ordinal (Rose, Pevalin and O'Reilly, 2005). These CFA measurement models were fit within each cohort using the same indicators. The level of measurement of the variables was taken into account using polyserial and polychoric correlations between the observed variables, and logit and ordered logit regression where appropriate. Factor scores were generated and these were divided into quintiles (equal fifths). Full and detailed information regarding the construction of the latent variable and the assessment of model fit is provided in appendix B.

Handling Missing Data

An advantage of using CFA is that missing data can be handled using full information maximum likelihood (FIML) estimation. Enders and Bandalos (2001) describe the FIML approach as follows. FIML computes a case-wise likelihood function using only those variables that are observed for case *i*. The case wise likelihood of the observed data is obtained by maximizing the function. This is shown in equation 4.3.

Equation 4.3

$$\log L_i = K_i - \frac{1}{2}\log|\Sigma_i| - \frac{1}{2}(x_i - \mu_i)' \Sigma_i^{-1}(x_i - \mu_i)$$

Where, K_i, is a constant that depends on the number of complete data points for case *i*, x_i, is the observed data for case *i*, and μ, and ∑, contain the parameter estimates of the mean vector and covariance matrix, respectively, for the variables that are complete for case *i*. The case-wise likelihood functions are accumulated across the entire sample and are maximised as shown in equation 4.4.

Equation 4.4

$$\log L(\mu, \Sigma) = \sum_{i=1}^{N} \log L_i$$

All available data are utilised during parameter estimation. The FIML algorithm does not impute the missing values, but the borrowing of information from the observed portion of the data is analogous to replacing missing data points with the conditional expectation of the missing value, given the values of the other variables in the model.

The results are unbiased as long as the missing data can be assumed missing completely at random (MCAR) or missing at random (MAR) (Enders and Bandalos, 2001; Newman, 2003). Data can be considered MCAR, if, as the name suggests the missingness is random, but this is rarely the case. If data were MCAR there would be no patterning in the missingness, so that everybody had an equal chance of non-response. Data can be assumed MAR if the missingness is correlated with other variables included in the analysis (Howell, 2012). That is, dependent upon the responses to other variables missingness is random. The other variables in the model provide information about the missingness. Specifically the other variables provide information about the marginal distributions of the incomplete variables. Where the assumptions of MAR are met, the estimates will be unbiased.

A much more difficult scenario occurs when data are missing not at random (MNAR). MNAR exists when even after accounting for all the available observed information, the reason for observations being missing still depends on the unseen observations themselves. Both FIML and multiple imputation will yield biased estimates when data are MNAR. However it has been suggested that when data are MNAR both FIML and multiple imputation will result in less bias than list wise deletion or other methods for dealing with missingness (Enders, 2010; Graham, 2009).

The concept of missing data will be discussed again in chapter 6 when an alternative method for handling missing data (multiple imputation) will be discussed.

Analysis

Data were analysed using Stata version 13 (StataCorp, 2013) and Mplus version 7.1 (Muthén and Muthén, 1998-2012). Firstly I consider the descriptive trends of child overweight for children of all ages in the birth cohorts (ages 3-16). This provides background information on the prevalence of child overweight within each cohort, how this prevalence changed with age, and the patterning by socioeconomic status. Within each cohort I formally test whether socioeconomic inequalities in child overweight have changed as the children age using a random effects model.

The main focus of the analysis is on comparing socioeconomic inequalities in child overweight for children aged 7-10. Within each cohort I create the measures of social class, parental education, and the latent measure of socioeconomic position as described previously. Within the MCS the latent variable is created with information from the clustering, strata and probability weight included in the analysis. I append these data sets to create one data set containing information from all three birth cohorts and a variable indicating cohort membership (1965/1980/2007). A probability weight variable was created that is held at 1 in the NCDS and BCS, and given the value of the "weight2" weight of the MCS[20]. Logistic regression analyses were then used to consider the interaction between the cohort membership variable (the time period), and each measure of socioeconomic status. To test whether the trends varied over

[20] Analyses were run with and without this probability weight, and using the values from the "dovwt2" probability weight to ensure the use of this weight was not driving the results. The use of this probability weight does not change the substantive findings compared to using no weight or compared to using the values of the dovwt2 weight in the MCS.

time by gender a three way interaction is considered between time period, socioeconomic status and gender.

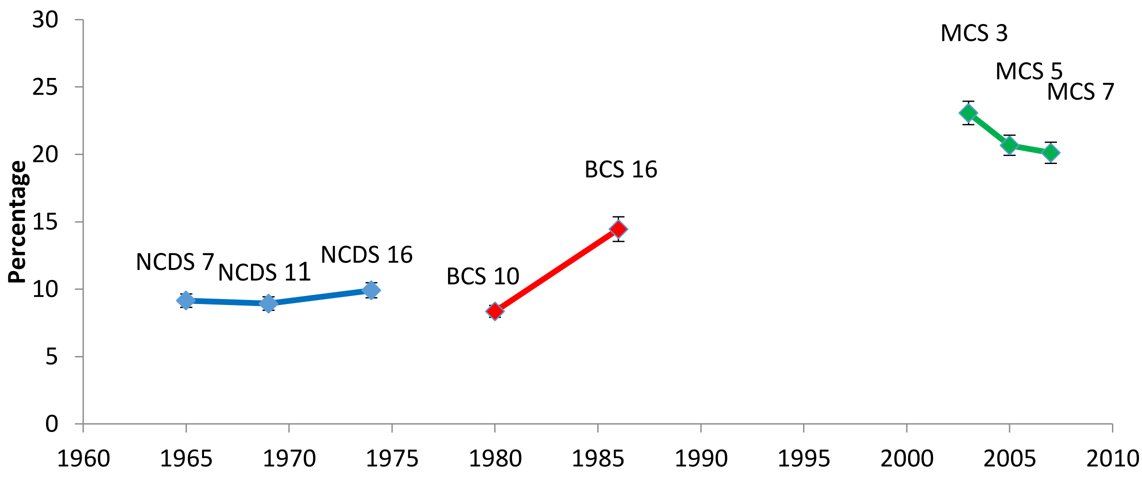
As all analyses presented here are essentially bivariate list wise deletion is applied to the missing data, with the exception of the latent variable model, whereby FIML is applied.

Results: Overall trends in child adiposity

The following section is based on children of all ages within the birth cohorts during childhood. The estimates span a time period of 42 years and include children from approximately ages 3-16. The estimated prevalence of overweight is presented for all available data in the three cohorts during childhood in figure 4.2. As can be seen in figure 4.2 the data that belong to each cohort have been labelled with the cohort name and the approximate age of the children at that sweep of data collection. As expected, this figure shows that the prevalence of childhood overweight has increased over time.

Overweight prevalence is relatively stable in the NCDS cohort as the children age. There are always approximately 9-10% of children in the cohort classified as overweight. The prevalence of overweight increased rapidly in the BCS cohort with age, increasing by 6 percentage points. Therefore a higher proportion of this cohort became overweight with age. For the MCS cohort, the prevalence of overweight is high compared to the NCDS and BCS cohorts, but the prevalence appears to be declining with age. This apparent decline with age may be result from an overestimation of child overweight by the IOTF criteria in very young children (those aged 3) (Monasta *et al.*, 2011), or it could represent a genuine patterning in child overweight status with age – it may be that the prevalence of overweight status declines with age for very young children.

Figure 4.2. Percentage of children classified as overweight (as defined by the IOTF criteria) in the three birth cohorts. Error bars represent 95% confidence intervals.



Figures 4.3 and 4.4 show these overall trends disaggregated by RGSC and by father's NS-SEC respectively. The RGSC and NS-SEC show a similar picture; in the NCDS cohort the prevalence of overweight appears to be increasingly patterned by socioeconomic status with age. The descriptive analysis suggests that between the age 11 measure (1969) and the age 16 measure (1974), the differential socioeconomic groups become more divergent in the prevalence of overweight. The prevalence of child overweight decreases in highest social class groups between age 11 and age 16, whereas the prevalence increases in the lowest social class groups between age 11 and 16. A formal test of the interaction[21] between RGSC and age (chi2(8)= 34.42, $p<0.05$), and NS-SEC and age (chi2(12)=22.02, $p<0.05$) suggested there was a statistically significant change in the relationship between social class and child overweight with age.

For the BCS cohort, the descriptive analysis suggests that the prevalence of child overweight is increasingly patterned by social class with age. All socioeconomic groups experience increases in child overweight between age 10 and 16, however the lowest socioeconomic groups experience the largest increases. Again the changes over time were tested formally with an interaction[22] between RGSC and age (chi2(4)= 5.72, $p>0.05$), and NS-SEC and age (chi2(5)= 5.68, $p>0.05$). These interactions suggested

[21] Interactions between age and SES tested in a Random-effects GLS regression

there were no statistically significant changes in the association between social class and child overweight with age in the BCS cohort.

In the MCS cohort, the descriptive statistics show that socioeconomic differences in overweight are already marked and are more divergent than in the BCS or NCDS, despite the younger age of the cohort members (ages 3,5 & 7). The descriptive statistics indicate that the prevalence of overweight was patterned by social class at each age of measurement in the MCS, but that differences between social class groups are getting larger with age. Interactions between social class and age suggest that the relationship between social class and child overweight did change by a statistically significant amount as children aged in the MCS cohort (RGSC: chi2(8)= 23.65, p<0.05| NS-SEC: chi2(12)= 28.98, p<0.05).

Figure 4.3. The percentage of children classified as overweight disaggregated by father's social class as defined by the Registrar Generals Social classification. Error bars represent 95% confidence intervals.

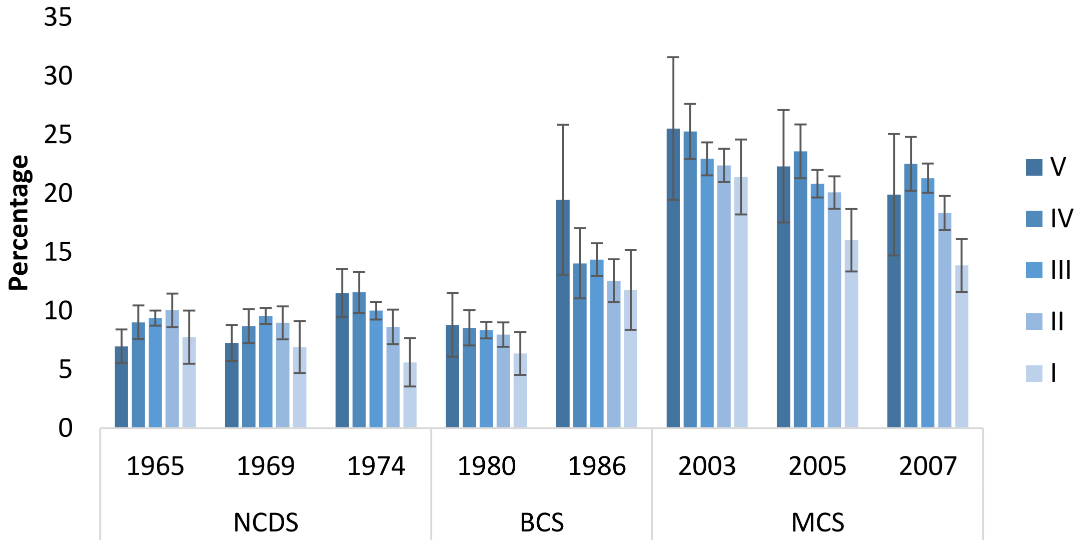
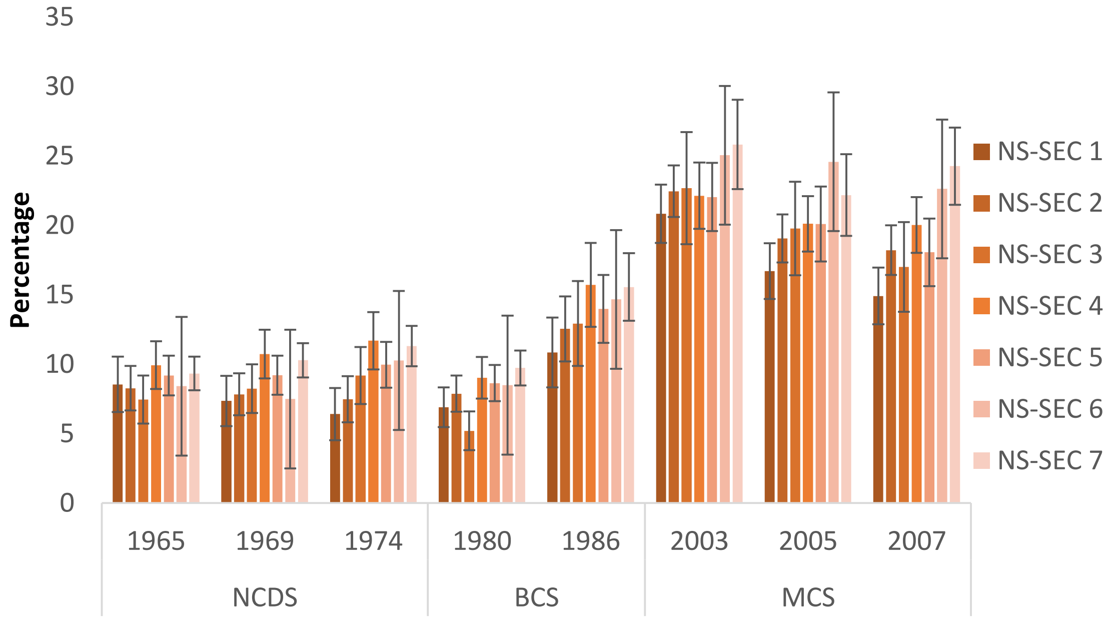


Figure 4.4. The percentage of children classified as overweight disaggregated by father's social class as defined by the National Statistics Socioeconomic Classification. Error bars represent 95% confidence intervals.



Results: Changing socioeconomic inequalities in child obesity for younger children (age 7-10)

The main analysis focusses on children aged between 7 and 10 only. As each data set has children of only one specific age, the cohort now represents a point in time. The NCDS cohort now represents 1965, because children were approximately aged 7 in 1965. The BCS cohort represents 1980, and the MCS cohort represents 2007. The results for the analysis of younger children are presented in table 4.3. Each measure of SES is considered separately, that is the models do not adjust for the other measures of SES.

NS-SEC

The first section of table 4.3 shows the results whereby SES is measured by the NS-SEC. As can be seen in table 4.3, there was a statistically significant difference in child overweight prevalence rates by NS-SEC in 1980 and 2007, but not in 1965. In 1980 children with fathers in higher managerial/professional occupations (7%) or intermediate occupations (5%) had the lowest prevalence of overweight, whereas children with fathers in routine manual occupations had a much higher prevalence (10%). There was a similar but stronger pattern in 2007 whereby children with parents in higher managerial/ professional occupations (15%) had a relatively lower prevalence rate of overweight compared to children from all other social classes (NS-SEC 2 19%; NS-SEC 3 20%; NS-SEC 4 20%; NS-SEC 5 21%; NS-SEC 6 24% and NS-SEC 7 22%). Children from the NS-SEC 6 category had the highest prevalence of overweight, which was nearly 10 percentage points higher than children in the NS-SEC 1 category. A formal test of the interaction between NS-SEC and time was significant ($\chi 2(12)=23.56$, $p<0.05$, $n=33357$), suggesting that the differences between the NS-SEC groups had become larger over time. However it is the very low prevalence rates of children in NS-SEC 1 in 2007 that is driving this interaction as there are only small differences between the other NS-SEC groups. The coefficients from this interaction model are presented in table 4.4.

RGSC

The results for the RGSC measure of SES are presented in the second block of table 4.3. The relationship between parental RGSC and child overweight status was not statistically significant in 1965 or in 1980. Nevertheless, in 1980 there is some evidence of a weak linear relationship, as the prevalence of child overweight consistently decreases from RGSC V (9%) to RGSC I (6%). In 2007, there is a stronger association between parental RGSC and child overweight. The prevalence rates increased with decreasing social class up to RGSC IV. As shown with the NS-SEC measure, there are very low prevalence rates of overweight for children in the highest RGSC group (RGSC I: 14%) compared to all other social class groups (RGSC II 18%, RGSC III 21%, RGSC IV 23%, RGSC V 20%). As might be expected from these descriptive statistics, a formal test of the interaction between RGSC and the time was significant ($\chi 2(8)=16.31$, $p<0.05$, $n=36129$), suggesting that the groups have become more different over time, although again this is largely driven by the low prevalence rates of the highest social class group in 2007. The coefficients from the interaction model are presented in table 4.5.

Table 4.3. The prevalence of overweight in three British birth cohorts by socioeconomic status.

	1965	**1980**	**2007**
	Percentage (SE)	Percentage (SE)	Percentage (SE)
Social Class			
NS-SEC 1	8 (1.0)	7 (0.7)	15 (1.0)
NS-SEC 2	8 (0.9)	8 (0.7)	19 (0.7)
NS-SEC 3	7 (0.9)	5 (0.7)	20 (1.2)
NS-SEC 4	10 (0.9)	9 (0.8)	20 (1.1)
NS-SEC 5	9 (0.7)	9 (0.7)	21 (1.5)
NS-SEC 6	8 (0.7)	9 (0.7)	24 (1.2)
NS-SEC 7	9 (0.6)	10 (0.6)	22 (1.2)
Chi Square/ F test	$\chi2(6)=5.60$, $p>0.05$, n=9382	$\chi2(6)=22.96$, $p>0.05$, n=10658	$F(5.55, 2159)=8.28$, $p<0.05$, n=13317
Interaction over time	$\chi2(12)=23.56$, $p<0.05$, n=33357).		
RGSC			
V	7 (0.7)	9 (1.4)	20 (2.5)
IV	9 (0.7)	9 (0.8)	23 (1.2)
III	9 (0.3)	8 (0.4)	21 (0.7)
II	10 (0.7)	8 (0.5)	18 (0.7)
I	8 (1.2)	6 (0.9)	14 (1.2)
Chi Square/ F test	$\chi2(4)= 6.07$, $p>0.05$, n=12787	$\chi2(4)=3.88$ $p>0.05$, n=10831	$F(3.70, 1441)=7.56$, $p<0.05$, n=12511
Interaction over time	$\chi2(8)=16.31$, $p<0.05$, n=36129		
Mother's Education			
Minimum education	10 (0.3)	9 (0.3)	22 (0.6)
Above Minimum education	8 (0.8)	7 (0.5)	19 (0.6)
n	12843	10064	13274
Interaction over time	$\chi2(2)=0.31$, $p>0.05$, n=36181		
Father's Education			
Minimum education	9 (0.3)	9 (0.4)	21 (0.6)
Above Minimum education	8 (0.5)	7 (0.5)	17 (0.7)
n	12784	9565	10118
Interaction over time	$\chi2(2) =2.91$, $p>0.05$, n=32467		
Latent SES measure			
Quint 1	8 (0.5)	9 (0.5)	22 (0.9)
Quint 2	10 (0.6)	9 (0.6)	24 (1.0)
Quint 3	9 (0.6)	9 (0.6)	21 (0.9)
Quint 4	10 (0.5)	8 (0.6)	19 (0.8)
Quint 5	9 (0.6)	7 (0.5)	15 (0.9)
Chi Square/ F test	$\chi2(4) =5.64$, $p>0.05$, n= 13295	$\chi2(4) =11.23$, $p<0.05$, n= 12138	$F(3.80, 1480)= 13.59$, $p<0.05$
Interaction over time	$\chi2(8) =22.11$, $p<0.05$, n=39269		

Notes: Figures presented for MCS were derived taking into account the complex survey design using the 'svy' commands in Stata. Probability weight used was 'weight2' survey design weight.
F-tests presented for MCS instead of chi square tests due to using the 'svy' commands in analysis.
For testing interactions over time, MCS responses were only weighted using the probability weights.

Table 4.4. The interaction between NS-SEC and time period dummies

	OR	S.E	z	p	95% confidence Intervals	
NS-SEC[+]						
2	0.96	0.16	-0.21	0.83	0.69	1.34
3	0.86	0.16	-0.81	0.42	0.60	1.23
4	1.18	0.19	1.02	0.31	0.86	1.62
5	1.08	0.17	0.50	0.61	0.80	1.47
6	0.98	0.15	-0.11	0.91	0.72	1.34
7	1.10	0.16	0.65	0.52	0.82	1.47
Time period^						
1980	0.79	0.14	-1.34	0.18	0.57	1.11
2007	1.91	0.28	4.38	0.00	1.43	2.54
NS-SEC#Time period						
2 # 1980	1.19	0.27	0.80	0.43	0.77	1.85
2 # 2007	1.36	0.25	1.62	0.11	0.94	1.96
3 # 1980	0.86	0.22	-0.59	0.55	0.52	1.42
3 # 2007	1.60	0.33	2.23	0.03	1.06	2.41
4 # 1980	1.13	0.25	0.57	0.57	0.74	1.74
4 # 2007	1.17	0.23	0.84	0.40	0.81	1.71
5 # 1980	1.18	0.25	0.78	0.44	0.78	1.78
5 # 2007	1.36	0.27	1.54	0.12	0.92	2.00
6 # 1980	1.27	0.27	1.12	0.26	0.84	1.94
6 # 2007	1.83	0.34	3.23	0.00	1.27	2.64
7 # 1980	1.32	0.27	1.38	0.17	0.89	1.96
7 # 2007	1.53	0.28	2.35	0.02	1.07	2.18
constant	0.09	0.01	-18.27	0.00	0.07	0.12
n	33 357					
Fit statistics	R2[a]	log pseudolikelihood[b]			BIC	
	0.0357	-12640.863			-321798.913	

[a] pseudo R2

[b] pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.

Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.

[+]NS-SEC 1 (Higher managerial and professional occupations) is the reference category

[^] 1965 is the reference category

 # refers to an interaction

Table 4.5. The interaction between RGSC and time period dummies

	OR	S.E	z	p	95% confidence Intervals	
RGSC[+]						
IV	1.23	0.18	1.39	0.16	0.92	1.65
III	1.13	0.15	0.91	0.36	0.87	1.48
III	1.32	0.20	1.82	0.07	0.98	1.77
I	1.01	0.19	0.04	0.97	0.69	1.47
Time period[^]						
1980	1.12	0.24	0.53	0.60	0.73	1.71
2007	3.02	0.63	5.32	0.00	2.01	4.54
RGSC#Time period						
IV # 1980	0.79	0.19	-0.97	0.33	0.48	1.28
IV # 2007	0.93	0.21	-0.34	0.74	0.59	1.45
III # 1980	0.83	0.19	-0.81	0.42	0.54	1.30
III # 2007	0.89	0.19	-0.53	0.59	0.58	1.36
II # 1980	0.68	0.16	-1.59	0.11	0.43	1.09
II # 2007	0.64	0.14	-1.97	0.05	0.41	1.00
I # 1980	0.70	0.21	-1.20	0.23	0.39	1.26
I # 2007	0.61	0.17	-1.84	0.07	0.35	1.03
constant	0.09	0.01	-18.81	0.00	0.07	0.11
n	36 129					
Fit statistics	$R2^a$	log pseudolikelihood [b]			BIC	
	0.0327	-13383.181			-352150.239	

[a] pseudo R2
[b] pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.
Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.
[+]RGSC V is the reference category
[^] 1965 is the reference category
 # refers to an interaction

Parental education

The results for parental education are shown in the third and fourth section of table 4.3, mother's education is shown first, followed by father's education. Higher levels of mother's education reduced the odds of a child being overweight in all three birth cohorts, the same was true for father's education. There was also reasonable stability in the size of the odds ratios[23] for mother's education over time (OR=0.82 in 1965,

[23] OR should be interpreted such that an OR of 1 means equal odds of being overweight, OR over 1 mean an increase in the odds of being overweight, and OR below 1 mean a decrease in the odds. OR are always compared to another group, they are relative. OR represent the difference in the odds of "success", or in this case the odds of being overweight compared to a reference category. The odds are simply the probability of success/probability of failure. The ratio is calculated by dividing the odds of one group success by the odds of another. For example imagine 20% of males are overweight and 33% females are overweight. The odds of being overweight for males is (0.2/(1-0.2)=0.25) 0.25 and for females (0.33/(1-0.33))=0.5. The OR for being overweight for females then is (0.5/0.25)=2.00.

OR=0.77 in 1980 and OR=0.81 in 2007). The interaction between mother's education and time was not statistically significant ($\chi2(2)=0.31$, $p>0.05$, n=36181). The coefficients from this full interaction model are presented in table 4.6.

The difference in the odds of being overweight by father's education was smallest in 1965 (OR=0.86), but the odds ratios in 1980 and 2007 were quite similar (OR = 0.73 in 1980 and 0.77 in 2007). The change in the odds ratios could suggest that father's education is a more important correlate of child overweight in 1980 and 2007 than in 1965, but the interaction between father's education and time was not statistically significant ($\chi2(2)=2.91$, $p>0.05$, n=32467). This suggests that changes in the relationship between father's school leaving age and child overweight has not changed by a statistically significant amount over time. The coefficients from this full interaction model are presented in table 4.7. The relationship between father's education and child weight status is explored in more detail in chapter 6.

As a robustness check, parental education in 2007 was coded as staying in full time education past the age of 18, instead of the age of 16. This reduced the predicted odds ratios in 2007 for main respondents (OR=0.77) but had little impact on the predicted odds ratios for partner respondents (OR=0.77). The interaction with time was not significant for main respondents education ($\chi2(2) =2.33$, $p>0.05$) or partner respondents education ($\chi2(2)=2.63$, $p>0.05$).

Table 4.6. The interaction between whether or not mother was at school past minimum leaving age and time period dummies

	OR	S.E	z	p	95% confidence Intervals	
Mum school -stayed[+]	0.82	0.06	-2.72	0.01	0.71	0.95
Time period[^]						
1980	0.96	0.05	-0.74	0.46	0.86	1.07
2007	2.63	0.13	19.07	0.00	2.38	2.91
Mum school#Time period						
stayed # 1980	0.95	0.10	-0.52	0.61	0.76	1.17
stayed # 2007	0.96	0.09	-0.45	0.65	0.80	1.15
constant	0.11	0.00	-64.83	0.00	0.10	0.11
n	36 181					
Fit statistics	R2 [a]	log pseudolikelihood [b]			BIC	
	0.0316	-13583.02			-352474.252	

[a] pseudo R2
[b] pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.
Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.
[+]did not stay at school past the minimum school leaving age is the reference category
[^] 1965 is the reference category
 # refers to an interaction

Table 4.7. The interaction between whether or not father was at school past minimum leaving age and time period dummies

	OR	S.E	z	p	95% confidence Intervals	
Dad school - stayed[+]	0.86	0.06	-2.05	0.04	0.74	0.99
Time period[^]						
1980	0.99	0.05	-0.11	0.91	0.89	1.11
2007	2.53	0.13	17.81	0.00	2.29	2.81
Dad school#Time period						
stayed # 1980	0.85	0.10	-1.42	0.16	0.68	1.06
steayed # 2007	0.86	0.08	-1.54	0.12	0.71	1.04
constant	0.10	0.00	-65.35	0.00	0.10	0.11
n	32 467					
Fit statistics	R2 [a]	log pseudolikelihood [b]			BIC	
	0.0277	-12020.185			-313101.504	

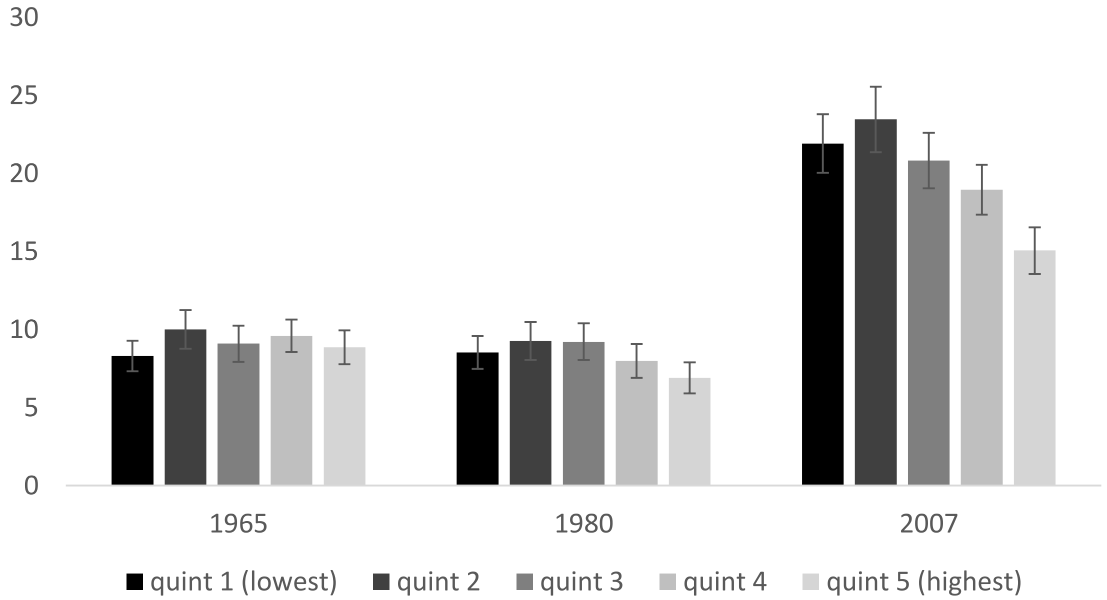
[a] pseudo R2
[b] pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.
Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.
[+]not staying at school past the minimum school leaving age is the reference category
[^] 1965 is the reference category
 # refers to an interaction

Latent measure

The results for the latent measure of SES are shown in the final section of table 4.3. In 1965 there was no evidence for social patterning in the prevalence of overweight using the latent measure of SES. In 1980, the highest SES quintile had the lowest prevalence rate of child overweight and this increased with decreasing SES for the 4 highest SES quintiles (7%, 8%, 9%, and 9% respectively). In 2007 the prevalence of overweight for children in the bottom four quintiles, that is the lowest 80% of the SES distribution ranged from 19-24%, whereas only 15% of children in the highest SES quintile were classified as overweight. As shown in figure 4.5, in 2007 there is a clear gradient in overweight prevalence by SES quintile for quintiles 2-5.

The relationship between the latent measure of SES and child overweight did vary significantly over time ($\chi2(8)=22.11$, $p<0.05$, $n=39269$), but as appears to be a consistent finding across the NS-SEC, RGSC and the latent measure of SES, it is the very low prevalence rates of the highest socioeconomic group in 2007 that are driving the interaction over time.

Figure 4.5. The proportion of children (aged between 7-10) classified as overweight by parental socioeconomic group over three time points. Error bars represent 95% confidence intervals.



Did these socioeconomic trends vary by gender?

To test whether the relationship between the different measures of SES and child overweight varied over time in the same or in different ways for males and female children, three way interactions were tested looking at the interaction between child sex, the measure of socioeconomic status and time. None of these interactions were statistically significant suggesting that the relationships changed over time in a similar way for males and females. These results are reported in full in appendix C.

Robustness check 1: Principal Components Analysis

As an alternative to CFA I constructed Principal components scores[24] using principal components analysis (PCA). The input matrix for this analysis was a set of polychoric correlations[25]. PCA is a descriptive data reduction technique that converts a set of correlated variables into a set of uncorrelated variables called principal components. The first component is always the component which describes the largest variance in the data. This procedure is related to factor analysis, but rather than assuming that the observed variables are predicted by the latent variable, the latent variables (the components) are predicted by the observed variables. PCA creates the latent variables based on a linear combination of the observed variables, therefore changing the observed variables will the interpretation of the latent construct. This is in direct contrast to factor analysis. PCA is favoured by some researchers, because it involves making fewer assumptions about the relationship between the observed variables and the 'components'. The results based on PCA are very similar to those presented in this chapter. The results can be found in appendix D.

Robustness check 2: Different specifications of the latent variable model

As well as using a different methodology to create a composite measure of SES, Several different specifications for deriving a latent measure of socioeconomic status were run to assess whether allowing the residuals to covary, using missing data methods or the use of different estimators made a large difference to the results. In all cases the pairwise correlations between the factor scores generated through these different specifications were very high (>.95). Furthermore the standardised coefficients did not change a great deal between different specifications, suggesting

[24] *Principal components Analysis is a descriptive technique for maximising the amount of variance in a set of variables accounted for by a smaller set of components.*

[25] *In stata the 'polychoric' command fits polyserial, pearson and polychoric correlations dependent upon the shape of the data. the matrix produced may contain all different types of correlation dependent upon the scale of measurement of the variables.* Polychoric correlation coefficients are a measure of association for ordinal variables which rests upon an assumption of an underlying joint continuous distribution.

that the findings would be very similar regardless of how the latent variable measurement model was specified. The results from this analysis are shown in appendix E.

Robustness Check 3: Accounting for increasing ethnic variation in the samples

Whilst ethnicity is a crucial issue in inequalities research, the small numbers of children from ethnic minority backgrounds in the NCDS and BCS70 cohorts makes it difficult to consider the impact of ethnicity on overweight. Some ethnic groups have lower prevalence rates and some higher prevalence rates than the 'white' ethnic groups (Jebb, Rennie and Cole, 2004; Wardle *et al.*, 2006), therefore it is not feasible to create a white/ non-white dummy variable to consider the changing impact of ethnicity on child overweight. The proportion of the sample in the birth cohorts comprising of children from ethnic minorities has increased over time. Given there is evidence that socioeconomic status can vary by ethnicity (Nazroo and Karlsen, 2001; Platt, 2011),it is plausible that the finding that inequalities by SES have widened over time is the result of confounding between SES and ethnicity. Therefore the increasing proportion of children from ethnic backgrounds needs to be accounted for. As a robustness check all analyses were run on a 'white only' sample, whereby all children from ethnic minority backgrounds were excluded from the sample. The results from this analysis are very similar to those presented here and are available in Appendix F.

Discussion

The findings of this study indicate that inequalities in childhood weight status by social class (NS-SEC or RGSC) or latent measures of SES have widened over time. There has been however a stability in the influence of parental education on children's probability of being overweight. These widening socioeconomic inequalities appear to be driven by the relatively low increases in prevalence rates of those in the highest social class or socioeconomic groups, compared to all other SES groups. This finding is interesting because it challenges previous suggestions that widening inequalities in child overweight are being driven by the lowest socioeconomic groups (Stamatakis *et al.*, 2005; Stamatakis, Wardle and Cole, 2010). Indeed the results presented here tend to suggest that the lack of increasing prevalence in the higher socioeconomic groups over time distinguishes them from all other socioeconomic groups.

The discordance between the findings presented here, and those of previous research may be due to differences in the granularity of the measurement of SES. Previous research has used binary or three category measures of SES, largely due to sample size restrictions (Stamatakis *et al.*, 2005; Stamatakis, Wardle and Cole, 2010; White *et al.*, 2007). It may be that these less granular measures of SES were masking the unique trend for the highest socioeconomic group, by combining this group with other social groups. Despite disagreeing on which SES groups are driving the widening inequalities in SES over time, the results presented here do generally support the previous evidence for widening inequalities over time (Stamatakis *et al.*, 2005; Stamatakis, Wardle and Cole, 2010). Although, the results presented here place emphasis on the other end of the socioeconomic spectrum.

There are potential policy relevant implications from this paper. Previous recommendations have been to target the lower socioeconomic groups, as evidence indicated that the lower socioeconomic groups were experiencing a faster rate of increase in child obesity prevalence than other social groups (Stamatakis *et al.*, 2005; Stamatakis, Wardle and Cole, 2010). Indeed there has been a focus on lower socioeconomic groups for obesity intervention (Hollar *et al.*, 2010; Horodynski *et al.*, 2011; Tyler and Horner, 2008), however the current results suggest that the middle socioeconomic groups are also in need of intervention, when compared to the highest socioeconomic group. Perhaps consideration needs to be given not only to why the disadvantaged are more at risk for overweight and obesity in childhood, but also to what is “protecting” children in the higher socioeconomic groups from the obesogenic changes in the environment (Egger and Swinburn, 1997). One potential explanation could be that the highest socioeconomic groups, or the “elite”, are the only ones who have the resources to shield their children from the obesogenic environment (Egger and Swinburn, 1997).

In 1965 there was no statistically significant relationship between social class or socioeconomic status and the proportion of children classified as overweight. This is congruent with previous research which used measures of parental social class to investigate the probability of a child being classified as overweight (Power and

Moynihan, 1988; White *et al.*, 2007). However, the results presented here show a statistically significant association between parental education and child overweight status in 1965. This association persists in 1980 and in 2007, and the magnitude of the effect is reasonably stable, even when the measure of parental school leaving age is raised to 18 instead of 16 in the MCS cohort.

Parental education has not been considered by other researchers who have considered socioeconomic inequalities in the trends in child overweight (Stamatakis *et al.*, 2005; Stamatakis, Wardle and Cole, 2010; White *et al.*, 2007). But the relationship between parental education and the probability of a child being overweight or obese is reasonably well established (El-Sayed, Scarborough and Galea, 2012a; Greenlund *et al.*, 1996; Lamerz *et al.*, 2005; Shrewsbury and Wardle, 2008). The stability in the influence of parental education could suggest that the pathways, through which higher levels of parental education reduce child obesity, such as the ability to access and utilise health information (Williams *et al.*, 2013) or through the modelling of lifestyle behaviours (Patrick and Nicklas, 2005; Savage, Fisher and Birch, 2007), were just as important before the obesity epidemic began, and have remained important despite changes in the environment which make it easier to become obese (Egger and Swinburn, 1997).

The differential findings for parental education versus social class and the latent measure of socioeconomic status reinforces the point that socioeconomic indicators are not interchangeable, and that different indicators are measuring unique concepts which may have differential relationships with outcomes (Bukodi and Goldthorpe, 2012; Goldthorpe, 2012). However it may be potentially problematic to measure changes in SES over time through occupation based measures, because there have been changes in the types of occupations available, the requirements for obtaining these occupations and the rewards and prestige for certain occupations (Gallie, 2001; Wyatt and Hecker, 2006). Equally the meaning of continuing past compulsory education and the returns from doing so have changed (Greenaway and Haynes, 2003).

The latent measure of SES is a measure of relative social position which ranks families within each time period. It is plausible then that this latent measure is not as open to this bias created by the changes in occupations and education over time, because it compares parents with other parents at the same time period. Therefore a latent measure may more accurately reflect access to social and economic resources at a given time period, because it is ranking children based on their families social position and access to these resources (Oakes and Rossi, 2003) compared to all other children in the sample for that specific time point.

The analysis presented in this chapter has several strengths which make it unique. The birth cohort data sets are large and rich data sources which provide excellent data on a range of childhood issues. Prior to this analysis, these data had not been utilised to look at trends in socioeconomic inequalities in child obesity or child obesity trends more generally. Much of the previous literature is based around two data sets only, the HSE and the NSHG (Chinn and Rona, 2001; Stamatakis *et al.*, 2005; Stamatakis, Wardle and Cole, 2010; Stamatakis *et al.*, 2010; White *et al.*, 2007). However, there is a large time gap between the BCS70 cohort and the MCS cohort, which likely explains why these data have not been used. There is no way to tell from the results presented here what happened between 1980 and 2007. However, the available literature can shed some light on this issue. Previous research suggests that the socioeconomic inequalities in the probability of a child being overweight did not begin to widen until the late 1990's/early 2000's (Stamatakis *et al.*, 2005), and they continued to widen thereafter (Stamatakis, Wardle and Cole, 2010). This may provide a time frame for when the differences observed between 1980 and 2007 began to emerge.

Summary

This chapter provided information on the long terms trends in childhood obesity and how socioeconomic inequalities in childhood weight status have changed over time. The results are largely dependent upon the measurement of socioeconomic status, if we look at parental education, it appears that the proportional differences in childhood overweight have remained reasonably similar over time. However measures of social class and a latent measure combining social class with attribution aspects of SES suggests that differences have become more divergent over time. One indicator of socioeconomic status that could not be included in the analysis in this chapter is income. This is because the measures of income in the NCDS and BCS are potentially problematic (Erikson and Goldthorpe, 2010). Most importantly for the chapter above, familial income is not measured in the NCDS until the children are aged 16. More generally there is a large amount of missingness in income in the NCDS and BCS, the income data reported by the self-employed is of low quality, and from a comparative perspective because Income has been constructed in different ways in these cohorts it is not straightforward to make comparisons (Erikson and Goldthorpe, 2010).

There has been a great deal of interest in the role of income and income inequalities on childhood outcomes. With regards to child obesity, there is an expectation that low income causes obesity, because food choices are limited by the money available to buy food. Healthier foods are expensive, whereas high calorie, energy dense foods are relatively inexpensive. This expected relationship between income and child obesity is influencing policy makers to consider policies based on a redistribution of income to tackle child obesity and inequalities within child obesity. The relationship between income and child obesity is considered in the next chapter, using data from the MCS.

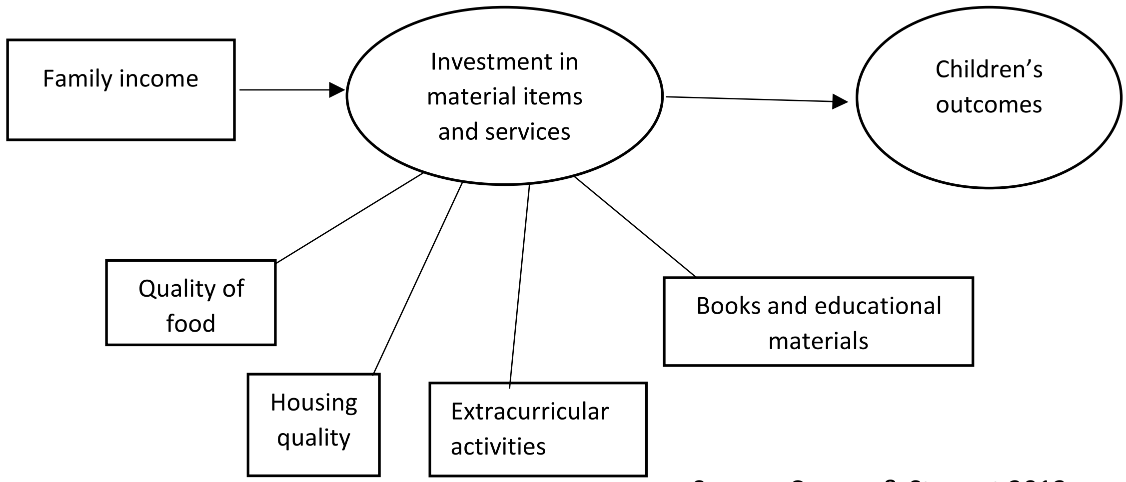
Chapter 5: The relationship between low parental income and childhood overweight and obesity.

This chapter addresses whether or not there is a relationship between familial income and childhood obesity. This chapter begins by introducing readers to the literature on the influence of parental income on childhood outcomes more generally. There is a rich literature looking at the relationship between familial income and other childhood outcomes, most notably educational outcomes. Upon reviewing the literature it is clear that there is a gap regarding the relationship between income and childhood health outcomes including child overweight and obesity. Furthermore the vast majority of the evidence looking at familial income and childhood outcomes comes from the US.

Children from low income or ‘poor’ families generally have worse outcomes than children from more affluent families. This is true for a range of child outcomes including health, educational and behavioural outcomes (Mayer, 2010; Mayer, 1997; Mayer, 2002). There are several explanations for this relationship. The most popular explanation, informed by economic theory on human capital, suggests that more affluent parents can spend more on, or “invest” more in their children and this leads to better outcomes (Cooper and Stewart, 2013; Jenkins and Schluter, 2002b; Mayer, 2010). This is known as the economic investment model and is shown in figure 5.1. The child related expenditures or investments that might be considered important for children’s development and consequent life chances include things such as nutrition, clothing, learning resources and cultural activities.

This particular explanation for the relationship between income and children’s outcomes is enticing, because form a policy perspective it is easy to address the problem. The life chances of poor children can be lifted simply by increasing the investments in poor children either through increasing the income that poor families have to spend on their children, or through increasing the amount spent on children from poor families within institutions such as schools and health services.

Figure 5.1. The economic investment model

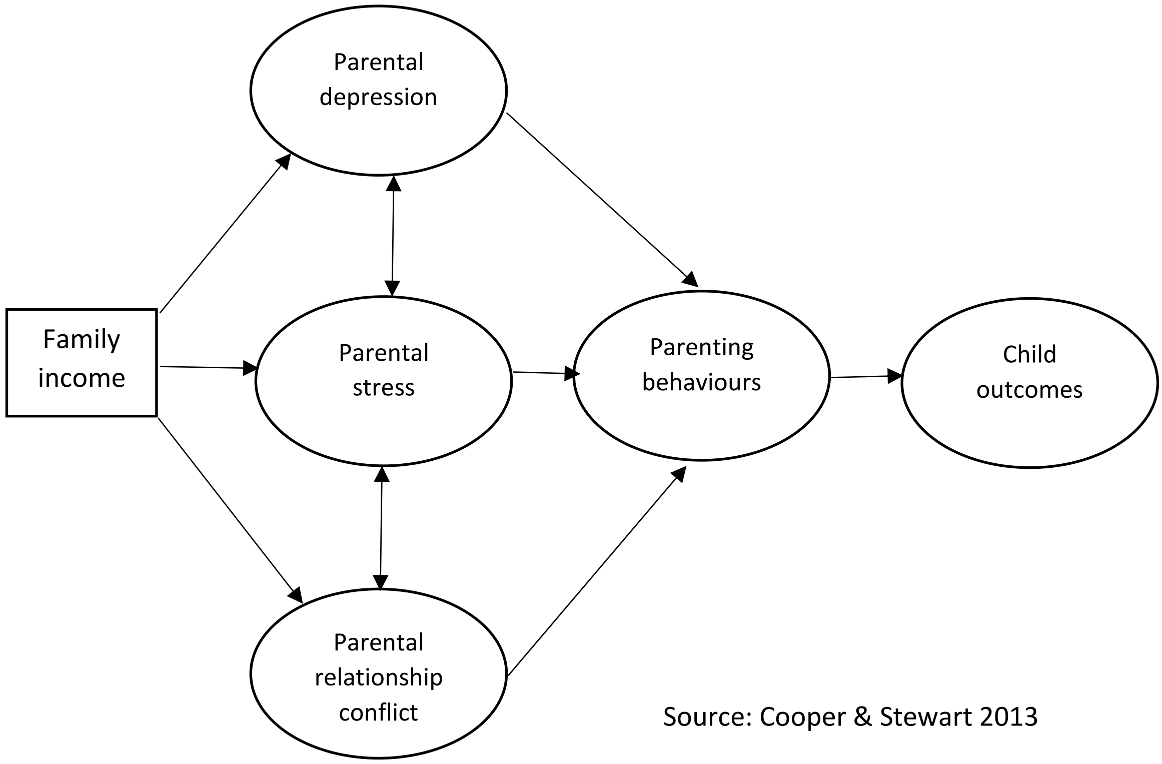


Source: Cooper & Stewart 2013

The second potential explanation for the relationship between low income and child outcomes concerns parenting behaviours (Cooper and Stewart, 2013; Jenkins and Schluter, 2002a). This explanation, based on the family stress model put forward by Conger and colleagues (1994), suggests that having a low income increases parental stress levels, and this stress hinders parent's ability to provide quality care, which has negative impacts upon children's outcomes (Conger, Conger and Martin, 2010; Conger *et al.*, 1994). Therefore income is hypothesised as having an indirect effect on childhood outcomes through the influence that low income has on parenting. This increased stress has reinforcing consequences, parents with high levels of stress are more likely to have poor physical health, poor mental health and poorer spousal relationships. This is known as the parental stress model, and is depicted in figure 5.2 below.

The third possible explanation, is that other characteristics of the parents are correlated with income, for example parental education, parental behaviours, motivation, attitudes & beliefs, and it is these characteristics that impact upon children's outcomes (Mayer, 2010; Mayer, 1997; Mayer, 2002). In this explanation the observed association between income and childhood outcomes is not causal and actually reflects characteristics of the parents that tend to be correlated with income and 'good' parenting. If this explanation were true, increasing income in itself would have little impact upon children's outcomes (Mayer, 1997).

Figure 5.2. The parental stress model



What does the evidence look like for the relationship between income and child outcomes more generally?

There is some debate as to whether income has a causal influence on childhood outcomes, and how important income is. Susan Mayer (1997/2002/2010) reviewed the evidence on the relationship between income and children's outcomes and found that the effect of income was small. She suggested that this was because parental characteristics that employers were willing to pay for, such as reliability, good health and honesty, improved children's life chances independent of their effect on parental income, and that children with parents who have these characteristics do well even when their parents have low incomes. She suggests that the things that do actually help increase children's life chances cost very little, and that when people experience increases in income they do not purchase these things, but rather they purchase more luxurious items. Mayer highlight the difficulties of obtaining causal estimates of the impact of income on child outcomes, stating:

"People's income depends on their skills, their work efforts and other factors. These factors can also affect children's outcomes. Research that estimates the correlation between parental income and children's outcomes cannot tell us anything about the causal relationship between parental income and children's outcomes because such estimates do not control for all such parental characteristics." (Mayer, 2002, p. 12)

In contrast Kerris Cooper & Kitty Stewart (2013) suggest there is strong evidence for a causal relationship between parental income and children's outcomes and that the influence of income on children's outcomes is non-trivial. They systematically reviewed 34 studies from the USA, UK, Canada, Norway and Mexico. Although, the majority of these studies came from the USA (22/34) and only 4 came from the UK. Based on analyses of all studies they suggest that children from low income families have poorer outcomes in part because they are poorer, not just because of the correlation between income and parental characteristics. They found supporting evidence for the economic investment model and the parental stress model for the mechanisms through which income was influencing children's outcomes. Across the 34 studies there were reasonably consistent results for the impact of parental income on children's educational and cognitive outcomes, and to a lesser extent on their social and behavioural outcomes, but the results for the influence of parental income on children's physical health outcomes were inconsistent, with 3/8 studies finding no effect. The results were less convincing for the four UK studies, with 2/4 studies in total finding no significant relationship between income and child outcomes.

Mayer (2002) has also found the least consistent results for childhood health outcomes, which she attributed also to the lack of high quality research looking at the relationship between parental income and childhood health. In her own words:

"surprisingly few studies estimate the effect of parental income on children's health and those that do often use weak measures of income…or unrepresentative samples" (Mayer, 2002, p. 37)

Therefore there is clearly a need for more high quality research considering the relationship between income and childhood health outcomes, there is also a need for more high quality research considering the relationship between income and child health outcomes specifically in the UK context.

Motivation for considering the role of income on child obesity

As discussed in the first and fourth chapters, childhood obesity has a socioeconomic gradient, with children in more advantaged families less likely to be overweight or obese (El-Sayed, Scarborough and Galea, 2012a; Observatory, 2012; Shrewsbury and Wardle, 2008). There is an expectation that income is associated with child obesity, and that there may be a causal relationship between income and child obesity. This expectation has arisen because the different indicators of socioeconomic status are often treated as though they are interchangeable (Bukodi and Goldthorpe, 2012) and because there is evidence for a relationship between income and child obesity from other countries (Costa-Font and Gil, 2013; Freedman *et al.*, 2007; Jo, 2012). This expectation has been expressed by politicians and in key policy documents in the UK. For instance, in a speech given to the food and drink industry, Anne Soubry, the minister responsible for public health, claimed that the low cost of unhealthy foods was responsible for child obesity levels, and that the poorest children were most likely to be obese (Anna Soubry, 2013). As a further example, in a recent obesity government policy paper titled 'Healthy Lives, Healthy people: A Call To Action on Obesity' (2011), it is claimed that:

"*..income, social deprivation and ethnicity have an important impact on the likelihood of becoming obese. For example, women and children in lower socioeconomic groups are more likely to be obese than those who are wealthier."* (HM Government, 2011, p12).

The claim that the low cost of unhealthy foods is driving child obesity rates and inequalities within them has been put forward in the media repeatedly (Gallagher, 2013; Peck, 2013; Williams, 2011). This explanation for child obesity reflects the economic investment model, in that an assumption is being made that poor families do not have the resources to buy healthier foods and therefore the problem can be fixed by increasing the means available with which to buy food. This approach leads to two obvious policy responses: either to give more money to low income families to spend on food, or increase the price of unhealthy food and make healthy food more affordable. The latter approach is receiving renewed support (Obesity Steering Group, 2013) after David Cameron initially suggested a "fat tax" at the conservative conference in October 2011 (Press Association, 2011). There is currently a debate regarding whether high sugar foods should also be taxed (BBC NEWS, 2014). Although there is little being said about whether healthier foods would be made more affordable for low income families.

The assumption held by some that SES is synonymous to income is potentially problematic. If income is not in fact the key driver of inequalities in child obesity, it may wrongly influence those in positions of power to make or support changes that are potentially ineffective in tackling child obesity.

What does the evidence say regarding the relationship between income and child obesity specifically?

Within the UK context the findings are mixed as to whether there is an association between familial income and child overweight and obesity, but the majority of studies find that higher incomes are associated with a lower risk of a child being overweight or obese (El-Sayed, Scarborough and Galea, 2012a; Griffiths *et al.*, 2010b; Jotangia, 2005; Observatory, 2012). Research published using the MCS indicates that the bivariate relationship between household income and child weight status was not significant when the children were aged 3 (Hawkins *et al.*, 2009) but was significant when the children were aged 5 (Griffiths *et al.*, 2010a). This suggests that the association may change with age.

Indeed the relationship may vary by sex as well as age. Analyses of different age groupings of children in the Health Survey for England (HSE) data came to different conclusions regarding the association between income and child overweight and obesity. One study considered children aged 2-15 and found a strong association between income and child overweight and obesity for female children but not male children (Observatory, 2012), whereas when children were grouped into an age band of 2-10 year olds and were not disaggregated by sex, the observed association between income and child overweight and obesity was much weaker (Jotangia, 2005).

Research from America also indicates that the relationship between income and child overweight or obesity varies by the child's age, with a stronger association between income and child overweight and obesity amongst older children (Freedman *et al.*, 2007). Research from America has also shown that the relationship between income and child overweight can vary by ethnicity (Freedman *et al.*, 2007; Wang and Zhang, 2006). Amongst children aged 6-11 and 12-18 Freedman and colleagues (2007) found that increasing income was associated with a decrease in child overweight and obesity in 'white' and 'Mexican American' ethnic groups, but the reverse pattern found amongst children from 'black' ethnic groups. Similar results were found by Wang & Zhang (2006) who found that for African American children, higher incomes were associated with a higher risk of overweight and obesity.

These analyses do not necessarily tell us if a relationship between income and childhood obesity exists, because other aspects of socioeconomic status such as parental education and social class are correlated with income, and are often not included in the models (Babey *et al.*, 2010; Eagle *et al.*, 2012; Freedman *et al.*, 2007; Gigante *et al.*, 2013; Griffiths *et al.*, 2010a). Additionally other characteristics of the parents such as motivation, attitudes and beliefs may also be correlated with income. Multiple regression can be used to account for some of these potential confounding influences. Where multiple regression has been used to look at the relationship between income and child weight status, there are still problems. Firstly, the measurement of income has not been ideal. The scale on which income is measured has not always been adequately considered, nor has the distinction between the "transitory" and "permanent[26]" components (Brophy *et al.*, 2009; Matijasevich *et al.*, 2009). Yet, leading economists (Jenkins and Schluter, 2002a) and sociologists (Erikson and Goldthorpe, 2009) agree this distinction is crucial when looking at intergenerational relationships.

There are exceptions to this. Washbrook, Gregg & Propper (2013) used a time-averaged measure of income from two points in time and included measures of

[26] Permanent income is a measurement of average income over time, in which temporary fluctuations in income do not have much effect upon consumption patterns or economic welfare (Friedman, 1957). Transitory income on the other hand is a one shot measurement of income, which is likely to contain larger amounts of measurement error and may not adequately reflect long term economic wellbeing.

parental education and social class in their modelling. They found only a very weak bivariate relationship between income and child fat mass in the UK. Through decomposition analysis they demonstrated that the majority of the relationship could be attributed to parental education (60%) and the rest to social class (40%), suggesting no direct relationship between familial income and child fat mass (Washbrook, Gregg and Propper, 2013).

In contrast, a study based on an American cohort of children that also used time-averaged income measures found evidence for an association between income and child overweight (Jo, 2012). This study found that family income increases the probability of child obesity for very poor families, but that income only has a negative relationship with children's obesity rate once a certain income threshold is achieved. Jo (2012) explains the nonlinear relationship between income and child obesity by suggesting that very poor families may scarcely be able to afford any type of food whereas the moderately poor may opt for cheap but highly calorific foods. Equally Jo (2012) suggests that the very poor may not be able to afford a home entertainment system, which results in the children burning more calories. However, there are differences between the UK and the USA which mean that research conducted in America is not necessarily applicable in the UK context. There are differences in the system of social stratification, health care systems, institutions and policies which support the economic, social and physiological wellbeing of citizens. It may well be that income has a stronger association with child obesity in America than in the UK.

The aim of the analysis

The aim of the analysis presented in this chapter is to provide additional evidence on whether or not there is a relationship between familial income and child obesity in the UK context. This analysis will make several important contributions to the literature by providing additional evidence as to whether income is associated with child overweight and obesity, improving upon the measurement of income utilised in other studies, considering the relationship between low income and child overweight status independent of other aspects of SES, and utilising the longitudinal component of the MCS to consider whether changes in income are associated with changes in overweight and obesity. Due to concerns over data quality income was not considered in the previous chapter, however the superior quality of the income data in the MCS means that it is possible to look at this relationship with these data.

Methodology

The sample for the Millennium Cohort Study is described in detail in chapter 3. I will be using the ‘svy’ commands in stata version 13 (StataCorp, 2013) with the dovwt2 probability weight[27]. The weight “dovwt2” reflects the inverse probability of selection into the sample, adjusted for the probability of sweep level non-response. The probability of non-response is based on observable characteristics of the families who participated in at least one prior sweep in which data was collected (Plewis, 2007a).

The outcome variable used in most of the analyses is taken from sweep 4 (age 7). Out of the 19,177 possible families, there were 17,031 eligible families at sweep 4, with productive responses achieved from 13,857 (81% of those eligible) families (14,043 children). Height and weight data were available for 13,813 children. 14 of these children were missing measurements of income, leaving 13,799 children in the sample. The characteristics of the sample for this analysis are described in Table 5.1, along with the characteristics of those who are missing from the analysis, but were productive at sweep 4 (244 (2%) in total). Children who were excluded from the sample tended to have parents with lower income, lower levels of education, lower social class status, were more likely to report having a longstanding illness, be of non-white ethnic origin and be male.

Measuring Overweight

Whether the child was overweight/obese was measured at age 7 by the IOTF criteria described in chapter 2 (Cole *et al.*, 2000). The relevant age appropriate cut off values were calculated by linear interpolation of the IOTF half year BMI values, based on the sex and age (to the nearest 10th of a year) of the children at measurement of height and weight. Height and weight measurements were taken by trained interviewers using standardised protocols. As shown in table 5.1 approximately 20% of the sample

[27] As a robustness check the models with run with just the survey design weight not adjusted for longitudinal attrition from the sample. The model were also run using an alternative method of dealing with the complex survey design whereby the strata variable was entered into the regression model, the standard errors were clustered at the electoral ward level and the probability weight were applied. The results do not noticeably differ from those presented here.

are classified as overweight and obese. Weight was measured using high precision scales. Weight was measured to the nearest 0.1kg after removing heavy objects from pockets/hair etc and whilst wearing light indoor clothing. Height was measured to the nearest millimetre with parental assistance using a Leicester Stadiometer (Gray, Gatenby and Huang, 2010).

Measuring Income

Income is measured using OECD-modified equivalised weekly net income, which is weighted by the household size and composition (Anyaegbu, 2010). The equivalisation is based on a couple with no children, which is set at a value of 1. A value of 0.67 is given to the first adult, 0.33 for the spouse, 0.33 for each additional child aged between 14-18 and 0.2 for children under 14 (Anyaegbu, 2010; Rosenberg, 2012). The aim of this equivalisation is to take into account that a higher amount of income is needed when more people are in the household to achieve a comparable standard of living as a lower income would achieve in a smaller household. For example, a single adult in a household is estimated to need 0.67 times the income of a household comprised of two adults in order to achieve a comparable standard of living. Whereas a household with two adults and two children under the age of 14 is estimated to need 1.4 times the amount of income as a household with just two adults in order to achieve a comparable standard of living. Table 5.1 indicates that the mean weekly equivalised income of families in the sample is £346 per week. The proportion of children missing from the sample is very small and any bias this causes in the sample should not have a large influence on the results.

The measurement of income in the MCS is very comprehensive, particularly in the third and fourth sweeps of data collection. As well as asking parents about gross and net earnings from employment, including earnings form second jobs, irregular work and self-employment, a raft of questions relating to income from numerous benefits and grants were also including in the measurement of income (Hansen *et al.*, 2010). Due to the comprehensive nature of the measurement, in the MCS the measurement of income forms a good assessment of how much funds families have available to them.

Table 5.1. Characteristics of the sample for analysis in the chapter.

	Sample	**Not in Sample**
N	**13799**	**244**
Overweight	20%	N/A
Equivalised Income Mean (SD)	346 (197)	269 (167)
Quintiles of Income	Mean (range)	
Quint 1	134 (40 - 172)	
Quint 2	214 (172 - 258)	
Quint 3	307 (258 - 360)	
Quint 4	424 (360 - 502)	
Quint 5	666 (502 - 1258)	
Main Respondent Education	%	%
NVQ 1	8	9
NVQ 2	27	27
NVQ 3	15	12
NVQ 4	29	19
NVQ 5	6	3
Overseas	3	2
None	11	27
Social Class	%	%
High Manag/prof	16	12
Lower Manag/prof	28	17
Intermediate	13	10
Small emp/self employed	9	8
Lower Sup & Technical	7	7
Semi Routine	13	20
Routine	7	9
No Classification	8	17
Longstanding Illness/disability	%	%
yes	19	29
no	81	60
Not Reported	0	11
Ethnicity	%	%
White	86	73
Other Ethnic group	3	5
Black/Black British	2	2
Pakistani/Bangladeshi	4	12
Indian	3	5
Mixed	1	4
Sex	%	%
Female	49	37
Male	51	63

Values weighted using dovwt2 survey design & attrition weight.

Income was not observed for 8% (1574/18552) of the families in sweep 1, 14% (2314/15590) in sweep 2, 11% of families in sweeps 3 (1629/15246) and 11% of families in sweep 4 (1579/13857). Information on age, housing tenure, labour market status of main and partner respondents, sampling strata, region of residence, whether or not receiving benefits, ethnicity, highest NVQ qualification, type of accommodation, number of children in household and lone parent status were used to impute (estimate) the missing values for income by the survey organisers using interval regression. This imputed data was deposited with the MCS. More details about this procedure for predicting income can be found in the MCS guide to the dataset, from page 81-85 (Hansen *et al.*, 2010). Using this imputed data is potentially problematic, particularly as income is a key variable in the analysis. Imputed information is only as good as the model upon which the imputation is based (Carpenter and Kenward, 2013). In this case, the model appears to be very good, as income is likely to be well predicted by the covariates listed above. Also, the amount of imputed data is small, so if income was not well predicted it is unlikely to have a large influence on the results[28].

Measures of income from a single point in time (transitory income) are subject to measurement error (Hyslop and Imbens, 2001). Measurement error is likely to downwardly bias parameter estimates, and may mask relationships that exist between family income and child obesity. This measurement error can be reduced by averaging income across different time points (time-averaged income) (Blanden, Gregg and Macmillan, 2013; Solon, 1992). Time-averaged income was generated by averaging across the four sweeps of data collection covering 7-8 years. Transitory income was measured at sweep 4 only, when the children are approximately 7 years old. In the results section transitory and time-averaged income measures are compared to demonstrate that the strength of the relationship varies depending on the measure of income chosen.

It is important to consider the functional form of the relationship between income and child overweight before attempting to model it. The functional form refers to the shape of the relationship. Fitting a linear functional form when a relationship is not

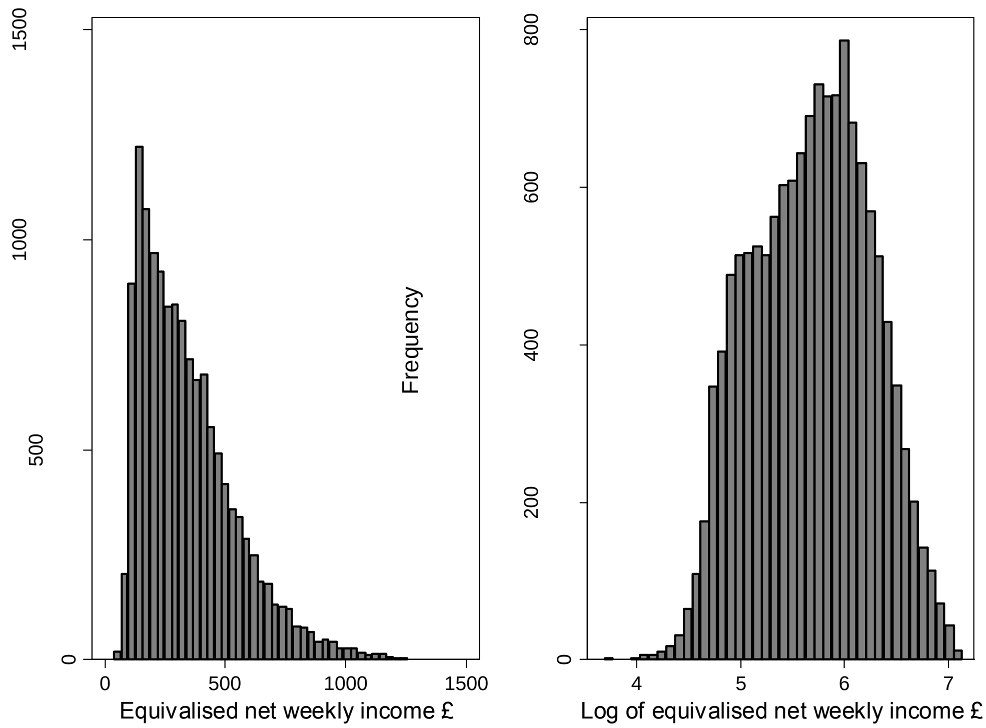
[28] As a sensitivity check, the analysis was also conducted using the observed income data i.e non-imputed income. These results were very consistent whether using imputed or non-imputed income.

linear may underestimate the true extent of the relationship between income and overweight (Cooper and Stewart, 2013). The shape of the relationship between familial income and child overweight will be assessed using non parametric methodology called local polynomial smoothing, which will be described in the analysis section. The relationship between the log of income and child overweight will be examined. If the relationship is log-linear, income will be treated as continuous, which will increase the power in the analysis. If the linearity assumption does not hold, income will be divided into equal fifths (quintiles), so that different portions of the income distribution can be compared without assuming linearity. These quintiles will be weighted using survey design and sweep non-response weights. The lowest income quintile will be used as the reference category, making comparisons with the lowest income quintile straightforward.

The log of income is utilised rather than income for two reasons. Firstly, the distribution of income is highly positively skewed, as shown in figure 5.3. By using the log transformation it is possible to change the shape of the distribution so that it is a closer approximation of a normal distribution. Secondly, and more importantly, it makes more sense to interpret the log of income. The logarithmic scale changes the interpretation of income from additive to multiplicative. This takes into account the proportional value of increased in income. For example an increase in income of £5000 per annum for somebody earning £20,000 per annum is proportionately bigger than an increase of £5000 for somebody earning £100,000, this is taken into account using the log transformed income as the difference between £25,000 (log = 10.12) and 20,000 (log = 9.90) is bigger than the difference between £105,000 (log = 11.56) and £100,000 (log = 11.51).

The distribution of equivalised weekly income is shown on the left hand side of figure 5.3. As can be seen this distribution is positively skewed, with a tail on the right hand side of the distribution which represents families with very high levels of equivalised income. This is equivalised net weekly income so the distribution represents the amount of income families actually have to spend, conditional upon the number of adults and children in the household. The log of equivalised weekly net income is shown on the right hand side of figure 5.3. This distribution is not perfectly normally distributed, but it is a much closer approximation. There appears to be two peaks in the distribution, a peak at around the log value of 4.7 (roughly £109 equivalised), and another peak at around the value of 6 (roughly £403 equivalised). The lower peak of the distribution is largely comprised of children who are in families where the only source of income is welfare benefits, these are mostly lone parent families, or families in which both parents are not in paid employment.

Figure 5.3 The distribution of equivalised weekly income and the log of equivalised weekly income in the MCS data.



Other Socioeconomic characteristics

As one of the aims of the paper is to try and separate the effect of income from other socioeconomic characteristics. Both parental education and social class are included as covariates in the model. Once other socioeconomic characteristics are included in the model we are measuring the 'conditional income effects', or the effect of income net of other socioeconomic characteristics. The literature clearly shows that conditional income effects are smaller than unconditional income effects (Jenkins and Schluter, 2002a). Therefore it is expected that including these other socioeconomic characteristics will reduce the estimated association between income and child overweight. Considering the conditional income effects could be problematic if we assume that other socioeconomic characterises, i.e education and social class are on the causal pathway between income and child overweight status. If this assumption were true, it would lead to an underestimate of the association between income and child overweight.

However it is unlikely that parental income causes parental education and social class. Indeed the causal pathway is most likely to operate from education, to social class and income (Kuha and Goldthorpe, 2010). Logically, education opens the doors to certain occupations and not to others. These occupations will vary in terms of social class (defined in this context by employment relations, chances for progression, employment stability and control/autonomy over work) and income. Whilst occupation based social class and income are likely to be correlated, as certain occupations are rewarded with higher income than others, these two constructs are not measuring the same thing and income does vary within social classes (Blanden, Gregg and Macmillan, 2010).

These three aspects of socioeconomic status are correlated, as shown in table 5.2. However the correlations are not so high to suspect that collinearity would be problematic. Collinearity diagnostics[29] also suggest that the inclusion of these three variables in the same model is not problematic. Nevertheless I am aware of arguments for not including different aspects of socioeconomic status in the same model. Blau (1999) notes the dangers of including variables that may be jointly chosen with household income. Therefore these covariates are added sequentially. Model fit can be assessed by the inclusion of these covariates and the results can be interpreted for the unconditional income effect, the effect of income conditional upon education and the effect of income conditional upon education and social class.

[29] The tolerance of the variables, the percent of variance in the covariate that cannot be accounted for by the other covariates, is high, meaning that a large proportion of the variance in the variables cannot be accounted for by the inclusion of the other variables. The Variance Inflation Factor (VIF) is 1/tolerance. High VIF Numbers indicate that collinearity may be problematic. The VIF is 1.48 on average for these variables, ranging from 1.43 – 1.60, well below the cut off values of between 5-10, in which collinearity may be problematic (Belsley, Kuh and Welsch, 2005)

Table 5.2 Polychoric correlations between different aspects of SES to be included in models.

	Income	Main Social Class	Partner Social Class	Main NVQ	Partner NVQ
Income	1.00				
Main Social Class	-0.59	1.00			
Partner Social Class	-0.60	0.43	1.00		
Main NVQ	0.46	-0.56	-0.38	1.00	
Partner NVQ	0.43	-0.35	-0.58	0.41	1.00

Notes: The 'polychoric' command was used in stata. This command takes into account the level of measurement of the data and calculates the appropriate type of correlation. When both variables have 10 or fewer observed values, a polychoric correlation is calculated, when only one of the variables takes on 10 or fewer values (i.e., one variable is continuous and the other categorical) a polyserial correlation is calculated, and if both variables take on more than 10 values a Pearson's correlation is calculated

Parental education

Parental education was measured using the highest NVQ equivalent qualification of main responders and partner responders. NVQ's were originally created to simplify the qualifications awarded to vocational learning. However in the context of measuring educational achievement, they bring together both vocational qualifications and academic based qualification into a comparable format. This is important as just using academic based measures of education, such as years of schooling or highest academic qualification may result in an underestimation of the education level of people who undergo specific vocational training. Parental qualifications are categorised into NVQ levels 1-5, overseas qualifications, or no qualifications. Higher NVQ levels relate to a higher level of qualification. NVQ level 5 is equivalent to obtaining a postgraduate degree i.e a masters or doctorate degree, while NVQ level 1 and 2 are equivalent to finishing secondary school education, with level 1 referring to lower achievement on final examinations. The highest NVQ equivalent qualification achieved across all sweeps of data collection was created using information reported at each sweep on further academic or vocational qualifications achieved (Rosenberg, 2012). NVQ levels were derived by the survey organisers using the criteria outlined in table 5.3.

The main parental respondent was the child's mother in almost all sweeps for almost every child. In sweep 1, 99% of main responders were mothers, 98% in sweep2, 97% in sweep 3 and sweep4. Where NVQ level was missing, it was replaced with the last reported highest NVQ level equivalent qualification of the main respondent from previous sweeps[30]. Partners education is taken from the highest NVQ level reported at sweep 4. In sweep 4, 89% of responses to the partner questionnaire were from natural fathers and a further 7% were from father figures, such as step fathers, adoptive fathers and foster fathers. Partner respondent's education was missing in 4% of cases where there were two parents in the household.

Table 5.3. The derivation of NVQ levels in the MCS cohort.

Academic/ vocational Qualification	**NVQ level**
Higher Degree and Postgraduate qualifications	5
First Degree (including B.Ed.)	4
Post-graduate Diplomas and Certificates	5
Diplomas in higher education and other higher education qualifications	4
Teaching qualifications for schools or further education (below degree level)	4
A/AS/S Levels/SCE Higher, Scottish Certificate Sixth Year Studies, Leaving Certificate or equivalent	3
O Level or GCSE grade A-C, SCE Standard, Ordinary grades 1-3 or Junior Certificate grade A-C	2
CSE below grade 1/GCSE or O Level below grade C, SCE Standard, Ordinary grades below grade 3 or Junior Certificate below grade C	1
Other academic qualifications (incl. some overseas)	N/A
None of these qualifications	N/A
Professional qualifications at degree level e.g. graduate member of professional institute, chartered accountant or surveyor	5
Nursing or other medical qualifications (below degree level)	4
NVQ or SVQ level 4 or 5	4
HND, HNC, Higher Level BTEC/RSA Higher Diploma	4
NVQ or SVQ Level 3/GNVQ Advanced or GSVQ Level 3	3
OND, ONCM BTEC National, SCOTVEC National Certificate	3
City & Guilds advanced craft, Part III/RSA Advanced Diploma	3
NVQ or SVQ Level 2/GNVQ Intermediate or GSVQ Level 2	2
BTEC, SCOTVEC first or general diploma	2
City & Guilds Craft or Part II/RSA Diploma	2
NVQ or SVQ Level 1/GNVQ Foundation Level or GSVQ Level 1	1
BTEC, SCOTVEC first or general certificate/SCOTVEC modules	1
City & Guilds part 1/RSA Stage I,II,III/Junior certificate	1
Other vocational qualifications (incl. some overseas)	N/A
None of these qualification	N/A

[30] This is an example of using Last Observation Carried Forward (LOCF). In many applications LOCF is problematic because there are likely to be changes in the variable over time. For this application it is only possible for there to be positive changes, i.e the highest level of qualification cannot be reduced, only increased. And for the vast majority of observed cases, highest qualification obtained does not change.

Social Class

Social class was measured using the NS-SEC (Rose, Pevalin and O'Reilly, 2005), which is described more fully in the previous chapter. The principles underlying the NS-SEC are primarily employment relations (Rose, Pevalin and O'Reilly, 2005). Social class was derived from respondents' last known occupation so that social class could be assigned to those who were currently out of the labour force. Emily Beller (2009), put forward a convincing argument for including information from both mother's and father's social class in models interested in intergenerational relationships. She present both a theoretical and empirical argument that both parents social class provides valuable information about the families class resources, net of the effect of the other parents social class. Empirically she shows that including information from both parents, and not just the social class of one parent results in better fitting models (Beller, 2009).

In this case model fit was tested using highest household social class, derived by taking either main or partner respondent's social class and using main and partner respondents social class as separate covariates, i.e including both main and partner respondents social class. Comparisons of model fit showed using the highest household social class rather than mothers and father's social class independently resulted in a better fitting model[31]. Please see appendix G for model fit criteria.

Demographic Variables

This chapter also aims to separate out the effect of income from parental characteristics that are likely correlated with income. Therefore a rich list of demographic variables are included in the modelling process. All demographic variables were chosen because they were plausibly associated to both familial income and child overweight. These included; Parental report of child's ethnicity (6 category census classification), whether or not the child, main and partner respondent had a longstanding illness/disability, region of residence, lone parent status and parental age. Around 22% of the children resided in lone parent households. Child's ethnicity was missing for only one case. This missing value was replaced with ethnicity reported in previous sweeps of data collection. Information on whether or not the child had a

[31] To ensure this decision did not have a large influence on the results, the models were run with mother's and father's social class included as separate covariates and the results were substantively no different from those presented here.

longstanding illness, and whether the main respondent had a long standing illness was only missing in less than 1% of cases. In two parent families information from partner respondent's long standing illness was missing in 15% of cases. There were no missing values for region or main parental respondent age or partner respond age, where a partner was present in the household. A missing data category was included for those missing information on long standing illness so that these cases were not lost from the analytic sample.

Ad-hoc methods for missing data

To ensure the same sample is used in each analysis, missing dummy variables were used to keep cases that had missing information on the covariates in the models. For continuous variables this involves setting the value of the missing response to the mean value, and including a variable indicating that this response was missing. For categorical variables, this involves including a missing variable category.

The use of 'ad-hoc' methods for replacing missing values has been deemed inappropriate (Carpenter and Kenward, 2013). There are however, advantages to this 'ad-hoc' approach. With large numbers of covariates in the model, it is almost inevitable that some item non-response will be present. Where questions are more sensitive they are less likely to elicit responses. This item specific non-response is seldom random, so for example people with very low or very high incomes may be less likely to report income (Riphahn and Serfling, 2005). Where full case analysis is conducted, only the sample which responded to all possible covariates is included in the analysis. This is problematic because it is unlikely to be a representative sample. The coefficients, for example for income may not reflect the true effect of income because the sample now represents a group of people with particular characteristics. These people are likely reasonably well off, highly educated, in good health and also may possess other characteristics such as perfectionism or tenacity which drove them to answer the questions.

In a simple regression the use of missing categories and missing dummies does not alter the point estimates for each covariate. I demonstrate this in table 5.4 below where I consider the relationship between father's education (the variable with the highest level of missingness) and child overweight status. For multiple regression the same is also true, the point estimates are identical to full case analysis, except if a person did not respond to one covariate, for example age, but they did respond to education, their response to education is included in the estimation of the education coefficients. This approach could be conceptually compared to looking at pairwise correlations, rather than correlations following listwise deletion. Specification 'a' in table 5.4 refers to estimates obtained through the use of a missing data category, whereas specification 'b' refers to estimates obtained through complete case analysis. The estimated odds, standard errors and t statistics are identical for the categories of father's education.

The only difference is that we estimate the relationship between the missing values and child overweight and obesity. This missing group will be highly heterogeneous and includes families where there is no father present and where fathers were present but they didn't respond for various reasons. There are more sophisticated techniques available for handling missing data that allows for the modelling of these missing values, FIML was discussed in chapter four and multiple imputation will be discussed in chapter six. Complete case analyses of these data are available in Appendix H.

Table 5.4. Comparing the coefficients and standard errors for analysis using missing data category (a), and for complete case analysis (b).

	Odds		**Std. Err.**		**t**	
	a	b	a	b	a	b
NVQ 1			Reference category			
NVQ 2	-0.19	-0.19	0.12	0.12	-1.53	-1.53
NVQ 3	-0.24	-0.24	0.14	0.14	-1.70	-1.70
NVQ 4	-0.39	-0.39	0.12	0.12	-3.21	-3.21
NVQ 5	-0.57	-0.57	0.15	0.15	-3.82	-3.82
Overseas	-0.29	-0.29	0.16	0.16	-1.79	-1.79
None	0.08	0.08	0.15	0.15	0.52	0.52
Missing	0.00		0.12		-0.02	
_cons	-1.20	-1.20	0.11	0.11	-10.75	-10.75
n	a=13799 b=10486					

Analysis

Firstly, the bivariate relationship between income and child overweight will be explored to establish what the functional form of the relationship between income and child obesity is. Secondly differences in the bivariate relationship between transitory and time averaged income will be assessed to highlight the advantages of using time-averaged measures of income. The strength of the bivariate association between familial income and child obesity will be considered. Potential confounding variables will be controlled for to establish whether they explain the relationship between familial income and obesity. Confounding variables are those that have both a relationship with familial income and child overweight. These covariates will added to the model in blocks, so that the effect of each 'block' on the income coefficients can be assessed. If the magnitude of the coefficients reduces after adding in covariates, this means part of the originally observed effect of income on child overweight is actually the result of the correlation between income and the newly added covariates. The covariate blocks include demographics (model 2), main parental respondent education (model 3), partner respondent education (model 4) and social class (model 5). Social class was added to the models last because, although there is income variability within social class groups (Blanden, Gregg and Macmillan, 2013), it is potentially difficult to identify and separate their effect.

As the MCS data are longitudinal it is possible to look at whether changes in familial income influence changes in child weight status. However in order to look at changes over time measures of transitory income have to be considered. Fixed and random effect models are considered. Several robustness checks were conducted to check whether the results were a statistical artefact, whether they resulted from the measure of income used, or whether they were specific to the MCS cohort.

What functional form should income take?

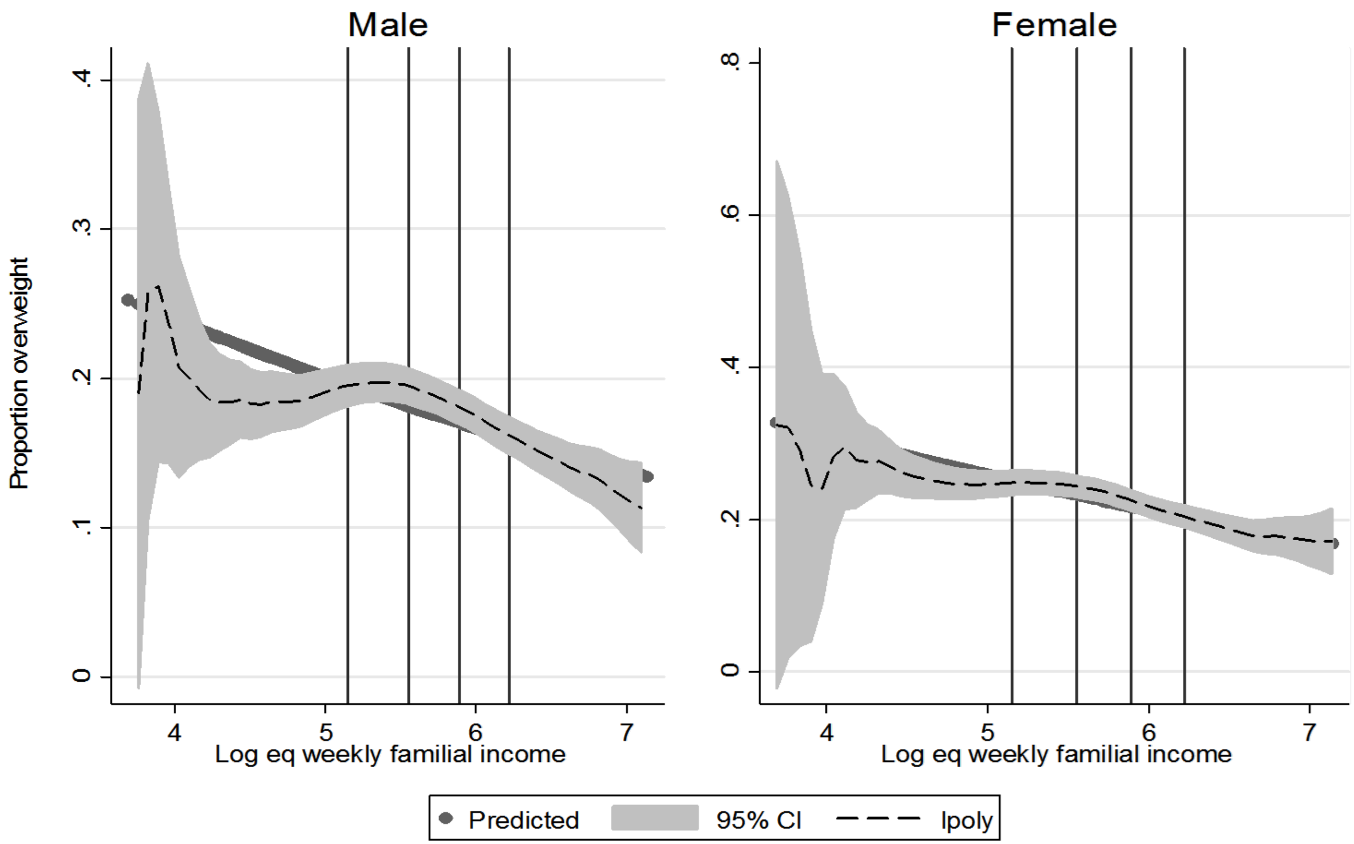
The shape of the relationship between income and child obesity was assessed using local polynomial smoothing. Local polynomial smoothing is a non-parametric technique for plotting the relationship between two variables, without assuming any functional form (i.e linear or quadratic) (Gutierrez, Linhart and Pitblado, 2003). It is effectively a smoothed scatter plot of the relationship between the independent variable, X (income), and the dependent variable, Y(child overweight and obesity). For each point on the scatter plot, X_0, a low order weighted least squares regression is fit using only points that are within a certain neighbourhood from X, this is the bandwidth. In this case the appropriate bandwidth was estimated using the Rule of Thumb (ROT) estimator[32].

Figure 5.4 shows the relationship between the log of time-averaged familial income (x-axis) and the probability of the child being overweight (y-axis). Figure 5.4 demonstrates that a linear functional form would be acceptable for females, but probably not for males. There is deviation from linearity for males in the bottom 20% of the income distribution, evidenced by the 95% confidence intervals not consistently touching the predicted linear line. Figure 5.4 suggests that, for males, the likelihood of overweight increases with increasing familial income within the bottom 40% of the income distribution.

For considering the impact of low income on child overweight status, income was divided into quintiles (equal fifths). The choice of quintiles is consistent with the existing literature (Gigante *et al.*, 2013; Jotangia, 2005), and different number of income groups are considered in the robustness section. Quintiles allow us to compare children in the poorest 20% of households to the rest of the income distribution. Where income is modelled as a continuous variable, a quadratic term is tested and fit where appropriate.

[32] *The ROT is a plugin estimator of the asymptotically optimal constant bandwidth, which minimizes the conditional weighted mean integrated squared error.*

Figure 5.4. Local polynomial smooth of the relationship between log-income and child overweight.



Notes: Vertical lines represent income quintiles. Vertical axes are not identical for males and females. The appropriate bandwidth was estimated using the Rule of Thumb (ROT) estimator, a plugin estimator of the asymptotically optimal constant bandwidth, which minimizes the conditional weighted mean integrated squared error. The confidence intervals do not take into account the clustered nature of the data, this means that the shaded grey area is too narrow.

Comparing transitory income and time-averaged income

Estimates of the relationship between time-averaged income and childhood obesity can be found in Table 5.5 (females) and Table 5.6 (males). The results are presented as marginal effects, therefore they represent the difference in the probability of a child being overweight for the income quintiles, compared to the lowest income quintile. The left hand most column (Model 0) presents estimates using transitory income and the adjacent column (Model 1) using time-averaged income. It was hypothesised that the association between income and overweight status would be stronger when using time-averaged income, due to there being less measurement error. Tables 5.5 and 5.6 show that this is the case. For males and females the difference in the probability between income quintile 1 and 5 is larger when using time-averaged income measures (Male -0.04: Female -0.08), compared to transitory income measures (Males -0.02: Females -0.05). Furthermore for males, the difference between the highest and lowest income quintile is statistically significant with time-averaged income measures, but not with transitory income measures. Using transitory measures of income results in the estimation of the relationships between income and child overweight being lower than using measures of income which are averaged over time.

This is perhaps not surprising from both a theoretical and empirical standpoint. Over short periods of time families can smooth consumption and thus smooth their living standards (Mayer, 2002). Therefore a transitory measure of low income may not reflect living standards as families can use savings or borrow money to alleviate the impact of low income in the short run. Empirically, transitory income has a bigger variance than permanent income, therefore estimated correlations will always be smaller for this reason.

Logistic regression models

Females

The results for females are presented in Table 5.5. Marginal effects show the difference in the probability of being classified as overweight, compared to the lowest income quintile, when all other covariates are held at their means. The unadjusted relationship (Model 1) shows that the difference in the probability of being classified

Table 5.5. The relationship between quintiles of income and the probability of overweight for females. Results are presented as marginal effects.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Tran Income Q1	Reference category					
Tran Income Q2	-0.00 (0.02)					
Tran Income Q3	-0.01 (0.02)					
Tran Income Q4	-0.03 (0.02)+					
Tran Income Q5	-0.05 (0.02)**					
Tran missing	0.30 (0.19)					
TA Income Q1		Reference category				
TA Income Q2		-0.00 (0.02)	0.01 (0.02)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)
TA Income Q3		-0.02 (0.02)	-0.01 (0.02)	-0.00 (0.02)	0.00 (0.02)	0.01 (0.02)
TA Income Q4		-0.03 (0.02)	-0.02 (0.02)	-0.00 (0.02)	0.01 (0.02)	0.02 (0.02)
TA Income Q5		-0.08 (0.02)***	-0.06 (0.02)**	-0.04 (0.02)+	-0.03 (0.02)	-0.01 (0.03)
N	6,830	6,830	6,830	6,830	6,830	6,830

+ $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

Model 0 = Transitory Income; Model 1 = Time averaged Income; Model 2 = Model1 + Ethnicity, longstanding illness/disability of child, main and partner respondent, region, parental age; Model 3 = Model 2 + Main respondent education; Model 4 = Model 3 + Partner respondent education; Model 5 = Model 4 + Highest social class of parents.
Estimates weighted using dovwt2 survey design & attrition weight

as overweight is lower for income quintiles three (-0.02), four (-0.03) and five (-0.08), compared to the lowest income quintile. However, only the difference between the highest and lowest income quintile is statistically significant, with the probability of a child being classified as overweight 8 percentage points lower than children in the lowest income quintile. This suggests that the relationship is being driven by children in families at the very top end of the equivalised income distribution, because for 80% of the income distribution, there is no significant difference in the probability of children being overweight, and the differences in the probability of being classified as overweight are not so large. The differences between Q1 and Q5 are not due to the low income group; rather the high income quintile is exceptional.

This difference between Q1 and Q5 is somewhat reduced after adjustment for demographic characteristics (Model 2). Children in income quintile 5 have a 6 percentage point lower probability of being classified as overweight than children in the lowest income quintile. The difference between the highest and lowest income quintiles remains statistically significant. Therefore, the difference in the probability of child overweight for those in Q1 and Q5 cannot be fully explained by the differential demographic characteristics of the income quintiles. The literature suggested that the relationship between income and child overweight and obesity may vary by ethnicity. This was formally tested using an interaction between income and ethnicity ($F(19,371)=1.01$, $p>0.05$). There was no evidence that the relationship between income and child overweight varied by a significant amount for different ethnic groups.

The statistically significant difference between Q1 and Q5 does not persist after adjustment for main respondent's education (Model 3), although the difference is approaching significance and is below the $p<0.10$ threshold. Holding demographic characteristics and main respondent's education constant, children in the highest income quintile have a probability of being classified as overweight that is 4 percentage points lower than the lowest income quintile. After adjustment for partner respondent's education (Model 4), there are no statistically significant differences between income quintiles. Furthermore there is very little difference in the probability of being classified as overweight amongst any of the income quintiles compared to the lowest income quintile. The difference between the highest and lowest income quintile is completely attenuated after adjustment for social class (model 5). This means there is almost no difference in the probability of being overweight by income quintiles, after adjustment for all covariates in the model. The full table of coefficients is available in appendix I.

Males

The results for males are presented in Table 5.6. The unadjusted relationship (Model 1) suggests that children in the lowest income quintile are significantly more likely to be classified as overweight than those in the highest income quintile. The probability of being classified as overweight is 4 percentage point lower in the highest income quintile compared to the lowest. Interestingly, for male children the highest probability of being overweight is in the second income quintile, not the lowest income quintile. Children in income quintile 2 have a 3 percentage point increase in the probability of being classified as overweight compared to the lowest income quintile. Therefore there are statistically significant differences in the probability of being classified as overweight between income quintile 2 and 5.

After adjustment for demographic characteristics (Model 2), there are no longer statistically significant differences between the highest and lowest income quintiles. Children in the highest income quintile have a 2 percentage point decrease in the probability of being classified as overweight compared to the lowest income quintile. This attenuation from Model 1 to Model 2 suggests that the difference in probability of child overweight for male children in the lowest income quintiles compared to the highest income quintile is partially explained by the differential demographic characteristics of the children in the highest and lowest income quintiles.

After controlling for demographic characteristics statistically significant differences emerge between the lowest income quintile and the second income quintile. Children in income quintile 2 have a four percentage point higher probability of being classified as overweight than children in the lowest income quintile. This could be expected given the shape of the relationship shown in figure 5.4 for males, with a peak in the probability of child overweight in the second income quintile. Again the interaction between ethnicity and income was tested ($F(20,369)=0.89$, $p>0.05$) but there was no evidence that the relationship varied by ethnic group. After successive adjustment of the model for main respondent education (Model 3), partner respondent education (Model 4) and social class (Model 5), the significant difference between Q1 and Q2 persists, with children in the second income quintile consistently estimated to have a higher probability of being classified as overweight compared to the lowest income quintile (model 3: +0.04, model 4: +0.04, model 5: +0.04). Male children in income quintile 3, the middle quintile, also have a higher probability of being classified as overweight than those in the lowest income quintile once demographic characteristics have been controlled for. The difference between the middle income quintile and the lowest income quintile is not statistically significant, but does approach significance once partner respondent's education has been controlled for (Model 4).

Table 5.6. The relationship between quintiles of income and the probability of overweight for males. Results are presented as marginal effects.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Tran Income Q1	Reference category					
Tran Income Q2	0.03 (0.02)+					
Tran Income Q3	0.03 (0.02)+					
Tran Income Q4	-0.00 (0.02)					
Tran Income Q5	-0.02 (0.02)					
Tran missing	0.06 (0.22)					
TA Income Q1		Reference category				
TA Income Q2		0.03 (0.02)+	0.04 (0.02)*	0.04 (0.02)*	0.04 (0.02)*	0.04 (0.02)*
TA Income Q3		0.01 (0.02)	0.03 (0.02)	0.03 (0.02)	0.04 (0.02)+	0.04 (0.02)+
TA Income Q4		-0.02 (0.02)	-0.00 (0.02)	0.01 (0.02)	0.02 (0.02)	0.02 (0.02)
TA Income Q5		-0.04 (0.02)*	-0.02 (0.02)	-0.01 (0.02)	0.01 (0.02)	0.02 (0.02)
N	6,969	6,969	6,969	6,969	6,969	6,969

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 0 = Transitory Income; Model 1 = Time averaged Income; Model 2 = Model1 + Ethnicity, longstanding illness/disability of child, main and partner respondent, region, parental age; Model 3 = Model 2 + Main respondent education; Model 4 = Model 3 + Partner respondent education; Model 5 = Model 4 + Highest social class of parents.
Estimates weighted using dovwt2 survey design & attrition weight

Therefore, for reasons not explained by the demographic characteristics included in the model, parental education or social class, male children in families in Q2 have a higher probability of being classified as overweight than children in families in the lowest income quintile. It is improbable that this finding is an artefact of using quintiles, rather than some other number of income groups, because figure 5.4 clearly shows a peak in the probability of overweight just below the middle of the income distribution. However, sensitivity of the results to the use of different numbers of income groups is explored in the robustness section. The probability of being classified as overweight in Q2 is not statistically significantly different from any other income quintile, except than Q1, after adjustment for parental education.

The nonlinear association between income and overweight for males suggests that for the poorest families, males are less likely to be overweight. This is similar to the findings of Jo (2012), except that here it is only applicable to males. Jo (2012) explained this finding suggesting that children from very poor families were either more active, due to a lack of home entertainment systems, or did not have access to sufficient quantities of food to become overweight, due to families not having enough money to purchase large amounts of food.

Another potential explanation could be that families in quintile 2 represent the 'working poor'[33] whereas those in the lowest income quintile represent families mostly living on welfare benefits alone. A high proportion of parents in the lowest income quintile are unemployed (main respondent: 78%, partner respondent (where present): 50%). Whereas in the second income quintile, a higher proportion of the parents are in employment (52% main respondents and 86% partner respondents (where present)). Therefore whilst parental income may be lower in the lowest income group, it may be that families who belong to the working poor are under more strain, due to having to balance work, home and childcare commitments than families where the parent or parents do not work. In the lowest income quintile it may be that, even though there is less money to spend on food, there is more time dedicated to meal preparation. It may be that parents who are not in employment spend more time engaging in physical activities with children because they have time in which to do so.

Alternatively as discussed further in the conclusions section, this may be the result of measurement error for those who report very low incomes (Brewer, Etheridge and O'Dea, 2013). However, the actual size of the effect is quite small, and there is no obvious reason that the potential explanations given above shouldn't hold true for females also.

Coefficients for the different aspects of SES

As one of the aims of this chapter is to differentiate between different aspects of SES and income, it is potentially interesting to consider the relationship between the other aspects of SES included in the model and the probability of a child being classified as overweight. The coefficients for the other aspects of SES included in the models are

[33] The working poor have been defined as people who spend 27 weeks or more in a year "in the labor force" either working or looking for work but whose incomes fall below the poverty level.

presented in table 5.7. The results for males are shown on the left hand side and the results for females are shown on the right hand side.

The probability of being overweight does not differ significantly by main respondent's education level for males or females, holding demographic characteristics and equivalised income constant. For females however, the changes in probability are in the expected direction, with increasing main respondent's education associated with decreases in the probability of a child being overweight. The difference between NVQ 1 and NVQ 3 and NVQ 1 and NVQ 5 are significant at the ten percent level.

Interestingly, when partner respondents education is included in the model specification (model 4), the probability of a child being overweight does vary significantly by partner respondent's education for both males and females, with higher levels of partner respondents education associated with decreases in the probability of being overweight. Indeed for male children there is a difference of 8.5 percentage points in the probability of being overweight when partner respondents have NVQ level 1 qualifications, compared to NVQ level 5 qualifications. For females there is a difference of 6.5 percentage points. The different probabilities by partner respondent's education level remain statistically significant and of similar magnitude for male children after social class is included in the model (model 5). For female children the difference in probability between NVQ 1 and NVQ 5 is only significant at the 10 percent level after adjustment for social class.

There are no statistically significant differences in the probability of being overweight by highest household social class, holding parental education, demographic characteristics and equivalised income constant. For female children the differences between NS-SEC 1 and 5 (difference is 7.9 Percentage points), and NS-SEC 1 and 6

Table 5.7. The marginal effects for the other aspects of SES (parental education and social class), when all other covariates are held at their means.

	Males			Females		
	Model 3	Model 4	Model 5	Model 3	Model 4	Model 5
Main NVQ 1	reference			reference		
Main NVQ 2	0.027	0.027	0.028	-0.032	-0.030	-0.030
	(0.020)	(0.020)	(0.020)	(0.027)	(0.027)	(0.027)
Main NVQ 3	0.004	0.008	0.007	-0.052	-0.048	-0.045
	(0.022)	(0.021)	(0.021)	(0.030)+	(0.030)	(0.031)
Main NVQ 4	-0.020	-0.014	-0.013	-0.042	-0.032	-0.031
	(0.023)	(0.022)	(0.022)	(0.029)	(0.030)	(0.031)
Main NVQ 5	-0.035	-0.026	-0.022	-0.067	-0.052	-0.051
	(0.028)	(0.028)	(0.029)	(0.037)+	(0.038)	(0.039)
Main overseas	0.047	0.050	0.051	-0.032	-0.025	-0.025
	(0.034)	(0.034)	(0.034)	(0.048)	(0.048)	(0.048)
Main none	-0.005	-0.007	-0.010	-0.024	-0.027	-0.027
	(0.026)	(0.025)	(0.025)	(0.031)	(0.031)	(0.032)
Part NVQ 1	reference			reference		
Part NVQ 2		-0.051	-0.051		-0.010	-0.007
		(0.027)+	(0.027)+		(0.031)	(0.031)
Part NVQ 3		-0.061	-0.061		-0.001	-0.000
		(0.029)*	(0.029)*		(0.037)	(0.037)
Part NVQ 4		-0.059	-0.058		-0.046	-0.042
		(0.029)*	(0.029)*		(0.030)	(0.031)
Part NVQ 5		-0.085	-0.082		-0.065	-0.060
		(0.031)**	(0.031)**		(0.033)*	(0.033)+
Part overseas		-0.040	-0.036		-0.070	-0.069
		(0.040)	(0.041)		(0.037)+	(0.037)+
Part none		-0.004	-0.003		0.014	0.014
		(0.033)	(0.033)		(0.037)	(0.036)
Part missing		-0.003	-0.009		0.022	0.024
		(0.039)	(0.038)		(0.045)	(0.044)
NS-SEC 1	reference			reference		
NS-SEC 2			0.023			0.024
			(0.014)			(0.019)
NS-SEC 3			0.028			-0.012
			(0.021)			(0.024)
NS-SEC 4			0.012			0.028
			(0.022)			(0.027)
NS-SEC 5			0.043			0.079
			(0.038)			(0.041)+
NS-SEC 6			0.033			0.054
			(0.021)			(0.028)+
NS-SEC 7			0.002			0.021
			(0.023)			(0.027)
NS-SEC missing			0.082			0.011
			(0.042)+			(0.039)
N	6,968	6,968	6,968	6,829	6,829	6,829

(difference is 5.4 percentage points) are approaching significance, with children from families classified as NS-SEC 5 & 6 more likely to be classified as overweight.

Changes in income and changes in weight status

The longitudinal aspect of the MCS means that it is possible to consider whether changes in familial income lead to changes in child weight status. There are two types of models for looking at whether changes in familial income influence changes in children's weight status: the fixed effects (FE) model and the random effects (RE) model. The FE method controls for unobserved fixed characteristics of individuals such as their genetic make-up or ability, which, arguably do not change over time. The individual becomes their own control and estimation only uses information from within the same person over time so that time invariant covariates, such as gender, cannot be estimated. This is one method for addressing omitted variable bias, whereby unobserved characteristics may bias estimates for the observed characteristics in a model. However, the FE method cannot account for unobserved characteristics of the individuals which do change over time (Allison, 2009; Maddala and Lahiri, 2009).

The RE method is more efficient because it uses both within-person and between-person variance, whereas the FE method only uses within variance. Within the RE model coefficients for time-invariant covariates can be estimated. Whilst there are benefits to the RE model, it does make more assumptions than the FE model. The RE model assumes that the residuals are normally distributed, and further it assumes that the residuals are not correlated with the other covariates in the model (Clarke *et al.*, 2010). This is a very strict assumption to meet, and in most instances won't be a valid assumption to make. However, these assumptions only need to be met if we are interested in obtaining a causal estimate.

The random effects estimator is a weighted average of the between and within level variance, so that if most of the variance is within people, i.e if there were a lot of time points in which to observe changes, the estimates we get become closer to a fixed effect. However with such a small number of observed time points in the present study, the majority of the variance will be between people, rather than within. For example within a family, weekly equivalised income may increase by £100 over the seven years, but between families the range in incomes will be much larger than this. Common practise is to estimate a fixed effect model and a random effects model and compare the coefficients using a Hausman test (Hausman, 1978). If the Hausman test is significant it suggests that there are systematic differences between the estimates of the FE and RE models, which are likely attributable to violations of the RE assumption, therefore the model which makes the fewest assumption, the FE, should be used. In this particular case, the Hausman test indicated that the FE model was preferable to the RE model[34].

However, Clark and Linzer (2012) argue that the Hausman test is not always a sufficient statistic for deciding between RE and FE specifications. Whilst the Hausman test, can test for bias in the parameter estimates, it does not take into account the trade-off between bias and variance. Clark and Linzer (2012) consider the role of the sample size and number of observations per unit, the level of correlation between the variable of interest and the unit level effects as well as the extent of within unit variation on the estimates obtained from FE and RE models. They suggest that when there is a small amount of within unit variation and with a small number of observations per unit, as is the case here, that a RE model is preferable to a FE model, because the efficiency gained in the estimation is of more benefit than the cost of the bias introduced into the estimates (Clark and Linzer, 2012). Therefore I discuss the results for a RE model in this chapter as well as the results of the FE model.

Unfortunately due to the data that is available, both the fixed effect and random effects model were estimated with transitory income at each time point. This introduces measurement error and most likely downwardly biases the estimated relationship between income and child overweight (Cooper and Stewart, 2013). In order to estimate a FE model, there must be changes in the outcome variable. Therefore children need to change weight status classification i.e from overweight to not overweight or vice versa. The majority of children did not change weight status classification over the time period under study i.e they either stayed in the not

[34] For the Hasuman test both models were estimates as linear probability models

overweight category or stayed in the overweight category between ages 3-7. Therefore a Linear probability model (LPM) was used instead in a conditional FE logit[35]. The linear probability model estimates changes in the probability of being overweight. Separate models were run for males and females. Income was measured as the standardised log of equivalised weekly income and was included as a continuous variable.

The FE LPM models are adjusted for parental age and illness/disability status of the children, because these covariates varied with time and have a plausible association with familial income and child overweight status. The results suggest that for both males (EST=-0.002, SE=0.01, p>0.05, n within=8706, n =21 757) and females (EST=0.001, SE=0.001, p>0.05, n within=8401, n=21 243) there was no evidence of an association between income changes and the probability of being overweight. Standard deviation increases in the log of equivalised weekly income resulted in very small changes in the probability of being classified as overweight. The random effect logit models were also adjusted for child's ethnicity, region of residence, and parental education. There was no evidence that changes in the standardised log of equivalised income were associated with changes in the odds of a child being overweight for males (OR=1.01, SE=0.03, p>0.05, n within=8704, n=21 727) or females (OR=1.04, SE=0.04, p>0.05, n within=8397, n=21 215).

Robustness Check 1: A different measure of low income – income poverty

The definition of poverty is not having enough income or material resources to have a standard of living that is considered acceptable within the society in which they live (Bradshaw *et al.*, 2008). The measure of poverty most commonly used in practise is a measure of relative income poverty, throughout Europe this involves calculating 60% of the median income. Clearly this measure of poverty doesn't match the definition of poverty, and this income based measure has been criticised for being arbitrary (Coudouel, Hentschel and Wodon, 2002). For the purposes of this analysis the relative poverty measure provides another measurement of what constitutes low income.

[35] The FE models were estimated using the conditional FE logit (3300 cases), and with BMI as the outcome variable. The results are consistent with those presented here.

Income poverty is measured in two ways for the poverty analysis; firstly income poverty was calculated from the time-averaged income measure ((0.6*median(308.41) = poverty threshold(185.04)). Secondly, income poverty was defined as being below the OECD income poverty threshold within each sweep of data collection. Sequential logistic regression models were used to investigate the influence of poverty on child obesity and FE models were used to investigate the influence of changes in poverty on changes in the probability of being overweight.

The results for the poverty measure based on time-averaged income are presented in table 5.8. In the unadjusted model the difference in the probabilities are in the expected direction but are not statistically significant, with poverty increasing the probability of a child being classified as overweight. However, after adjustment for demographic characteristics the differences in the probabilities reduce, suggesting no difference in the probability of a child being classified as overweight for those in poverty. After adjustment for main respondent's education, the difference in probability inverts to the opposite direction, so that poverty is associated with a decreased probability of being classified as overweight. The probability of being classified as overweight for children in poverty continues to decrease compared to Children not living in poverty with successive adjustment for partner respondent's education and social class.

The relationship between the incidence of poverty and child weight status was also investigated. The incidence of poverty was coded to reflect the number of sweeps in which the child was classified as living in poverty (0-4), i.e 1 reflects only being in poverty at one time point and 4 reflects being in poverty at all four time points. The differences in the probability of children who were never in poverty being classified as overweight, compared to any other frequency of poverty was small and insignificant after adjustment for parental education. The results from this analysis are available in the appendix J.

Table 5.8. The relationship between poverty (based on permanent income) and child overweight for males and females. Results presented as marginal effects.

Males	Model 1	Model 2	Model 3	Model 4	Model 5
Below poverty line	0.015	-0.004	-0.014	-0.021	-0.023
	(0.013)	(0.016)	(0.017)	(0.017)	(0.017)
	6969	6969	6969	6969	6969
Females	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Below poverty line	0.024	0.003	-0.013	-0.018	-0.027
	(0.015)	(0.017)	(0.018)	(0.017)	(0.018)
	6830	6830	6830	6830	6830

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 1 = unadjusted, Model 2 = Model1 + Ethnicity, longstanding illness/disability, region, parental age; Model 3 = Model 2 + Main respondent education; Model 4 = Model 3 + Partner respondent education; Model 5 = Model 4 + Highest social class of parents.
Not below poverty line is reference category
Estimates weighted using dovwt2 survey design & attrition weight

Again FE analysis was utilised to investigate whether changes in poverty status had any impact on the probability of a child being classified as overweight. The FE analysis also showed that changes in poverty status (i.e a move into poverty from not being classified as in poverty) did not have an effect on children's overweight status. For males a change in poverty status was associated with 0.00 (SE 0.01, p>0.05, n=8734) decrease in the probability of being classified as overweight. For females a change in poverty status was associated with a 0.01 (SE=0.01, P>0.05, N=8419) increase in the probability of being classified as overweight. The results provide little evidence that children in poverty are more likely to be overweight.

Robustness check 2: Comparing coefficients across logistic regression models.

It has been known for some time that directly comparing coefficients across logistic regression models is not always the same as directly comparing coefficients between linear regression models. Changes in the coefficients can not only result from confounding, but can also result from rescaling of the underlying latent variable (Karlson, Holm and Breen, 2012; Mood, 2010). Karlson, Holm & Breen (2012) present a method of making coefficients comparable across sequential logistic regression models. The results presented here have been checked using this method, and there was little discernible difference in the estimates or standard errors. These results, as well as an explanation of this problem, are available in appendix K.

Robustness check 3: Different numbers of income groups

The choice of using income quintiles is relatively arbitrary, quintiles retain large sample sizes within each income group and permit comparisons between relatively low and high income groups, but it is possible that the choice of quintiles is driving the results. The use of quintiles may either be masking patterns for the very lowest income groups, in which case a larger number of income groups would be preferable, or there may not be enough power in each income group to detect differences, in which case a smaller number of income groups would be preferable. Therefore as well as considering income quintiles and poverty as measures of low income, different number of income groupings were considered, and income was entered into the models as a continuous variable. The results, from these analyses, which are available in appendix L, are substantively no different to those presented here.

Robustness Check 4: repetition in another dataset

The reliability of the findings was tested using another recent birth cohort within the UK, the Growing up in Scotland survey (GuS). The Growing up in Scotland survey (GuS), covers a similar time period and population to the MCS and includes children of a similar age group (age 5-6). These results are presented in appendix M. In the GuS dataset, there is very little evidence of a relationship between low income and child overweight also.

Robustness Check 5: Accounting for underweight

One thing that hasn't been considered thus far is the potential for low income families to have increased proportions of children who are classified as underweight. There are BMI based measures for underweight which were created in the same way to the IOTF criteria for overweight, using the same sample (Cole *et al.*, 2007). Multinomial logistic regression was used to include underweight children in the analysis. The outcome variable was then defined as being underweight/ healthy weight/ overweight. Multinomial logit regression was used instead of ordinal logit for two reasons: firstly, descriptive statistics suggested that the relationship between income and weight status is likely nonlinear, with low income increasing the probability of being classified as underweight as well as overweight. Secondly, ordinal logistic regressions work on the assumption of proportional odds, in that the effect of income is assumed to be equal in the underweight, healthy weight and overweight groups. However, the multinomial logit assumes that the groups are unrelated and so information on the ordering of the outcome is lost. The results from the multinomial logit are presented in appendix N and do not differ from those presented here.

Conclusions

There is an expectation that a relationship between child overweight and obesity and low familial income/poverty exists (Anna Soubry, 2013; HM Government, 2011). The results presented here suggest there is, at most, a weak bivariate relationship between familial income and excessive childhood weight, which disappears after adjusting for other aspects of socioeconomic status. In all model specifications, there was no effect of income or poverty after adjustment for parental education. This finding is in agreement with other research (Washbrook, Gregg and Propper, 2013).

With regard to the three potential pathways through which income could influence child outcomes laid out at the start of this chapter: the economic investment model, the parental stress model and the correlation with other parental characteristics, the results here lend themselves to the third explanation. The relationship between income and overweight is explained by the inclusion of other parental characteristics (demographic characteristics and education) in the models. Susan Mayer (1997) also comes to similar conclusions in her work looking at the relationship between income and child outcomes more generally, suggesting that it isn't the amount of income itself, but the parental characteristics that result in low income, or are the result of low income, that explain the relationship we observe between income and child outcomes.

The findings presented here suggest that income is not the driving force behind the social inequalities that exist in child obesity (El-Sayed, Scarborough and Galea, 2012a; Stamatakis, Wardle and Cole, 2010). The results also suggest that changes in familial income or poverty status have little impact on child weight status. However, the FE results need to be interpreted with caution. Due to the problems with measuring income accurately, and the noise which this measurement error introduces, the FE models are likely to have underestimated the impact of changes in income on changes in child weight status (Cooper and Stewart, 2013).

One of the problems in the health literature is the use of different measures of socioeconomic status (Bartley, 2004), with the prevailing belief that these different measures are interchangeable (Bukodi and Goldthorpe, 2012). This paper emphasises that income and SES are not interchangeable concepts and separates the correlation between income and child weight status from the correlation between other aspects of SES and child weight status. The relatively weak bivariate correlation between income and child overweight perhaps suggests that income may not even be a useful proxy for socioeconomic status in this context, if income were to be used as the only measure of SES. This raises questions about its use in other policy making contexts. Indeed, Goldthorpe has put forward a convincing argument for using measures of social class rather than income when interested in an individual's economic situation (Erikson and Goldthorpe, 2009; Erikson and Goldthorpe, 2010; Goldthorpe and McKnight, 2004). This is because social class, as defined by occupations in the manner in which it is in the NS-SEC, contains information on economic stability and the future earning potential of individuals.

Whilst the findings of this analysis are UK specific, there are methodological strengths that can be applied internationally. The distinction between transitory and time-averaged income is not often considered in the sociological literature. But as demonstrated in the results section, time-averaged income is less prone to measurement error (Blanden, Gregg and Macmillan, 2013; Solon, 1992). Also, the potential problems with using same sample sequential logistic regression models are often overlooked (Karlson, Holm and Breen, 2012; Mood, 2010). This paper addresses these problems, by using a Stata package designed by Karlson, Holm & Breen (2012). Furthermore this paper focusses on a particular age group of children which provides specific information for policy making for this age group.

There are limitations to this paper which need to be considered. Firstly, measuring income may be particularly problematic for those who report the lowest incomes. The underreporting of income can result in comparably well off people being grouped in the lowest income category (Brewer, Etheridge and O'Dea, 2013). Also the measure of income does not take into account wealth, such that for example a lower income may result in a better lifestyle for somebody who owns a house outright, than for somebody who is still making mortgage payments or renting. Therefore the differences between low income groups and other parts of the income distribution could be underestimated.

Secondly, the effect of income on children's weight status may not be immediate, but may take time to manifest (Cooper and Stewart, 2013). If changes in familial income do result in changes in diet then it may take time for changes in dietary intake to result in changes in weight status. Time lags in the effect of income are not considered in this paper, but are an obvious candidate for future research. Thirdly, only young children are considered in this paper, but evidence indicates that the importance of income differs with age for different child outcomes (Cooper and Stewart, 2013). Lastly, measures of poverty used in this analysis focussed only on income – as this was the key variable of interest in the paper. But, research points to the importance of using material deprivation, perceived financial situation, and the receipt of means tested benefits as better indicators of relative poverty (Bradshaw, 2001).

The evidence presented here suggests no immediate effect of low income/poverty on child overweight status in the UK. Establishing which aspects of socioeconomic inequality (income, social class, social status, education) are associated with disproportionalities in child obesity is a vital step in understanding and tackling child obesity in the UK. The effects of these different aspects of inequality, including income/income poverty, also need to be considered for different age groups of children.

The focus of this paper was on the relationship between income and child weight status. The inclusion of parental education in the models had the largest impact on the income coefficients, which suggested that the income coefficients were actually picking up on the relationship between parental education and child overweight, due to the correlation between income and parental education. However, the coefficients for the other aspects of SES suggested that there could be a statistically significant relationship between partner respondent's education and child overweight status, even after accounting for demographic characteristics of the parents and child, familial equivalised income, main respondent's education, and household social class. There was little evidence for an independent association between main respondent's education or household social class and child overweight and obesity.

The analysis in chapter 4 showed that the relationship between parental education and child weight status was in the same direction between 1965 and 2007 for children aged 7-10, with increased education resulting in lower child overweight prevalence for both mothers and fathers. The results presented in this chapter suggest that the correlation between parental education and familial income can partially explain the relationship between income and child overweight. Therefore the relationship between parental education and child weight status warrants further investigation. The next chapter focusses upon the relationship between mother's education, father's education and child weight status.

Chapter 6: Parental education and child overweight and obesity, differentiating the role of mothers and fathers

Parents have a large influence on children's eating behaviour and activity levels (Lindsay *et al.*, 2006), the two key aspects for energy balance (Foresight, 2007). Research has demonstrated the importance of parents in children's food consumption and food preferences (Fernandez-Alvira *et al.*, 2013; Patrick and Nicklas, 2005; Scaglioni *et al.*, 2011). Parental influence goes beyond the food they provide and the activities they encourage; the way children are fed, the way the parents present themselves as role models, the foods and activities that parents make available and accessible in the home (Savage, Fisher and Birch, 2007), and the parenting style they adopt all influence children's lifestyles (Rudolf, 2010). Furthermore, specific behaviours linked to obesity in childhood such as the type of foods consumed by the child (Fernandez-Alvira *et al.*, 2013; Wijtzes *et al.*, 2013), the number of calories consumed (Dubois *et al.*, 2011), the amount of physical activity and sedentary behaviour that children partake in (Fairclough *et al.*, 2009; Stalsberg and Pedersen, 2010) are patterned by levels of parental education.

Parental education has an association with child obesity (El-Sayed, Scarborough and Galea, 2012a; Gable and Lutz, 2000; Shrewsbury and Wardle, 2008; Singh, Siahpush and Kogan, 2010). Parents with more years of education, or higher educational qualifications, are less likely to have overweight or obese children. Parental education in the UK context is generally measured by the mother's education (El-Sayed, Scarborough and Galea, 2012a; Shrewsbury and Wardle, 2008), or derived from the highest education level of either the mother or father (Gable and Lutz, 2000; Rohner and Veneziano, 2001; Shrewsbury and Wardle, 2008). Few studies focus upon the role of father's education or consider its effect independently of mother's education (De Vito *et al.*, 1999; Eidsdottir *et al.*, 2013; Giampietro *et al.*, 2002; Klein-Platat *et al.*, 2003; Lamerz *et al.*, 2005; Wake, Hesketh and Waters, 2003). This makes it very difficult to establish what impact, if any, father's education has on child obesity.

The potential influence of father's on children's development has received less attention than the role of mother's. Traditional views about parenthood cast the role of the mother as more important for children's development (Cohen, 1993; Veneziano, 2004), not only because mother's spend a greater amount of time with children (Craig, 2006; Sullivan, 2010), but also because the role of the mother was defined by more caring, emotional and nurturing behaviours, whereas traditional views of fatherhood are as the financial providers and protectors (Hood, 1993; Thompson and Walker, 1989). However, the later part of the twentieth century saw an increased interest in the role of father's in children's development (Radin, 1994), as the emergence of the "new father" meant an increase in male involvement in parenting and the meaning of fatherhood to men (Cohen, 1993).

Evidence indicates that in recent years, fathers are playing a much more active role in childcare and household tasks (Bianchi *et al.*, 2000; Sayer, Bianchi and Robinson, 2004; Sullivan, 2010). Whilst, on average, father's overall time spent with children is still less than mother's, the difference has declined substantially (Sayer, Bianchi and Robinson, 2004; Sullivan, 2010). The increased participation of father's may increase their influence on their children's development. Or the influence of father's may have been unrecognised, due to a focus on mother-child relationships in earlier work (Cohen, 1993; Veneziano, 2004). Nevertheless, there is an emerging literature on the importance of fathers. The developing evidence indicates that father's, or father figures, are indeed influential in children's development, sometimes even more so than mother's (Amato, 1994; Barrera and Garrisonjones, 1992; Grant *et al.*, 2000; Khaleque and Rohner, 2012; Rohner and Veneziano, 2001; Sarkadi *et al.*, 2008).

Fathers may be more influential for son's health outcomes and mother's more influential for daughter's health outcomes (Loureiro, Sanz-de-Galdeano and Vuri, 2010; Perez-Pastor *et al.*, 2009; Thomas, 1994). This may be because children take their same sex parent as a role model for themselves (Korupp, Ganzeboom and Lippe, 2002). Or it may be because parents interact differentially with children of different sexes. Indeed there is evidence to suggest that father's may interact preferentially with sons (Cox *et al.*, 1989; Lamb, 1977a; Lamb, 1977b; Lamb, 1977c; Radin, 1994; Starrels, 1994), spending more time with male offspring than female offspring. The same may be true for mothers and daughters. The literature on eating disorders suggests that the mother's relationship with food, and her perceptions of her body image are the biggest predictors of her daughter's relationship with food and body image (Brown and Ogden, 2004; Lombardo *et al.*, 2012; Wertheim *et al.*, 2002). However, there is a dearth of evidence for these same sex parental-child relationships specific to child obesity, and the evidence which does exist is inconsistent (Freeman *et al.*, 2012; Leary, Davey Smith and Ness, 2010; Mostazir *et al.*, 2013; Perez-Pastor *et al.*, 2009).

The role and influence of father's on child obesity has been highlighted as a crucial gap in the research evidence (Rudolf, 2010). The phenomenon of assortative mating[36] (Mare, 1991; Vandenburg, 1972), means that the measured effect of mother's education on child obesity may to a greater or lesser extent actually reflect the influence of the omitted father's education in previous studies[37]. The aim of this research paper is, therefore, to establish whether father's education has a substantial independent influence on child obesity. This is potentially important information for policy makers and for child obesity interventions, and is an area that needs more research focus and attention.

Congruent with the current literature, it is hypothesised that children with mother's who have higher levels of education will be less likely to be overweight or obese. However it is also hypothesised that there will be an inverse relationship between father's education and child obesity, and that father's education will partially explain the relationship between mother's education and child obesity. It is also hypothesised that the effect of mother's and father's education will differ by the child's gender, so that the effect of mother's education on child obesity will be stronger for female children, and father's education will be stronger for male children.

One important consideration is disentangling the contribution of father's education and father's economic contribution to the household, and the potential impact on child development (Cohen, 1993; Davis and Perkins, 1996; Meyer and Garasky, 1993; Rosenberg and Wilcox, 2006). Parental education is correlated with other aspects of socioeconomic status such as income and social class (Torssande and Erikson, 2010). Indeed some may argue that education to some extent determines social class (Kuha and Goldthorpe, 2010), and income (Blanden, Gregg and Machin, 2002), because higher levels of education increase the likelihood of obtaining high paying, stable

[36] Assortative mating refers to the non-random selection of mates based on similar characterisitcs. In education research it has been shown that people with similar levels of education are much more likely to partner than people with disparate levels of education (see Mare (1991)).
[37] There is variability in parental levels of education. Polychoric correlations suggest that the correlation is 0.53.

occupations with preferential working conditions. Therefore if other aspects of socioeconomic status are not controlled for in the modelling process, any observed effect of father's education conditional upon mother's education may actually reflect the father's economic contributions.

To the best of the author's knowledge there are fewer than a handful of studies considering the impact of father's education on child obesity which go beyond bivariate analysis (Klein-Platat *et al.*, 2003; Lamerz *et al.*, 2005; Wake, Hesketh and Waters, 2003), and these are not based on UK data. The evidence from these studies suggest little to no significant effect of father's education on child obesity, after adjusting for demographic and socioeconomic covariates. It is therefore hypothesised that if there is any effect of father's education on child obesity, it will be explained by his economic contributions to the household measured through income and social class.

As suggested by the hypotheses stated in the text above, the research questions addressed by this paper are:

1) To what extent is the link between mother's education and child obesity distorted by excluding the role of father's?
    a. Do mother's and father's education have an independent association with child obesity?
    b. How does the influence of mother's education compare to the influence of father's education?
2) Does mother's education have a stronger influence on daughter's obesity than sons? And does father's education have a stronger influence on son's obesity than daughters?
3) Is the effect of parental education independent of other socioeconomic and demographic characteristics of the parents?

Method

Sample

The MCS contains responses from a "main" and a "partner" respondent, who are carers for the children. Table 6.1 shows the relationship between these respondents and the cohort member (the child(ren)) at approximately age 7. As shown in table 6.1, in the majority of cases the main and partner respondents are the cohort member's natural mother and father respectively. There tends to be more variation in who completes the partner questionnaire. Dummy variables were created to identify responses from natural mothers and natural fathers. Dummy variables were also used to identify instances where mothers had completed the partner respondent questionnaire, or where fathers had completed the main respondent questionnaire. These were used throughout the data coding to obtain accurate measures of mother's and father's responses.

Table 6.1. The relationship between respondents to the main and partner questionnaires and the cohort member (the child).

	Main Respondent		Partner Respondent	
	n	*%*	*n*	*%*
Not applicable	0	0.00	2,954	21
Natural mother	**13,571**	**97**	**268**	**2**
Natural father	**398**	**3**	**9,878**	**70**
Adoptive mother	**13**	**0**	**1**	**0**
Adoptive father	**1**	**0**	**28**	**0**
Foster mother	**1**	**0**	**1**	**0**
Step mother	**3**	**0**	**41**	**0**
Step father	**6**	**0**	**681**	**5**
Grandmother	41	0	2	0
Grandfather	2	0	17	0
Other, female	3	0	6	0
Other, male	3	0	164	1
Natural Sister	1	0	1	0
Natural Brother	0	0	1	0

Due to the complex nature of modern family structures, measuring the impact of parental education is not a straight forward process. As this paper is specifically interested in the influence of father's education measured independently from mother's education, it is imperative that there be both a mother and father present in the family. Therefore the sample for the main analysis contains only children from two parent households.

The potential for genetics to play some role in the relationship between parental education and child obesity (Bouchard and Loehlin, 2001), suggested that the distinction between biological parents and non-biological parents was important.

Some non-biological parents are present at birth and others come into children's lives much later on. The non-biological parents may be the only parental figure the child has known, but in some cases they are not, making it difficult to determine how much that particular parent has contributed to the child's development. Therefore responses that were not from natural mothers or fathers were not included in the main analysis. Thus the sample in the main analysis only includes children from two parent families, where the mother and father are the biological parents of the child and where they are both resident in the household. However information from non-biological parents is considered in the second part of the paper, where the robustness of the results in different family forms is assessed.

This analysis is concerned with data collected during sweep 4 of the MCS. Out of the 19,177 possible families, there were 17,031 eligible families at sweep 4. Productive responses were achieved from 13,857 (81% of those eligible) families (14,043 children) [38]. After excluding respondents who were not mother's and father's (grandparents, siblings, or other) there were productive responses for 13995 children. Height and weight data were available for 13,765 (98%) of these children. There were responses from 10,135 biological parent pairs, with height and weight data available for 9837 children. Only information from one child per family was included in the analysis, giving a sample size of 9,705 children. The 752 two parent families, where one or both of the parents are a non-biological parent are considered separately.

The characteristics of this sample are described in table 6.2, along with the characteristics of those who are missing from the analysis, because they are missing information on the dependent variable (overweight). There are only 124 children missing overweight status, just over 1% of the two parent biological parent sample. The distribution of characteristics in table 6.2 suggests that the people excluded from the sample are on average more likely to be male children, from an ethnic minority background – particularly Pakistani/Bangladeshi and to be in a worse socioeconomic position than the people included in the sample. However, the level of missingness is very small and is unlikely to introduce bias into the results.

[38] Productive responses are those where there is some information collected on any instrument in sweep 4.

Table 6.2. Characteristics of the sample for analysis, and children not in the sample due to non-response on height or weight at age 7.

	Sample	not in sample
N	9705	124
Weighted N	*9329*	*112*
Gender	**%**	**%**
female	49	35
Overweight	%	
	20	
Mothers Education	**%**	**%**
No qualifications	9	21
ISCED 3C *(GCSE/NVQ 1& 2)*	32	33
ISCED 3A *(A/AS level/ NVQ 3)*	15	15
ISCED 5B *(NVQ 4&5/CertHE/ DipHE)*	16	11
ISCED 5A *(Bacholerette)*	18	11
ISCED 6 *(Masters/PhD)*	6	4
Missing	4	5
Fathers Education	**%**	**%**
No qualifications	7	15
ISCED 3C *(GCSE/NVQ 1& 2)*	28	28
ISCED 3A *(A/AS level/ NVQ 3)*	14	14
ISCED 5B *(NVQ 4&5/CertHE/ DipHE)*	15	8
ISCED 5A *(Bacholerette)*	15	10
ISCED 6 *(Masters/PhD)*	8	2
Missing	13	23
Ethnicity Child	**%**	**%**
white	86	72
mixed	2	4
indian	2	2
pakistani	6	17
black	2	0
other	1	6
Social Class Father	**%**	**%**
Hi Managerial	18	15
Lo Managerial	25	10
Intermediate	5	2
Small employer	18	17
Lower supervisor/Technical	12	11
semi-routine	9	14
Routine	11	18
missing	3	13
Equivalised weekly Income		
mean (SE)	405 (7.7)	326 (20.4)

Indicative levels of qualifications listed next to ISCED levels for reference. This list does not include all qualifications included in the ISCED category. Further information regarding the ISCED measure is available in appendix O.

To take into account the complex sampling design, as well as attrition from the MCS survey over time, data are analysed using the ‘svy’ commands in Stata version 13 (StataCorp, 2013) with the ‘dovwt2’ probability weight. The ‘svy’ commands take into account the clustering units and the strata of the sample in the calculation of standard errors. The probability weights are not only used in the calculation of standard errors but they are also used to obtain correct point estimates (StataCorp, 1985-2013). The weight “dovwt2” reflects the inverse probability of selection into the sample, adjusted for the probability of attrition from the sample over time. The probability of attrition is based on observable characteristics of the families who dropped out of the MCS sample from sweep to sweep (Plewis, 2007a). All analyses presented in the main analysis are analysed using the ‘svy’ commands[39].

Dependent variable

Overweight

Whether the child was overweight was measured at approximately age 7 by the International Obesity Task Force (IOTF) criteria, described in chapter 2 (Cole *et al.*, 2000). Within the sample 20% of children are classified as overweight.

Independent Variables

Parental Education

Mother’s and father’s education was measured independently following the principles of the International Standard Classification of Education (ISCED). ISCED is an instrument for compiling internationally comparable education statistics. The ISCED was chosen because it includes both academic and vocational qualifications within the same framework, much like the NVQ equivalents measure of education used in the previous chapter. But the ISCED also accounts for the different opportunities for further education and different opportunities within the labour market that these different types of qualifications afford (UNESCO-UIS, 2006).

[39] As a robustness check the models with run with just the survey design weight not adjusted for longitudinal attrition from the sample, and an alternative method of dealing with the complex survey design was employed. The results do not noticeably differ from those presented here.

The ISCED 97 covers information pertaining to the level of education and the field of education. The level of education broadly reflects the degree of complexity of the content of the educational programme as well as the knowledge, skills and capabilities required of participants for completion. Within these levels, programmes are further classified by the destinations to which they lead i.e whether they allow access to further education or to the labour market, and by the orientation of the programme, whether it is general, pre-vocational, or vocational.

There are seven levels of education in ISCED 97 from level 0-6; 0 is Pre-primary education, 1 is Primary education, 2 is Lower secondary education, 3 is Upper secondary education, 4 is Post-secondary non-tertiary education, 5 is Tertiary education (first stage), and 6 is Tertiary education (second stage). The ISCED was coded into the MCS data following the OECD guidelines, shown in Appendix O. Information on the parent's highest academic and vocational qualifications, as well as the age at which they finished full time continuous education was used to create the ISCED measure in the MCS. Highest academic and vocational qualifications of the mother and father were coded using information from every sweep of data collection, in which parents reported any new qualifications. Where qualification information was missing, the highest qualification reported at any sweep was utilised as the highest qualification.

As described in appendix O the ISCED levels 5a and 6 cannot be separately coded due to the way in which qualifications were recorded in the MCS, therefore ISCED 6 relates to all postgraduate education. Furthermore the very small numbers of parents with ISCED levels 0 and 1 made it necessary to combine ISCED levels 0-2 into a 'no qualifications' category for analysis. Table 6.3 shows the percentage of mother's who hold different levels of education within father's levels of education. Where mothers and fathers have the same level of education, this is shown by the percentage on the diagonal line, which is lightly shaded. As can be seen in table 6.3, there is variability in parental levels of education (correlation 0.53). For example within father's who have ISCED level 0-2, 33% of mothers have ISCED level 0-2 also, but 37% have ISCED level 3C, 10% have ISCED level 3A, and 14% have ISCED level 5B or above.

As a robustness check, the analyses were repeated with parental education measured using the highest NVQ equivalent of each parent, as well as the highest academic qualification only. The story remains very similar regardless of the measure of parental education utilised. The results using these different measures of parental education are available in Appendix P.

Table 6.3. Cross-tabulation of mother's and father's education level showing the percentage of mothers holding different levels of education within father's level of education.

	F ISCED 0-2	F ISCED 3C	F ISCED 3A	F ISCED 5B	F ISCED 5A	F ISCED 6	F ISCED MISSING	Total
M ISCED 0-2	33	9	6	4	1	1	20	9
M ISCED 3C	37	46	36	29	12	8	34	32
M ISCED 3A	10	17	20	17	10	8	11	15
M ISCED 5B	9	15	18	25	18	13	9	16
M ISCED 5A	4	8	14	17	45	43	7	18
M ISCED 6	1	2	4	5	10	26	2	6
M ISCED MISSING	6	3	1	2	2	1	16	4

M refers to Mother's education level
F refers to father's education level

Other covariates

As suggested in the introduction, parental income and social class are likely to be on the causal pathway between parental education and child overweight (Blanden, Gregg and Machin, 2002; Kuha and Goldthorpe, 2010). Therefore by including them as covariates in the model, the 'net' effect of education is considered, that is the effect of education above any effect on income and social class, rather than what is potentially the 'full' effect of education. This means there is potential for the effect of education to be underestimated. However, without including social class and income as covariates, it is not possible to differentiate the association between parental education and child overweight and other socioeconomic advantages and child overweight, because the education variable would pick up on these correlated but different socioeconomic aspects.

Income

The measure of income used was the OECD equivalised weekly net familial income (Rosenberg, 2012). Time-averaged income was generated by averaging OECD equivalised weekly net income across the four sweep of data collection covering 7-8 years, as described in the previous chapter.

Social Class

Social class was measured using the NS-SEC (Rose, Pevalin and O'Reilly, 2005), which was described in detail in chapter four. Where respondents were out of the labour force, social class was derived from respondents' last known occupation. Both mother's social class and father's social class were included as covariates, because the focus of the analysis was on the individual association between mother's and father's education with the probability of their child being overweight. Social class was measured at sweep 4, when children were approximately 7 years old.

Demographics

The demographic information that was included had to plausibly be linked to parental education and to whether or not a child is overweight. As the measure of parental education is the highest qualification obtained to date, consideration had to be given to what demographic information was associated with educational attainment in school, whether or not they choose to go into higher education and whether they chose to go back into education in adult life. Therefore the following demographic information measured at sweep 4 was included: whether or not the mother or father has longstanding illness/disability, parental age, region of residence, whether child has longstanding illness/disability, number of children in household, and parental report of child's ethnicity.

Parental body shape[40]

Parental obesity is one of the strongest predictors of child obesity (Reilly *et al.*, 2005), regardless of whether parents are the child's natural parents (Garn, Bailey and Cole, 1976). This is because parental BMI likely captures the genetic contribution of parent's body shape to child's body shape, but it also likely captures aspects of the shared familial environment which are pertinent to both the parents and child's body shape. However, education has a strong correlation with adult body shape (Devaux *et al.*, 2011), and there is evidence that this relationship is causal (Brunello, Fabbri and Fort, 2010; Devaux *et al.*, 2011; Kemptner, Jürges and Reinhold, 2011). Therefore including

[40] The term 'body shape' is used throughout this text to refer to all measures of adiposity including body fat, waist circumference, and BMI.

parental body shape as a covariate may reduce the estimated effect of parental education because it is a causal pathway, rather than a confounding variable.

BMI was available for both mother's and father's, and was taken from sweep 1, when children were 9 months olds, due to higher response rates. Measures of parental BMI taken at sweep 1 correlated highly with parental BMI taken at sweep 4 (mother's (0.86), father's (0.83)), suggesting a high degree of stability in parental BMI over time. Correlations between mother's education, father's education, income and social class are presented in table 6.4. The correlations are not excessively high ranging between 0.39 and 0.63 in size. These correlations, as well as collinearity diagnostics[41], suggest that collinearity is not problematic for this analysis (Mela and Kopalle, 2002).

Table 6.4. Correlations between mother's education, father's education, equivalised income, mother's social class and father's social class.

	Mother ISCED	Father ISCED	Income	Mother NS-SEC	Father NS-SEC
Mother ISCED	1.00				
Father ISCED	0.53	1.00			
Income	0.55	0.52	1.00		
mother NS-SEC	-0.61	-0.39	-0.57	1.00	
Father NS-SEC	-0.45	-0.63	-0.61	0.43	1.00

NS-SEC for correlation estimates measured using 3 category version of NS-SEC

Analytic strategy

The data were analysed using Stata version 13 (StataCorp, 2013). Firstly, to establish whether the relationship between mother's education and child overweight status replicated that found in the literature, a logistic regression model was run with child overweight as the outcome variable and mother's education as the predictor variable (model 0). Secondly, father's education was included in the model (model 1) to investigate the independence of mother's education and father's education on child overweight status. This addressed research question 1a, as to whether mother's and

41 Collinearity diagnostics were considered for mother's education, father's education, mother's social class, father's social class and average income. The tolerance of the variables, the percent of variance in the covariate that cannot be accounted for by the other covariates, is reasonably high (0.56-0.65), meaning that a good proportion of the variance in the variables cannot be accounted for by the inclusion of the other variables. The Variance Inflation Factor (VIF) is 1/tolerance. High VIF Numbers indicate that collinearity may be problematic. The VIF is 1.75 on average for these variables, ranging from 1.53 – 1.77, well below the cut off values of between 5-10, in which collinearity may be problematic (Belsley, Kuh and Welsch, 2005). This suggests that these variables are not measuring exactly the same thing and are therefore are not redundant in the regression pathways.

father's education have independent effects[42] on child overweight. Thirdly, the relative strength of the association between mother's education and child overweight, and father's education and child overweight was compared using Wald tests. This showed whether the impact of mother's education on child overweight status and father's education on child overweight status were significantly different, and hence addresses research question 1b.

Fourthly, possible same sex interactions between parents and offspring were investigated by including child's gender in the model as a covariate (model 2). Child's gender was then interacted with mother's education and father's education in the same model. The results of this interaction are reported in the text of the results section.

Lastly, Potential confounding/mediating variables were controlled to establish whether they explain the relationship between parental education and child overweight. Confounding variables are those that have both a relationship with parental educational attainment and child overweight. Mediating variables are those that may lie on the causal pathway between parental education and child overweight status. These covariates were added to the model in blocks, so that the effect of each 'block' on parental education could be assessed. These blocks were added to the main effects model (model 2), and not the interaction model. If the association between mother's education and child overweight, or father's education and child overweight reduced after adding in covariates, this means part of the originally observed effect of mother's/father's education on child overweight was actually the result of the correlation between mother's/father's education and the newly added covariates. The covariate blocks include, demographic characteristics (model 3), socioeconomic characteristics (Social Class: model 4 & Income: model 5) & Parental BMI (Model 6).

To ensure the same sample was analysed within each model, missing dummy variables were used to keep cases with item non-response. For continuous variables this

[42] Whilst the terms "impact" and "effect" are used throughout this paper for convenience, there is no assumption of causality, rather this paper deals with association.

involves setting the value of the missing response to the mean value, and including a variable indicating that this response was missing. For categorical variables, this involves including a missing variable category. These missing dummies do not alter the magnitude of the estimated coefficients.

Directly comparing coefficients across logistic regression models is not the same as directly comparing coefficients between linear regression models (Karlson, Holm and Breen, 2012; Mood, 2010). Changes in the coefficients can not only result from confounding, but can also result from rescaling of the underlying latent variable (Karlson, Holm and Breen, 2012). Karlson, Holm & Breen (2012) suggest a method of making coefficients comparable across nested logistic regression models. The results presented here have been checked using this method, and there was little discernible difference in the estimated odds ratios (OR) or standard errors. These results are available in appendix Q.

Results

To what extent is the link between mother's education and child obesity distorted by excluding the role of fathers?

The results are presented in table 6.5. For the full list of coefficients for all the variables included in the models please see appendix R. Model 0 and Model 1 are used to answer research question 1a & 1b. As can be seen in model 0, there is a reasonably strong relationship between mother's education and the odds of a child being overweight, with higher levels of maternal education associated with decreased odds of a child being overweight. This finding is in agreement with much of the current literature (Anderson, Butcher and Levine, 2002; Nardone *et al.*, 2010). Although it is interesting that only very high levels of maternal education (ISCED 5A OR=0.63 and ISCED 6 OR=0.67) are associated with significantly decreased odds of a child being classified as overweight. The odds ratios for the null model containing paternal education are also presented in appendix R. These show a stronger unadjusted association between father's education and child overweight status than mother's education and child overweight status.

After including paternal education (model 1), the association between mother's education and child overweight status is greatly reduced. There are no statistically significant differences in the odds of a child being overweight by mother's level of education and the difference in the odds for ISCED 5A (OR=0.84) and ISCED 6 (OR=0.92) compared to mother's with no qualifications were substantially reduced. Joint significance tests show that mother's level of qualification has no overall statistically significant association with whether or not a child is overweight ($F(6,383)=1.85$, $p>0.05$)

Table 6.5. Sequentially adjusted logit models showing the relationship between mothers and fathers education (measured by ISCED) and child overweight status. Results presented as odds ratios.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Mother No qualifications	Reference Category						
Mother ISCED 3C	1.01	1.14	1.15	1.14	1.19	1.20	1.14
	(0.11)	(0.13)	(0.13)	(0.14)	(0.15)	(0.15)	(0.14)
Mother ISCED 3A	0.86	1.01	1.03	1.00	1.05	1.06	0.99
	(0.10)	(0.12)	(0.12)	(0.13)	(0.14)	(0.14)	(0.14)
Mother ISCED 5B	0.91	1.11	1.13	1.05	1.09	1.12	1.05
	(0.10)	(0.14)	(0.14)	(0.13)	(0.15)	(0.16)	(0.15)
Mother ISCED 5A	0.63	0.84	0.84	0.78	0.80	0.83	0.84
	(0.08)***	(0.11)	(0.11)	(0.11)+	(0.12)	(0.13)	(0.13)
Mother ISCED 6	0.67	0.92	0.92	0.81	0.82	0.86	0.86
	(0.09)**	(0.14)	(0.14)	(0.13)	(0.14)	(0.15)	(0.15)
Mother missing	1.00	0.97	0.98	0.91	0.93	0.94	0.96
	(0.15)	(0.15)	(0.15)	(0.14)	(0.15)	(0.15)	(0.16)
Father No qualifications	Reference Category						
Father ISCED 3C		0.81	0.81	0.82	0.85	0.85	0.84
		(0.09)+	(0.10)+	(0.10)+	(0.10)	(0.11)	(0.11)
Father ISCED 3A		0.73	0.73	0.74	0.79	0.80	0.80
		(0.09)*	(0.09)*	(0.10)*	(0.10)+	(0.10)+	(0.11)+
Father ISCED 5B		0.74	0.74	0.73	0.79	0.81	0.81
		(0.10)*	(0.10)*	(0.10)*	(0.11)+	(0.11)	(0.11)
Father ISCED 5A		0.64	0.65	0.62	0.69	0.71	0.74
		(0.09)**	(0.09)**	(0.09)***	(0.10)*	(0.10)*	(0.11)*
Father ISCED 6		0.55	0.55	0.53	0.59	0.62	0.69

		(0.09)***	(0.09)***	(0.09)***	(0.10)**	(0.11)**	(0.12)*
Father missing		1.14	1.14	1.15	1.17	1.18	1.12
		(0.14)	(0.14)	(0.15)	(0.15)	(0.16)	(0.17)
	9,705	9,705	9,705	9,705	9,705	9,705	9,705

*+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001*

Model 0 = Mothers education,
Model 1 = Model 0 + Fathers education,
Model 2 = Model 1 + Child's gender
Model 3 = Model 2 + child's ethnicity, whether or not mother, father or child have longstanding illness or disability, mothers age, fathers age, region of residence, and number of children in the household
Model 4 = Model 3 + equivalised time averaged Income
Model 5 = Model 4 + Mother's social class & Father's social class.
Model 6 = Model 5 + Mothers BMI & Fathers BMI,

Conversely, joint significance tests of model 1 show that father's level of education does have an overall significant association with a child being overweight ($F(6,383)=6.00$, $p<0.05$). Model 1 shows a gradient in the relationship between father's level of education and a child's probability of being overweight. The decrease in odds accompanied by increases in father's education are suggestive of a dose response relationship. Children with father's with ISCED levels of 3A (OR=0.73), 5B (OR=0.74), 5A (OR=0.64) and 6 (OR=0.55) are significantly less likely to be classified as overweight than children with father's with no qualifications (level 0-2). Also the difference in the odds of a child being overweight when father's have ISCED level 3C (OR=0.81) compared to father's with no qualifications is significant at the ten percent level ($p=0.07$).

These results could suggest that the associations between mother's education and child overweight status reported elsewhere, are biased due to the omission of father's education (at least in circumstances where both natural parents are present). Therefore the reported relationship between mother's education and child overweight status, to some extent, reflects the correlation between mother's and father's education, and the correlations between father's education and child overweight status. Father's education may be a better indicator of household economic circumstances than mother's education, because mother's tend to be less consistently in the workforce and may take on occupational roles which are more flexible or convenient for childcare purposes, rather than roles which reflect their educational attainment (Lovejoy and Stone, 2012). This is explored in more detail in response to research question 3.

To establish whether father's education level had a significantly stronger association with child overweight status than their mother's education level, the difference between ISCED level 6 and the 'no qualifications' ISCED group (levels 0-2) was compared for mother's and father's. The test revealed a significant difference in the gradients ($F(1,388)=3.89$, $p<0.05$). This suggests that there is a difference in the strength of the association between mother's' and father's education and child overweight status, and that father's education has a stronger association with child overweight than does mother's education. It may be more intuitive to look at the differences in the predicted probabilities for children being classified as overweight with different combinations of parental education levels. In table 6.6 I present the predicted probabilities for four scenarios where mothers and fathers have differing combinations of the highest (ISCED 6) and lowest (no qualification) levels of education.

Table 6.6. The predicted probabilities of child overweight by different combinations of parental education.

Combinations of parental education	Predicted probability	95% confidence intervals
Children overweight on average	0.19	0.18 - 0.20
1: Mother & Father no qualification	0.23	0.19 - 0.27
2: Mother no qualifications & Father ISCED 6	0.14	0.10 - 0.18
3: Mother ISCED 6 & Father no qualifications	0.21	0.16 - 0.27
4: Mother ISCED 6 & Father ISCED 6	0.13	0.10 - 0.16

These predicted probabilities demonstrate a stronger association between father's education level and child overweight conditional upon mother's education level, compared to the association between mother's level of education and child overweight and obesity conditional upon father's education level. Where mother's level of education changes and father's level of education remains fixed (comparing scenario 1 with 3 and 2 with 4) there are very small changes in the predicted probabilities for child overweight and obesity (1 -2 percentage points). However, when father's level of education is altered, but mother's education remains fixed (comparing scenario 1 with 2, and 3 with 4) there are large changes in the predicted probabilities of child overweight (8 - 9 percentage points).

<u>Does mother's education have a stronger influence on daughter's obesity than sons? And does father's education have a stronger influence on son's obesity than daughters?</u>

Gender was included as a covariate in the model (model 2). There was a significant effect of gender, with the odds of females being classified as overweight 1.4 times higher than males. Gender was interacted with mother's and father's education levels. The results from the interaction model are shown in table 6.7. There was no evidence for an interaction between father's education and sex of child ($F(5,384)=0.74$, $p>0.05$)., or mother's education and sex of child ($F(5,384)=1.61$, $p>0.05$). This result was somewhat unexpected given that previous research from similar literatures show that father's had more influence on son's outcomes and mother's on daughter's outcomes

(Loureiro, Sanz-de-Galdeano and Vuri, 2010; Mostazir *et al.*, 2013; Perez-Pastor *et al.*, 2009; Thomas, 1994), and the evidence suggesting that parents interact preferentially with children of the same sex (Cox *et al.*, 1989; Lamb, 1977a; Lamb, 1977b; Lamb, 1977c; Radin, 1994; Starrels, 1994).

A possible explanation for this null finding may be that parental education captures influences on child obesity, which are equally beneficial for male and female children regardless of the potential increased interaction and involvement of parents with the same sex child, or parents of the same sex being particularly strong role models for children. This could be because in two parent households, many of the potential influences on child obesity are at the household or family level, and that individual parental education contributes to these household/family level factors. For example the food available in the household, the leisure activities participated in, the economic stability of the family may be similarly beneficial for male and female children's probability of being overweight.

Table 6.7. Parameter estimates for the full interaction between mother's education and sex of child, and father's education and sex of the child. Presented as odds ratios.

Mother's education	OR	Father's education	OR
Mother No Qualifications	Reference Category	Father No qualifications	Reference category
Mother 3C	1.36 (0.22)+	Father 3C	0.87 (0.15)
Mother 3A	1.32 (0.24)	Father 3A	0.72 (0.14)+
Mother 5B	1.24 (0.22)	Father 5B	0.84 (0.16)
Mother 5A	0.81 (0.16)	Father 5A	0.77 (0.16)
Mother 6	1.05 (0.27)	Father 6	0.60 (0.17)+
Mother Missing	0.83 (0.20)	Father Missing	1.32 (0.25)
Mother 3C#Female	0.74 (0.17)	Father 3C#Female	0.90 (0.21)
Mother 3A#Female	0.64 (0.17)+	Father 3A#Female	1.06 (0.28)
Mother 5B#Female	0.87 (0.24)	Father 5B#Female	0.80 (0.21)
Mother 5A#Female	1.08 (0.31)	Father 5A#Female	0.73 (0.20)
Mother 6#Female	0.81 (0.30)	Father 6#Female	0.87 (0.31)
Mother Missing#Female	1.35 (0.45)	Father Missing#Female	0.77 (0.20)
Female	1.93 (0.49)*		
N	9,705		

$p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

<u>Is the effect of parental education independent of other socioeconomic and demographic characteristics of the parents?</u>

Models 3, 4, 5 and 6 shown in table 6.5 are used to answer research question 3. While it has already been demonstrated that the association between mother's education and child overweight status was largely explained by the inclusion of father's education, it is of interest to know whether father's education has an independent effect from other socioeconomic and demographic characteristics of the family, or whether the effect of father's education is actually picking up on the effect of these previously omitted variables.

After controlling for demographic characteristics (model 3), the odds ratios changed very little. Indeed the association between father's education and child overweight is strengthened slightly by the inclusion of these demographic covariates. This is shown by the odds ratios moving further away from 1, particularly for father's with high levels of education.

Other socioeconomic variables were included, firstly social class (model 4) and then income (model 5). After controlling for mother's and father's social class, the differences in the odds by father's level of education were reduced. For father's with ISCED 3A (OR=0.79) and ISCED 5B (OR=0.79) the odds of having an overweight child were significant at the ten percent level. The statistically significant differences in the odds ratios persisted for father's with ISCED level 5A (OR=0.69) and 6 (OR=0.59), compared to father's with no qualifications. The overall association between father's education and child overweight status remained significant ($F(6, 383)= 3.69$, $p<0.05$). The inclusion of income did further reduce the association between father's education and child overweight status, but only very slightly. The overall association between father's education and child obesity remained statistically significant ($F(6, 383)= 3.27$, $p<0.05$).

As previously discussed it is likely that occupation based social class and income are to some extent a product of the level of education achieved (Kuha and Goldthorpe, 2010), therefore it is unsurprising that the direct effect of father's education on the probability of a child being classified as overweight is weakened by the inclusion of social class and income in the model as these are potential pathways through which parental education influences child obesity. The remaining difference in the odds of a child being overweight by father's level of education conditional upon social class and income suggests that father's education has a role to play in child overweight status above and beyond the influence that father's education has on social class and income. This suggests that father's education is not just associated with child overweight because it is proxying household economic circumstances. Although, the absence of full mediation here could reflect the dominance of education in determining socioeconomic position for health based outcomes (Winkleby *et al.*, 1992).

The final covariates added to the model were mother's and father's BMI. The association between father's education and child obesity weakened, as demonstrated by the reduced difference (shown by the odds ratios moving closer to 1) in the odds of child overweight status. Given the reasonably consistent associations found in the literature between adult body shape and education (Ball and Crawford, 2005; Mackenbach *et al.*, 2008b), as well as the potential pathways through which parents education is expected to influence children's obesity (food preferences, food availability, activities and other lifestyle behaviours), it was expected that some, if not all, of the effect of parental education on child overweight status would be explained by parental body shape. Indeed the relationship between father's education and child obesity was partially attenuated and the overall association between father's education and child overweight was no longer statistically significant ($F(6,383)=1.73$, $p>0.05$). However, there are still statistically significant differences in the odds of a child being overweight for father's with ISCED level 5A (OR=0.74) and 6 (OR=0.69), compared to father's with no qualifications.

Table 6.8. Parameter estimates for the interaction between mother's education and father's education. Presented as odds ratios.

Main Effects	OR	Interactions	OR
Mother No Qualifications		Mother 3C # Father 3C	1.28 (0.44)
Mother 3C	1.14 (0.30)	Mother 3C # Father 3A	0.73 (0.33)
Mother 3A	0.56 (0.21)	Mother 3C # Father 5B	0.47 (0.22)
Mother 5B	0.86 (0.33)	Mother 3C # Father 5A	0.68 (0.52)
Mother 5A	0.78 (0.39)	Mother 3C # Father 6	4.17 (4.38)
Mother 6	1.19 (1.00)	Mother 3A # Father 3C	2.61 (1.10)*
Mother Missing	1.59 (0.59)	Mother 3A # Father 3A	1.50 (0.83)
Father No qualifications		Mother 3A # Father 5B	0.88 (0.49)
Father 3C	0.62 (0.17)+	Mother 3A # Father 5A	0.53 (0.42)
Father 3A	0.91 (0.35)	Mother 3A # Father 6	7.00 (8.17)+
Father 5B	1.36 (0.59)	Mother 5B# Father 3C	1.31 (0.60)
Father 5A	1.22 (0.82)	Mother 5B # Father 3A	1.13 (0.63)
Father 6	0.13 (0.13)*	Mother 5B # Father 5B	0.66 (0.38)
Father Missing	1.02 (0.26)	Mother 5B # Father 5A	0.80 (0.62)
		Mother 5B # Father 6	6.04 (6.47)+
		Mother 5A # Father 3C	1.57 (0.85)
		Mother 5A # Father 3A	0.61 (0.39)
		Mother 5A # Father 5B	0.70 (0.44)
		Mother 5A # Father 5A	0.52 (0.43)
		Mother 5A # Father 6	4.79 (5.31)
		Mother 6# Father 3C	0.99 (0.95)
		Mother 6# Father 3A	0.67 (0.63)
		Mother 6# Father 5B	0.41 (0.42)
		Mother 6# Father 5A	0.39 (0.42)
		Mother 6# Father 6	2.31 (2.98)

Is there an interaction between mother's and father's education level?

The final form of analysis was to consider whether there was an additional benefit from having two highly educated parents in the household as opposed to just one, or whether certain combinations of parental levels of education reduced/increased the risk of child obesity. Given previous suggestions in the results that father's education has a stronger association with child overweight than does mother's education, it was important to investigate whether this finding held across all levels of parental education. For example, it might be expected that having a mother with a very high level of education negates the impact of having a father with a lower level of education.

The interaction terms were initially tested in the 'model 1' specification, whereby only mother's and father's education levels were included in the model. As shown in table 6.8, the interaction was not statistically significant with the ISCED measure of parental education ($F(35, 354)=1.31$, $p>0.05$). However there is a reasonably large amount of categories in the ISCED measure and this may be reducing statistical power, therefore the ISCED measure was collapsed into the following categories: Low (ISCED 0-2 or 'no qualifications'), medium (ISCED 3C and 3A) and high (ISCED 5B, 5A and 6). The interaction between mother's and father's education was tested with the smaller number of categories. Again there was no indication of interaction effects between mother's education and father's education ($F(9,380)=1.01$, $p>0.05$).

Considering different family structures

The analysis was repeated on the two parent families where parents were not both biological parents. In the majority of cases (706), the mother is the natural parent and the father is the non-natural parent. There are 35 cases where the father is the natural parent and 11 cases where neither parent is a natural parent. Overweight status was recorded for 736 children (98%). Due to the much smaller sample size and resulting empty cells in the ISCED measure for the joint distribution of mother's and father's education, parental education was measured using the low, medium & high classification described in the paragraph above.

As with the main analysis I considered whether mother's' and father's' education had independent effects on the probability of a child being classified as overweight. The results are presented in table 6.9. The tables shows the unadjusted relationship between mother's education and child overweight status (model A), father's education and child overweight status (Model B) and the conditional relationship between them (Model C). The results are broadly similar to those in the main analysis, though they are not statistically significant, which is probably due to the small sample size. When both mother's and father's education are included in the same model (Model C), the coefficients on mother's education are reduced, as shown by the odds ratios moving closer to 1 for mother's with middle or high education levels. Also, the coefficients on father's education change very little when adjusted for mother's education. The differences in the odds between low and high levels of education are larger for father's education than mother's education.

Table 6.9. logit models for subsample of non-biological parents showing the relationship between mother's and father's education (measured by ISCED) and child overweight status. Results presented as odds ratios.

	Model A	Model B	Model C
Mother low education	Reference Category		
Mother Medium education	*0.72*		*0.79*
	(0.25)		*(0.28)*
Mother High education	*0.71*		*0.82*
	(0.28)		*(0.34)*
Mother Missing	*0.29*		*0.27**
	(0.18)		*(0.18)*
Father low education	Reference Category		
Father Medium education		*0.83*	*0.85*
		(0.40)	*(0.44)*
Father High education		*0.55*	*0.55*
		(0.29)	*(0.32)*
Father Missing		*0.94*	*1.03*
		(0.48)	*(0.56)*
Constant	*0.29*	*0.26*	*0.32*
	(0.09)	*(0.12)*	*(0.16)*
N (weighted)	*865*	*865*	*865*

+ $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

Model A = Mother's Education
Model B = Father's Education
Model C = Mother's Education + Father's Education

Handling Missing Data

Missingness is a pervasive fact in longitudinal data. Since the early 1990s there has been a tremendous expansion in the availability of computing algorithms to fill-in or impute item non-response under the assumption that these items are ‘missing at random’ (MAR). An alternative algorithm was used in chapter four. As explained in chapter 4, Missing at Random (MAR) assumes that missingness depends on the observed responses to other variables observed. MCAR assumes that missingness does not depend on observed or unobserved characteristics of the individual. It is very rarely the case that responses are missing completely randomly. MNAR assumes that after accounting for all possible information, the missingness is still dependent on the unseen observations. In chapter 4 I applied FIML to account for missing data on the latent variable measure of SES, and I use more ‘ad-hoc’ methods in other areas of this thesis, here I apply multiple imputation.

Unlike FIML, Multiple imputation involves filling in the missing values multiple times and creating complete datasets. Multivariate imputation through chained equations is achieved through the use of the ‘ice’ command in stata. The chained equations approach to multiple imputation is a flexible approach and is well equipped for handling categorical data. The chained equations procedure operates through running a series of regression models, whereby each variable is modelled conditional upon other variables in the data. This means that each variable can be modelled according to it’s own distribution, which is particularly useful for categorical variables. For example, binary variables can be modelled using logit or probit regression, and ordinal variables can be modelled using ordinal logit or probit regression.

Multiple imputation with chained equations leads to unbiased population parameter estimates under the assumptions that the data are Missing at Random (MAR) (Wayman, 2003). This assumption asserts that after controlling for the variables in the imputation model any remaining missingness should be random. The missingness therefore depends only on the observed values and not on unobserved values (Graham, 2009). If the data do not meet the assumptions of MAR, the estimates could be biased. There is no way to fully to assess whether the assumptions of MAR hold, because the information on the missing values isn’t available, but imputation models should include variables which predict the missingess in the variables to be imputed. Typically imputation models contain all the variables that are in the substantive model of interest, but auxiliary variables that predict missingness and are not involved in the model of interest can be incorporated to enhance the imputations (Carpenter and Kenward, 2013).

Multiple imputation was used to estimate the values of the missing responses on the covariates. The amount of missingness is described in table 6.9. Chained equations were used in stata version 13 (StataCorp, 2013) via the 'ICE' command (Royston, 2009). The ICE algorithm works by initially filling all missing values in at random. Variables are sorted so that variables with the smallest amount of missing data are handled first. The first variable 'X_1' with at least one missing value is than regressed on the other variables, '$X_2,\ldots,X_k$. This estimation is restricted to those who do not have missing data on X_1. The missing values on X_1 are replaced with predictions form the regression model by drawing, with replacement, from the posterior predictive distribution of X_1 – the distribution of the missing data conditional upon the observed data. When X_1 is used in the subsequent prediction of the other variables, both the observed and predicted values are used. This is repeated for all the variables in the imputation model in turn. This is called a cycle. In order to stabilise the results this cycle is repeated a specified number of times for each single imputed dataset. The entire procedure is repeated '*M'* times to produce *M* number of imputed data sets (Carpenter and Kenward, 2013; Royston and White, 2011; Van Buuren, Boshuizen and Knook, 1999). In the analysis with the imputed data, the parameter estimates and standard errors are pooled across the imputed data sets using Rubin's rules (Rubin, 1987).

Table 6.10 below provides a summary of missingness for those items which will be used in the substantive analysis. The amount of missingness is small for most covariates (around 5-15%), with the highest amount of missingness in fathers BMI (17%). There were no clear or large patterns to the missingness. 7% of the sample were missing only responses to father's long-term illness/disability status and 6% were missing only responses to mothers BMI. All other patterns were only present in less than 3% of cases. The most common pattern of missingness (60%) was to have full information on all variables.

A major feature of complexity in MCS is the stratified and clustered nature of the design. Following the advice of Carpeneter and Kenward, the survey design and attrition weight (dovwt2) was included and interacted with all covariates in the imputation model (Carpenter, 2010; Carpenter and Kenward, 2013). After imputation, substantive regression analysis was conducted on the pooled data sets using the 'svy' commands in Stata (mi estimates: svy:), which take into account the 9 strata of the MCS, the clustering by electoral wards and the design and attrition weight.

Table 6.10. A description of the extent of missingness for the analysis

Variable	Number missing	Number observed	Missing (%)	Min	Max
M ISCED	453	9,376	5	1	7
F ISCED	1,302	8,527	13	1	7
M BMI	1,397	8,432	14	9.48	59.74
F BMI	1,733	8,096	18	10.78	74.97
M class	772	9,057	8	1	7
F Class	289	9,540	3	1	7
Income	11	9,818	0	40.1775	1206.85
F illness	1,404	8,425	14	0	1
M illness	98	9,731	1	0	1
Child overweight	124	9,705	1	0	1

M refers to Mother's and F refers to Father's
There was no missing data for region of residence, number of children in HH, mother's age, father's age, or child's ethnicity.
No imputed values of income were included in the pooled analysis as there was no missing values for income where all responses to child overweight were present.

The outcome variable, child overweight status, was included in the multiple imputation model (von Hippel, 2007). However, following the advice from von Hippel (2007) the imputed values for child overweight are discarded prior to the pooled analysis of the imputed data sets. The outcome variable is used in the imputation models so that the relationship between the variables to be imputed and the outcome is not underestimated, but the cases with imputed values for the outcome variable are dropped to reduce unnecessary measurement error in the estimates (von Hippel, 2007).

The following specifications were used for the imputation: Father's and mother's education were predicted from household income, their age, their social class, & child overweight status. Fathers and mothers social class were predicted from household income, their age, their education, & child overweight status. These equations were specified because it is well known that the specified variables provide a good prediction of the outcome variables i.e that social class, education and age will provide a good prediction of income. Child overweight status was included in these equations so that the correlation between child overweight status and these covariates was not artificially reduced by excluding the variable from the model. All other variables were predicted by the following (less themselves): household time averaged income, education of mother & father, age of mother and father, ethnicity of the child, child overweight status, number of children in the household, illness/disability status of mother and father and mother's and father's BMI (table 6.11). As a sensitivity check, instead of specifying the prediction equations, all variables were used in the prediction of missingness in all variables. The results from this analysis are reported here also (table 6.12).

Twenty different imputed data sets were created with twenty cycles of the algorithm and the results of the analysis are presented in table 6.11. Simulation studies suggest that with reasonably small amounts of missingness, as is the case here, that twenty imputations will produce incredibly similar results to using 100 imputed data sets with regard to power, the estimated coefficients and standard errors (Graham, Olchowski and Gilreath, 2007). As well as following recommendations for the appropriate number of cycles from by Van Buuren et al (1999) (Van Buuren, Boshuizen and Knook, 1999), trace files were used in stata to assess the number of cycles necessary for convergence. Trace files monitor the convergence of the imputation algorithm.

As a means of sensitivity checking the imputation models were estimated with the MCS with and without the survey design and attrition weights. Five imputation data sets were created instead of twenty, these were used to check how sensitive the results were to the number of imputation data sets created. Different numbers of cycles were also used. These sensitivity analyses can be found in appendix S. The results were very consistent regardless of these sensitivity measures.

Comparisons of table 6.11 and table 6.12 show that whether or not the specific prediction equations are specified, as was the case in table 6.11, or whether all variables were allowed to predict the missing values on all other variables, as in table

6.12, the substantive interpretation of the results is unchanged. The patterning of the odds is the same and the magnitude of the odds ratios and standard errors are reasonably similar. This suggests that adding in information for all the other variables, rather than those just specified in the prediction equations, did not change the predicted estimates a great deal. This is likely because education, social class and income can be reasonably well predicted from each other and age. Most importantly the results from handling the missing data presented in table 6.11 and 6.12, do not differ substantially from those reported in table 6.5, in which missing dummy variables and missing categories were used. The patterning of the results is the same and the magnitude of the coefficients is similar.

Table 6.13 shows the results from an analysis using listwise deletion, specification a, compared to the results using multiple imputation (equations specified, 20 data sets and 20 cycles, as shown in table 6.11), specification b. The first notable difference between the specifications is the sample size. In the fully adjusted model, model 5, there are only 5880 cases with responses to all covariates. In model 0, there is essentially no differences in the estimated odds ratios for mothers education using listwise deletion or multiple imputation, this is because the amount of missingness on mothers education is very small (<5%). However, as the sample size difference between the listwise deletion analysis and the multiple imputation analysis increases, so too do the differences in the predicted odds ratios. This difference is due to the fact that these analyses are taking place on different samples. In listwise deletion, only those who respond to everything are included in the model. People who respond to all questions likely have different characteristics to those who don't respond to every question, therefore the listwise deletion sample is a very bias sample, and becomes more bias the smaller it becomes. This is why in the main analysis missing data categories and dummy variables are included in the models, but also why other methods such as multiple imputation are utilised, to try and reduce the bias in the estimates.

Table 6.11 results from the analysis with the imputed data, where equations were specified.

	Model 0	Model1	Model2	Model3	Model4	Model5
Mother no qualifications	reference					
Mother ISCED 3C	0.99 (0.10)	1.06 (0.12)	1.09 (0.13)	1.13 (0.14)	1.14 (0.14)	1.06 (0.13)
Mother ISCED 3A	0.85 (0.09)	0.95 (0.11)	0.95 (0.12)	1.01 (0.14)	1.02 (0.14)	0.91 (0.13)
Mother ISCED 5B	0.89 (0.10)	1.03 (0.13)	1.00 (0.13)	1.04 (0.15)	1.07 (0.15)	0.97 (0.14)
Mother ISCED 5A	0.63 (0.08)***	0.80 (0.10)+	0.76 (0.11)+	0.78 (0.12)	0.82 (0.13)	0.81 (0.13)
Mother ISCED 6	0.67 (0.10)**	0.88 (0.14)	0.80 (0.13)	0.81 (0.15)	0.85 (0.15)	0.84 (0.15)
Father no qualifications		reference				
Father ISCED 3C		0.82 (0.09)+	0.83 (0.10)	0.86 (0.10)	0.87 (0.10)	0.86 (0.11)
Father ISCED 3A		0.73 (0.09)*	0.74 (0.10)*	0.79 (0.11)+	0.80 (0.11)+	0.81 (0.11)
Father ISCED 5B		0.72 (0.10)*	0.71 (0.10)*	0.77 (0.11)+	0.79 (0.11)+	0.80 (0.11)
Father ISCED 5A		0.62 (0.08)***	0.60 (0.08)***	0.67 (0.09)**	0.70 (0.10)*	0.73 (0.11)*
Father ISCED 6		0.53 (0.09)***	0.50 (0.09)***	0.57 (0.10)**	0.60 (0.11)**	0.67 (0.13)*
N	9,705	9,705	9,705	9,705	9,705	9,705

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 0 = Mothers education,
Model 1 = Model 0 + Fathers education,
Model 2 = Model 1 + child's gender, child's ethnicity, whether or not mother, father or child have longstanding illness or disability, mothers age, fathers age, region of residence, and number of children in the household
Model 3 = Model 2 + Mother's social class & Father's social class
Model 4 = Model 3 + equivalised time averaged Income
Model 5 = Model 4 + Mothers BMI & Fathers BMI,

Table 6.12 results from the analysis with the imputed data, where no equations specified.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Mother no qualifications	reference					
Mother ISCED 3C	0.99 (0.10)	1.07 (0.12)	1.09 (0.13)	1.14 (0.14)	1.14 (0.14)	1.05 (0.13)
Mother ISCED 3A	0.84 (0.09)	0.94 (0.11)	0.95 (0.12)	1.00 (0.14)	1.01 (0.14)	0.90 (0.13)
Mother ISCED 5B	0.89 (0.10)	1.02 (0.12)	1.00 (0.13)	1.03 (0.14)	1.06 (0.15)	0.96 (0.14)
Mother ISCED 5A	0.63 (0.08)***	0.79 (0.10)+	0.76 (0.11)+	0.78 (0.12)	0.80 (0.13)	0.80 (0.13)
Mother ISCED 6	0.66 (0.09)**	0.86 (0.13)	0.78 (0.12)	0.79 (0.14)	0.82 (0.15)	0.81 (0.14)
Father no qualifications	reference					
Father ISCED 3C		0.85 (0.11)	0.86 (0.11)	0.89 (0.12)	0.90 (0.12)	0.90 (0.13)
Father ISCED 3A		0.76 (0.11)*	0.76 (0.11)+	0.82 (0.12)	0.83 (0.12)	0.84 (0.13)
Father ISCED 5B		0.77 (0.12)+	0.75 (0.12)+	0.83 (0.14)	0.85 (0.14)	0.86 (0.15)
Father ISCED 5A		0.66 (0.10)**	0.63 (0.09)**	0.72 (0.12)*	0.74 (0.12)+	0.79 (0.13)
Father ISCED 6		0.56 (0.10)**	0.53 (0.09)***	0.61 (0.11)**	0.64 (0.12)*	0.73 (0.14)
N	9,705	9,705	9,705	9,705	9,705	9,705

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

No equations specified – all variables predicting missingness on all other variables. MI model weighted using MCS svy design and attrition weight "dovwt2". M=20, cycles=20.
Model 0 = Mothers education,
Model 1 = Model 0 + Fathers education,
Model 2 = Model 1 + child's gender, child's ethnicity, whether or not mother, father or child have longstanding illness or disability, mothers age, fathers age, region of residence, and number of children in the household
Model 3 = Model 2 + Mother's social class & Father's social class
Model 4 = Model 3 + equivalised time averaged Income
Model 5 = Model 4 + Mothers BMI & Fathers BMI,

Table 6.13 comparing the coefficients from list wise deletion (spec a) and multiple imputation (spec b).

	M0a	M0b	M1a	M1b	M2a	M2b	M3a	M3b	M4a	M4b	M5a	M5b
Mother no qualification	Reference											
Mother ISCED 3C	1.01	0.99	1.07	1.06	1.14	1.09	1.15	1.13	1.16	1.14	1.00	1.06
	(0.11)	(0.10)	(0.14)	(0.12)	(0.17)	(0.13)	(0.19)	(0.14)	(0.19)	(0.14)	(0.18)	(0.13)
Mother ISCED 3A	0.86	0.85	0.93	0.95	0.96	0.95	0.97	1.01	0.98	1.02	0.77	0.91
	(0.10)	(0.09)	(0.14)	(0.11)	(0.16)	(0.12)	(0.18)	(0.14)	(0.18)	(0.14)	(0.16)	(0.13)
Mother ISCED 5B	0.91	0.89	1.02	1.03	1.02	1.00	1.00	1.04	1.02	1.07	0.85	0.97
	(0.10)	(0.10)	(0.15)	(0.13)	(0.16)	(0.13)	(0.18)	(0.15)	(0.18)	(0.15)	(0.17)	(0.14)
Mother ISCED 5A	0.63	0.63	0.80	0.80	0.79	0.76	0.78	0.78	0.81	0.82	0.73	0.81
	(0.08) ***	(0.08) ***	(0.13)	(0.10) +	(0.15)	(0.11) +	(0.16)	(0.12)	(0.16)	(0.13)	(0.16)	(0.13)
Mother ISCED 6	0.67	0.67	0.80	0.88	0.77	0.80	0.73	0.81	0.76	0.85	0.64	0.84
	(0.09) **	(0.10) **	(0.13)	(0.14)	(0.14)	(0.13)	(0.16)	(0.15)	(0.17)	(0.15)	(0.16) +	(0.15)
Father no qualification	Reference											
Father ISCED 3C			0.85	0.82	0.89	0.83	0.95	0.86	0.96	0.87	1.02	0.86
			(0.10)	(0.09) +	(0.11)	(0.10)	(0.13)	(0.10)	(0.13)	(0.10)	(0.16)	(0.11)
Father ISCED 3A			0.78	0.73	0.79	0.74	0.88	0.79	0.89	0.80	0.92	0.81
			(0.10) +	(0.09) *	(0.11) +	(0.10) *	(0.14)	(0.11) +	(0.14)	(0.11) +	(0.16)	(0.11)
Father ISCED 5B			0.80	0.72	0.79	0.71	0.89	0.77	0.91	0.79	1.01	0.80
			(0.11)	(0.10)	(0.12)	(0.10)	(0.14)	(0.11)	(0.14)	(0.11)	(0.18)	(0.11)

				*		*		+		+		
Father ISCED 5A			0.71	0.62	0.68	0.60	0.80	0.67	0.82	0.70	1.00	0.73
			(0.10) *	(0.08) ***	(0.10) *	(0.08) ***	(0.13)	(0.09) **	(0.14)	(0.10) *	(0.19)	(0.11) *
Father ISCED 6			0.59	0.53	0.56	0.50	0.62	0.57	0.64	0.60	0.81	0.67
			(0.10) **	(0.09) ***	(0.11) **	(0.09) ***	(0.12) *	(0.10) **	(0.13) *	(0.11) **	(0.18)	(0.13) *
	9,259	9,705	8,215	9,705	7,273	9,705	6,944	9,705	6,944	9,705	5,880	9,705

Comparisons between listwise deletion, multiple imputation and the use of missing data categories and dummies can be achieved by comparing the coefficients in this table with the coefficients in table 6.6, which uses missing data categories in the analysis.
Model 0 = Mothers education,
Model 1 = Model 0 + Fathers education,
Model 2 = Model 1 + child's gender, child's ethnicity, whether or not mother, father or child have longstanding illness or disability, mothers age, fathers age, region of residence, and number of children in the household
Model 3 = Model 2 + Mother's social class & Father's social class
Model 4 = Model 3 + equivalised time averaged Income
Model 5 = Model 4 + Mothers BMI & Fathers BMI,

Another concern with the multiple imputation model was that there wasn't sufficient information in the model to predict parental BMI. Standard OLS regression indicated that at most, with all variables in the imputation model, only approximately 3-4% of the variation in observed parental BMI could be predicted. This casts doubt over the ability of the variables in the imputation model to accurately predict missing values in parental BMI[43]. Therefore it may be preferable to consider the results based on imputed data up to the specification of model 4, prior to the inclusion of parental BMI.

Robustness check: Repetition in another data set.

The finding that father's education may have a stronger association than mother's level of education for child overweight and obesity, and the very small association between mother's education and child overweight after accounting for father's education is surprising given the large body of literature showing the importance of the role of mother's education on child development. In order to test whether this finding is specific to the MCS data set the analysis was replicated in the Growing up in Scotland survey (GuS). This is a birth cohort of children born in Scotland, covering a similar time period to the MCS with outcome measures for child overweight when children were of a similar age. The findings, which are presented and discussed below, were largely consistent with the findings in the MCS. The inclusion of father's education reduces the association between mother's education and child overweight status so that it is no longer statistically significant. Furthermore the association between father's education and child overweight status remains statistically significant even after adjustment for demographic characteristics, time averaged income and social class.

The Growing up in Scotland Survey (GuS) is a longitudinal birth cohort survey of children born in 2004/2005. Follow up sweeps were conducted annually, starting when children were 10 months old, through to 2010/2011 when children were aged 5-6. Two parent households were selected for analysis where both natural parents were resident in the household during the first sweep of data collection and when children

[43] Other variables that are associated with BMI were sought to try and rectify this. Smoking status of the parents were included as auxiliary variables in different specifications of the imputation model but they did not change the results.

were aged 5-6 leaving a sample size of 2690. The cross sectional survey design weight from sweep 6 (2010/2011) was applied throughout the analysis.

Child overweight was defined by the IOTF criteria. Parental education was measured in the following categories for mothers and fathers separately:

- Degree or equivalent
- Vocational qualification below degree
- Higher Grade or equivalent *(equivalent to UK A level)*
- Standard Grade or equivalent *(equivalent to UK GCSE)*
- Other
- No qualifications
- Missing

The demographic information included mother's age (banded), father's age (banded), whether or not the child had a long standing illness, mother's ethnicity (white/other ethnic) and father's ethnicity (white/other ethnic). Time averaged income was generated by averaging income across all sweeps of data collection in which respondents reported their income. Mother's and father's social class were measured using the NS-SEC. Only mother's BMI was recorded in the survey. Therefore the final model specification is not comparable to the main analysis and the effect of father's body shape on the association between father's education and child overweight status cannot be assessed. However the demographic and socioeconomic information included in the model specification is very similar to that included in the analysis of the MCS.

The same modelling strategy was used, with covariates added in sequentially in blocks. The results are presented in table 6.14. In model 0 joint significance tests suggest that association between mothers education and child overweight status was significant at the ten percent level ($\chi^2(6)$= 11.41, p=0.08). The odds of a child being classified as overweight and obese were significantly higher when mother's had vocational qualifications (OR=1.37) or higher grade qualifications (OR=1.54), compared to mothers who had degrees. Children whose mothers had standard grade qualifications also had higher odds of overweight and obesity (OR=1.37), and this different was significant at the ten percent level.

However, after controlling for fathers education (model 1) the differences in the odds by mother's level of education are not statistically significant at the five percent level, although the direction of the odds ratios is still in the same direction with higher odds of child overweight and obesity when mothers have qualifications below degree level. Joint significance tests show that the overall association between mothers education and child overweight is no longer significant ($\chi^2(6)=4.79$, $p>0.10$). There are statistically significant differences in the odds of child overweight by father's level of education, with higher odds for fathers with vocational (OR=1.5) or higher grade qualifications (OR=1.9), compared to fathers with degrees. The differences in the odds remain statistically significant after adjustment for child's gender (model 2), other demographic characteristics (model 3), income (model 4), social class (model 5), and mother's BMI (model 6)

Table 6.14. Results from sequential logistic regression model in the GuS showing the relationship between parental education and child overweight. Results presented as odds ratios.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Mother Degree	Reference Category						
Mother VQ^	1.37 (0.16)**	1.23 (0.15)+	1.22 (0.15)	1.21 (0.15)	1.26 (0.16)+	1.18 (0.16)	1.12 (0.16)
Mother HG[a]	1.54 (0.31)*	1.36 (0.28)	1.36 (0.28)	1.33 (0.27)	1.36 (0.27)	1.28 (0.27)	1.25 (0.27)
Mother SG[b]	1.37 (0.23)+	1.23 (0.22)	1.23 (0.23)	1.17 (0.22)	1.23 (0.24)	1.16 (0.24)	1.10 (0.23)
Mother other	2.68 (2.09)	2.31 (1.95)	2.25 (1.89)	2.28 (1.97)	2.46 (2.12)	2.33 (2.03)	2.10 (1.72)
Mother none	1.17 (0.32)	1.07 (0.31)	1.07 (0.30)	0.97 (0.27)	1.05 (0.30)	0.95 (0.28)	0.86 (0.26)
Mother missing	2.80 (3.26)	1.99 (2.42)	1.81 (2.09)	1.74 (1.94)	1.77 (2.02)	1.60 (1.82)	1.62 (2.01)
Father Degree		Reference Category					
Father VQ^		1.50 (0.20)**	1.47 (0.20)**	1.50 (0.20)**	1.55 (0.21)**	1.43 (0.22)*	1.36 (0.21)*
Father HG[a]		1.90 (0.40)**	1.87 (0.39)**	1.84 (0.38)**	1.90 (0.39)**	1.85 (0.39)**	1.82 (0.39)**
Father SG[b]		1.21 (0.22)	1.21 (0.22)	1.23 (0.23)	1.29 (0.25)	1.18 (0.24)	1.08 (0.23)
Father other		0.47 (0.50)	0.52 (0.56)	0.46 (0.50)	0.49 (0.52)	0.49 (0.52)	0.50 (0.54)
Father none		1.19 (0.27)	1.16 (0.26)	1.12 (0.26)	1.19 (0.28)	1.13 (0.29)	1.13 (0.29)
Father missing		1.53 (0.43)	1.53 (0.43)	1.53 (0.44)	1.60 (0.47)	1.47 (0.44)	1.51 (0.46)
N	2,690	2,690	2,690	2,690	2,690	2,690	2,690

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

^VQ Vocational qualification below degree
[a]HG Higher Grade or equivalent
[b]SG standard Grade or equivalent

Model 0 = Mothers education,
Model 1 = Model 0 + Fathers education,
Model 2 = Model 1 + Child's gender
Model 3 = Model 2 + Mother's age, Father's age, whether child has longstanding illness or disability, mother's ethnicity, father's ethnicity.
Model 4 = Model 3 + equivalised time averaged Income
Model 5 = Model 4 + Mother's social class & Father's social class.
Model 6 = Model 5 + Mothers BMI

Conclusion

The results presented here suggest that father's education is a more important correlate of child overweight status than mother's education in two parent families for children around the age of 7. Indeed they suggest that the significant association between mother's education and child overweight is only observed because of the correlation between mother's and father's education levels. This may not be because father's education is a better proxy for household socioeconomic position, as the association between father's education and child overweight persists after adjustment for income and social class. This implies that it is important to include father's education as a predictor of child overweight status for younger children. This is in direct contrast to what happens in the majority of studies, whereby mother's education, but not father's education is included as a covariates (El-Sayed, Scarborough and Galea, 2012a; Shrewsbury and Wardle, 2008).

The role of father's education on child health outcomes has received little attention, despite the increased involvement of father's in two parent families (Cabrera *et al.*, 2000) and the potentially larger role that father's play in their children's development and health. This parallels the situation outlined by Emily Beller (2009), whereby despite the increasing participation of mother's in the labour market, there has been a persistence in using only father's social class to classify a families social position (Beller, 2009). This suggests that research hasn't kept pace with the vast changes that have taken place in family life. This may have consequences regarding conclusions that are drawn.

The evidence presented here suggests that ignoring the influence of father's leads to an exaggeration of the influence of mother's. This suggests that focussing resources solely on the mother may result in smaller gains than expected, and that including fathers in policy responses or child obesity interventions is important. However, without further knowledge of the mechanisms through which father's education is operating it would be premature to make policy suggestions at this stage.

The analysis presented here suggests that the relationship between father's education and child obesity does not operate through father's education having a larger impact on families' economic and financial resources, as this mechanism is tested in the models through the inclusion of social class and income. Indeed the analysis presented here suggests that it may be that some familial behaviour which contributes to child overweight and obesity is more influenced by father's education. There are some potential explanations for the association between father's education and child overweight status. It may be for example, that fathers are more influential on physical activity levels of the children. Fathers are more likely to engage in physically active forms of play with children than mothers, and are more encouraging of physical play (Veneziano, 2004).

Another potential explanation is that, on average, fathers are more influential in decisions regarding the food that is eaten by the family, and how often food is consumed outside the home (McIntosh *et al.*, 2011). For example a study in America found that father's use of fast food restaurants was more predictive of child's use of fast food restaurants than mother's use (McIntosh *et al.*, 2011). People with higher levels of education are more likely to have access to, to use, and to respond to health information (Meara, 2001; Walque, 2004). Therefore fathers with higher levels of education may make healthier decisions regarding food than fathers with lower levels of education. If fathers are more influential in deciding what the family eats, this would lead to the children with fathers with higher levels of education eating healthier food as well.

Alternatively, the influence of father's education on child overweight status may be indirect and may operate through the quality of the mother-father relationship, and father's participation in household tasks and childcare activities, with both being higher for more educated father's (Bianchi *et al.*, 2000; Sayer, Bianchi and Robinson, 2004; Sullivan, 2010; Tampieri, 2010). A healthier spousal relationship likely reduces familial stress and familial conflict, which have been linked to child obesity (Shankardass *et al.*, 2013; Stenhammar *et al.*, 2010). However, with little research in this area it is difficult to establish what the pathways might be. Future research should concentrate on trying to establish the pathway(s) from father's education to the probability of a child being classified as overweight. Establishing these pathways could lead to new insights into potential ways to help tackle child obesity and the inequalities within child obesity.

The findings of this research are limited to children of a specific age group in the UK. Future research could consider whether the role of mother's and father's education on child obesity varies over time. For instance with very young children, mother's education may play a larger role than father's education because in most cases, mother's take the primary responsibility for feeding very young children. Also, mothers may become particularly important for girls during adolescence, as they enter puberty. Another potential extension of this work is to consider whether this finding is UK specific, or whether the role of father's education on child overweight is important in other countries. It is hoped that this research will motivate other researchers interested in intergenerational relationships to use information from both parents, where possible, for family level variables such as socioeconomic status, and to consider the independent influences of mother's and father's in their own research.

Chapter 7: Conclusions

Child obesity has become a serious worldwide public health challenge. Currently, the UK has comparatively high prevalence rates of child obesity (OECD, 2013), with the prevalence increasing threefold since 1980. Research has pointed to severe and potentially long-term health, social, and psychological consequences to being overweight and obese during childhood (Guh *et al.*, 2009; Puhl and Heuer, 2010; Shafer and Ferraro, 2011). The UK government has done much to try and tackle the obesity crisis, with child obesity being high on the policy agenda (Department of Health, 2012; NHS, 2008). However, despite these efforts child obesity rates have remained high over the past decade (Health and Social Care Information Centre, 2013b), with evidence of increasing socioeconomic inequalities in child obesity (Stamatakis *et al.*, 2005; Stamatakis, Wardle and Cole, 2010). The majority of the evidence indicated that children from lower SES families are more likely to be overweight or obese. However, the lack of consistency in the measurement of SES is problematic for understanding the relationship between familial SES and child overweight and obesity.

The lack of uniformity in the measurement of SES, suggests that different indicators of SES are assumed to be interchangeable. This assumption is flawed for several reasons. firstly, in theory the causal mechanisms through which income, social class and education operate should be different (Adler and Newman, 2002). Secondly, evidence indicates that the strength of the relationship between socioeconomic status and child obesity varies depending upon the indicator used (Shrewsbury and Wardle, 2008). Thirdly, other research has shown that different indicators of socioeconomic status can have independent effect on childhood outcomes (Bukodi and Goldthorpe, 2012). Furthermore, because of the use of varying measures of socioeconomic status, it is not possible to determine exactly what aspects of socioeconomic status have a relationship with child obesity. Therefore it is not possible to formulate hypotheses about, and test, the potential mechanisms. This information is essential for the design of effective policies and interventions for reducing inequalities in child obesity

The aim of this thesis was to contribute to the literature on socioeconomic inequalities in obesity by looking specifically at the different aspects of SES and their relationship with child obesity. This thesis sought to answer three main questions regarding the relationship between parental SES and child obesity. 1) How has the relationship between parental SES and child obesity changed over time, and what can this tell us about inequalities in child obesity presently? 2) Is parental income associated with child obesity, and if so what is the size of this relationship? 3) What role do mother's and father's education play in child obesity risk?

The three key overarching findings of this thesis are: 1) that socioeconomic inequalities in child overweight and obesity are being driven by the highest socioeconomic group, not the lowest socioeconomic group, 2) that there are differential relationships between different aspects of SES and child overweight and obesity, and 3) that using socioeconomic information from both parents (where possible) is beneficial for understanding socioeconomic inequalities in child obesity. I discuss these findings in relation to the wider literature and their potential policy implications, and I highlight the future research that is needed. Limitations of this research are discussed in the context of the need for further research.

<u>Socioeconomic inequalities are being driven by the most advantaged, not the most disadvantaged.</u>

The wider literature pointed to strong evidence of socioeconomic inequalities in child obesity (El-Sayed, Scarborough and Galea, 2012a; Shrewsbury and Wardle, 2008), with a widening in socioeconomic inequalities in child obesity over time (Stamatakis *et al.*, 2005; Stamatakis, Wardle and Cole, 2010). Within this literature emphasis is repeatedly placed on the lowest socioeconomic group; with a much higher prevalence of child obesity in the lowest socioeconomic group. The assertion is that the lowest SES group has a disproportionately high prevalence of child overweight and obesity. Therefore it logically follows that reducing socioeconomic inequalities in obesity may result in a reduction in prevalence overall.

In contrast to this, the evidence that I have presented in this thesis does demonstrate socioeconomic differences in childhood overweight and obesity, but the driver of these differences is the highest socioeconomic group. In chapter 4, using more granular measures of SES than those used in previous research, I demonstrate that the widening inequalities in child overweight and obesity is being driven by the relatively low prevalence rates observed for those in the most advantaged social class.

Interestingly, all other social class groups have seen similar increases in child overweight and obesity prevalence over time. I also find the same pattern when I use a composite measure of SES and compare groups of equal sizes, thereby showing that this finding is not driven by the differential distribution on people within the social classes over time.

This pattern is shown again in chapter 5, where I consider the association between familial income and child overweight. The unadjusted association between income and child overweight showed no significant differences between the bottom 80% of the distribution, with much lower prevalence rates for children in families in the top 20% of the income distribution. This suggests that it is not the difference between the highest and the lowest socioeconomic groups that is driving the inequalities in child overweight, but the difference between the highest socioeconomic group and everyone else. From a policy perspective this implies that exclusively focussing upon lower socioeconomic groups may not be effective for reducing child obesity rates overall. This group does have a high prevalence of child obesity, but so do all other socioeconomic groups when compared to the highest SES group. These findings advocate the use of a society wide approach to tackling child obesity.

Further research is needed to discover exactly why the most advantaged in society have such low prevalence rates of child overweight/obesity. Sociological theory suggests that social groups can adopt their own set of norms, values and beliefs in order to distinguish themselves from other social groups. Indeed, Bourdieu (1984) proposed that social class identities are closely tied to food choice, body shapes and lifestyle behaviours (Bourdieu, 1984). Therefore a plausible line of enquiry for explaining the difference between the highest SES group and everyone else is that there are distinct sub-cultural norms, values and beliefs relating to food choice and lifestyle behaviours which are qualitatively different for this group. Exploring the reasons for the low prevalence rates of the highest socioeconomic group will be essential for further understanding inequalities in child obesity.

<u>There are differential relationships between different aspects of SES and child overweight and obesity</u>

A key theme in the thesis is the belief that different aspects of socioeconomic status are not interchangeable. The evidence presented confirms that this is indeed the case. When considering the trends in socioeconomic inequalities in child overweight in chapter 4, I find distinct differences in the trends by social class and parental education. Unlike social class, there was no evidence of changes in the relationship between parental education and child overweight over time. Although, the measure of parental education used has notable limitations. Nevertheless when I use finer grained measures of parental education as covariates in chapter five, and when I focus in greater detail on parental education in chapter six, I find differences in the shape of the relationship compared to income and social class. In particular, the relationship between parental education and child overweight is incremental, and is not driven solely by those with the highest levels of education. Furthermore I find that father's education has an association with child overweight and obesity *independently* from mother's education, social class and income.

As well as showing that these different indicators of social position have differential relationships with obesity and overweight, the work as presented highlights the importance of accounting for more than one aspect of SES in regression models, when focusing upon socioeconomic inequalities. In chapter 5 I argue that the belief of interchangeability of socioeconomic indicators has contributed to the belief that income has a strong relationship with child obesity. Whereas, the evidence suggests a weak bivariate association between income and child overweight and obesity, that falls away after accounting for parental education. Even when I exploit the longitudinal nature of the data and look only at changes within each family over time, I find no evidence of an association between familial income and child overweight and obesity. This might suggest that policies aimed at reducing child obesity by redistributing income to low income families may not be effective at tackling obesity specifically. Although, there is good evidence that increases in income would be beneficial for other child outcomes and hence this could indirectly influence child obesity rates (Cooper and Stewart, 2013).

Using socioeconomic information from both parents (where possible) is beneficial for understanding socioeconomic inequalities in child obesity.

When considering socioeconomic inequalities in child overweight and obesity, it is important to consider information from both parents (where this is possible). Emily Beller (2009) outlined the importance of combining information from both parents in the household when measuring social class. She states that the vast changes in society regarding women's increased participation in the labour market have not been mirrored by their inclusion in measures of household socioeconomic position. The exclusions of women's contribution to the socioeconomic position of the household leads to false conclusions being drawn regarding changes in the relationship between parental social class and children's outcomes. Similarly, I find evidence that the exclusion of father's educational level when measuring parental education and its association with child obesity leads to false conclusions being drawn regarding the relationship between parental education and child overweight and obesity.

The increased participation of women in the labour market has also been accompanied by a redefining of fatherhood and men's roles within the family (Cabrera *et al.*, 2000; Hood, 1993). An emerging literature within child development pointed to the importance of considering the role of fathers, as well as the role of mothers, for childhood outcomes (Rohner and Veneziano, 2001). In chapter 6 I consider the role of mother's and father's education independently. Therefore I considered the relationship between father's education and child overweight and obesity as well as mother's education and child overweight and obesity. I find evidence that father's education is a more important correlate of child overweight and obesity than mother's education. Mother's education may only be associated with child overweight and obesity due to its high correlation with father's education. This is not only because father's education is correlated with the family's financial and economic resources, as this pathway is explicitly tested.

This stronger association between father's education and child overweight and obesity is surprising. Other research has focussed on the role of mother's education, but there has been much less focus on the role of father's education. Therefore the focus of the work presented here was on testing the robustness of this finding. Numerous sensitivity checks, including repetition of the analysis in the Growing Up in Scotland survey, suggested this finding was unlikely to be spurious. The results also indicate that this finding holds for non-biological parents as well as biological parents. This suggests family based obesity interventions and policies should focus upon fathers/father figures as well as mothers.

Further research is needed to explain why father's education has a stronger association with child overweight and obesity. The literature points to fathers with higher levels of education participating more in childcare and household tasks, and having more harmonious relationships with their spouses, which may influence child obesity through a reduction in parental conflict and stress. In order to further this research, these potential mediators need to be tested. Finding out why father's education has such a strong association with child overweight will likely lead to more effective family based policies and interventions to reduce child obesity.

Limitations

Whilst numerous robustness checks have been carried out within each chapter there are potential limitations to this research and consequently the results I have found. Most of these are mentioned within the specific chapters to which they apply, but there are some limitations which are applicable to all the empirical chapters in this thesis. These limitations are discussed below.

The findings presented here are only applicable to children who belong to a particular age group, in chapter four this is children aged approximately 7-10 and in chapters five and six this is children up to the age of 7. It may be that the relationship between different aspects of parental SES and child overweight and obesity vary with age, such that, for instance parental education is particularly important for younger children, but familial income or social class become more important during adolescence. Therefore the findings presented here may only be applicable to children of similar ages. Future research should consider whether the findings presented here differ for children and young people of different ages, so that a picture of socioeconomic inequalities in overweight and obesity throughout childhood can be created. This would allow for the most effective policy responses to socioeconomic inequalities in child overweight and obesity for children of specific ages in childhood.

Not every potential aspect of SES or relative advantage/ disadvantage is considered within this thesis. In many instances SES is operationalised through occupation, income and education, and these are the three key aspects of SES which are considered within chapter four, five and six. However there are other aspects of social and economic wellbeing such as wealth and poverty, which have not been considered in as much detail here. A natural extension of this work for example would be to consider the relationship between wealth and child overweight and poverty and child overweight, controlling for other aspects of SES. It may well be that wealth and poverty interact with different aspects of socioeconomic status to create an even more nuanced picture of socioeconomic inequalities in child overweight and obesity. For example, in chapter five I find no relationship between familial income and child overweight and obesity, but this relationship may vary by familial wealth in terms of the assets they have. A larger amount of wealth in terms of assets and material resources may mean that a lower income is sufficient to achieve a good standard of living. Future research should build on the foundation of the research presented here and include other aspects of socioeconomic status/advantage and investigate the relationships between these different aspects.

A final limitation is that all of the evidence is correlational rather than causal. Without some form unethical intervention randomising access to different forms of education, occupation or manipulating changes in income, it is not possible to obtain causal estimates for the impact of different aspects of SES on child overweight and obesity. However, with the cohort data sets we can observe a great deal about large numbers of individuals. These data hold a rich account of the lives of individuals and permitted a rigorous examination of the influence of a variety of potential confounders. Furthermore, as is the case in chapter five, the longitudinal nature of the datasets can be exploited to get closer to obtaining casual estimates. Nonetheless the work presented here provides a comprehensive account of socioeconomic inequalities in child overweight and obesity, which addressed many of the limitations inherent in the existing literature.

Summary

This thesis provides an in depth investigation of the association between different aspects of SES and child overweight and obesity. The following policy recommendations follow from the findings presented in this thesis: 1) child obesity interventions should not be targeted at the lowest socioeconomic groups, but should be society wide. 2) Family based obesity interventions should focus on the father figure as well as the mother figure (in 2 parent families). 3) Increasing incomes for the lowest income group may not be effective for reducing child obesity.

This thesis contributes to the literature by providing a methodologically robust analysis of the relationship between the different aspects of SES and child overweight and obesity. Further research is needed to explain why these relationships exist. Most notably future research needs to consider why those in the highest socioeconomic group have such a distinctly low prevalence of child overweight and obesity, and why father's education is such a strong correlate of child overweight and obesity.

References

Abdullah, A., Wolfe, R., Stoelwinder, J. U., Courten, M. d., Stevenson, C., Walls, H. L. and Peeters, A. (2011). 'The number of years lived with obesity and the risk of all-cause and cause-specific mortality'. *Int. J. Epidemiol.,* 40 (4)*,* 985-996.

Adler, N. E. and Newman, K. (2002). 'Socioeconomic disparities in health: Pathways and policies'. *Health Affairs,* 21 (2)*,* 60-76.

Agerstrom, J. and Rooth, D. O. (2011). 'The role of automatic obesity stereotypes in real hiring discrimination'. *J Appl Psychol,* 96 (4)*,* 790-805.

Allison, P. D. (2009). *Fixed effects regression models*. London: SAGE.

Amato, P. R. (1994). 'Father-Child Relations, Mother-Child Relations, and Offspring Psychological Well-Being in Early Adulthood'. *Journal of Marriage and the Family,* 56 (4)*,* 1031-1042.

Anderson, P. M., Butcher, K. F. and Levine, P. B. (2002). Maternal employment and overweight children. NBER working paper series.

Andreyeva, T., Puhl, R. M. and Brownell, K. D. (2008). 'Changes in perceived weight discrimination among Americans, 1995-1996 through 2004-2006'. *Obesity,* 16 (5)*,* 1129-1134.

Anna Soubry, Minister of Public Health. (22nd January 2013, 2013). 'Delivering Healthy Growth'. *speech at the Food & Drink Federation conference,* .

Anyaegbu, G. (2010). 'Using the OECD equivalence scale in taxes and benefits analysis'. *Economic & Labour Market Review,* 4 (1)*,* 49-54.

Babey, S. H., Hastert, T. A., Wolstein, J. and Diamant, A. L. (2010). 'Income Disparities in Obesity Trends Among California Adolescents'. *American Journal of Public Health,* 100 (11)*,* 2149-2155.

Ball, K. and Crawford, D. (2005). 'Socioeconomic status and weight change in adults: a review'. *Social Science & Medicine,* 60 (9)*,* 1987-2010.

Bambra, C., Gibson, M., Sowden, A., Wright, K., Whitehead, M. and Petticrew, M. (2010). 'Tackling the wider social determinants of health and health inequalities: evidence from systematic reviews'. *Journal of Epidemiology and Community health,* 64 (4)*,* 284-291.

Barrera, M. and Garrisonjones, C. (1992). 'Family and Peer Social Support as Specific Correlates of Adolescent Depressive Symptoms'. *Journal of Abnormal Child Psychology,* 20 (1)*,* 1-16.

Bartlett, M. S. (1937). 'The statistical conception of mental factors'. *British Journal of Psychology,* 28*,* 97-104.

Bartley, M. (2004). *Health inequality : an introduction to theories, concepts and methods*. Cambridge: Polity Press. Available [Online] at: Table of contents http://www.loc.gov/catdir/toc/ecip041/2003006083.html.

Baum, C. L. and Ford, W. F. (2004). 'The wage effects of obesity: a longitudinal study'. *Health Economics,* 13 (9)*,* 885-899.

BBC NEWS. (2014). *Sugar tax may be necessary, England's chief medical officer says*. [Online]. Available at: http://www.bbc.co.uk/news/health-26442420.

Beller, E. (2009). 'Bringing intergenerational social mobility research into the twenty-first century: why mothers matter'. *American Sociological Review,* 74 (4)*,* 507-528.

Belsley, D. A., Kuh, E. and Welsch, R. E. (2005). *Regression diagnostics: Identifying influential data and sources of collinearity*. (Vol. 571): John Wiley & Sons.

Bianchi, S. M., Milkie, M. A., Sayer, L. C. and Robinson, J. P. (2000). 'Is Anyone Doing the Housework? Trends in the Gender Division of Household Labor'. *Social Forces,* 79 (1)*,* 191-228.

Biddle, S. J. and Asare, M. (2011). 'Physical activity and mental health in children and adolescents: a review of reviews'. *Br J Sports Med,* 45 (11)*,* 886-95.

Bland, J. M. and Altman, D. G. (1997). 'Statistics notes: Cronbach's alpha'. *Bmj,* 314 (7080)*,* 572.
Blanden, J., Gregg, P. and Machin, S. (2002). Education and Family Income. Mimeo: University College London.
Blanden, J., Gregg, P. and Macmillan, L. (2010). 'Intergenerational persistence in income and social class: the impact of increased inequality'. [Online], Working Paper No. 10/230. Available at: http://www.bristol.ac.uk/cmpo/publications/papers/2010/wp230.pdf.
Blanden, J., Gregg, P. and Macmillan, L. (2013). 'Intergenerational persistence in income and social class: the effect of within-group inequality'. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* 176 (2)*,* 541-563.
Bonthuis, M., Jager, K. J., Abu-Hanna, A., Verrina, E., Schaefer, F. and van Stralen, K. J. (2013). 'Application of body mass index according to height-age in short and tall children'. *PLoS One,* 8 (8)*,* e72068.
Bouchard, T. J. and Loehlin, J. C. (2001). 'Genes, evolution, and personality'. *Behav Genet,* 31 (3)*,* 243-273.
Bourdieu, P. (1984). *Distinction : a social critique of the judgement of taste*. London: Routledge & Kegan Paul.
Bradshaw, J. (2001). *Methodologies to measure poverty: More than one is best!* Paper presented at the International Symposium Poverty: Concepts and Methodologies.
Bradshaw, J., Middleton, S., Davis, A., Oldfield, N., Smith, N., Cusworth, L. and Williams, J. (2008). 'A minimum income standard for Britain: what people think'.
Brewer, M., Etheridge, B. and O'Dea, C. (2013). 'Why are households that report the lowest incomes so well-off?'.
Brewer, R. I. (1986). 'A Note on the Changing Status of the Registrar Generals Classification of Occupations'. *British Journal of Sociology,* 37 (1)*,* 131-140.
Brophy, S., Cooksey, R., Gravenor, M. B., Mistry, R., Thomas, N., Lyons, R. A. and Williams, R. (2009). 'Risk factors for childhood obesity at age 5: Analysis of the Millennium Cohort Study'. *BMC Public Health,* 9.
Brown, R. and Ogden, J. (2004). 'Children's eating attitudes and behaviour: a study of the modelling and control theories of parental influence'. *Health Educ Res,* 19 (3)*,* 261-71.
Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York ; London: Guilford Press.
Brunello, G., Fabbri, D. and Fort, M. (2010). 'Years of schooling, human capital and the body mass index of European females'.
Bukodi, E. and Goldthorpe, J. H. (2012). 'Decomposing 'Social Origins': The Effects of Parents' Class, Status, and Education on the Educational Attainment of Their Children'. *Eur Sociol Rev* 0(0)*,* 1-16.
Burnham, K. P. and Anderson, D. R. (2002). *Model selection and multimodel inference: a practical information-theoretic approach*: Springer.
Butler, N., Despotidou, S. and Shepherd, P. (1980). *1970 British Cohort Study Ten-year Follow-up: Guide to data available at the ESRC Data Archive*. City university, London: Social Statistics Research Unit.
Butler, N. R. (1980). Child Health and Education Study: Medical Examination From. University of Bristol: Department of Child Health Research Unit.
Byrne, B. M. (2011). *Structural equation modeling with Mplus : basic concepts, applications, and programming*. New York: Routledge Academic.
Cabrera, N. J., Tamisk-LeMonda, C. S., Bradley, R. H., Hofferth, S. and Lamb, M. E. (2000). 'Fatherhood in the twenty-first century'. *Child Development,* 71 (1)*,* 127-136.
Caird, J., Kavanagh, J., Oliver, K., Oliver, S., O'Mara, A., Stansfield, C. and Thomas, J. (2011). *Childhood obesity and educational attainment: a systematic review*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
Caliendo, M. and Lee, W.-S. (2013). 'Fat chance! Obesity and the transition from unemployment to employment'. *Economics & Human Biology,* 11 (2)*,* 121-133.

Carpenter, J. R. (2010). 'Multilevel multiple imputation allowing for survey weights'. [Online]. Available at: www.missingdata.org.uk. [Last accessed 26th february, 2014].
Carpenter, J. R. and Kenward, M. G. (2013). *Multiple imputation and its application*. Chichester: Wiley.
Cattell, R. B. (1983). 'Citation Classic - the Scree Test for the Number of Factors'. *Current Contents/Social & Behavioral Sciences,* (5)*,* 16-16.
Chinn, S. and Rona, R. J. (2001). 'Prevalence and trends in overweight and obesity in three cross sectional studies of British Children, 1974-94'. *BMJ,* 322 (7277)*,* 24-6.
Church, T. S., Thomas, D. M., Tudor-Locke, C., Katzmarzyk, P. T., Earnest, C. P., Rodarte, R. Q., Martin, C. K., Blair, S. N. and Bouchard, C. (2011). 'Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity'. *PLoS One,* 6 (5)*,* e19657.
Clark, T. S. and Linzer, D. A. (2012). Should I Use Fixed or Random Effects? The Society for Political Methodology.
Clarke, P., Crawford, C., Steele, F. and Vignoles, A. F. (2010). 'The choice between fixed and random effects models: some considerations for educational research'.
Cohen, T. F. (1993). 'What do fathers provide?: Reconsidering the economic and nurturant dimensions of men as parents.'. In J. C. Hood (Ed.), *Men, work, and family*. London: SAGE publications.
Cole, T. J. (2004). 'Children grow and horses race: is the adiposity rebound a critical period for later obesity?'. *BMC pediatrics,* 4 (1)*,* 6.
Cole, T. J., Bellizzi, M. C., Flegal, K. M. and Dietz, W. H. (2000). 'Establishing a standard definition for child overweight and obesity worldwide: international survey'. *BMJ,* 320 (7244)*,* 1240-3.
Cole, T. J., Flegal, K. M., Nicholls, D. and Jackson, A. A. (2007). 'Body mass index cut offs to define thinness in children and adolescents: international survey'. *BMJ,* 335 (7612)*,* 194.
Cole, T. J., Freeman, J. V. and Preece, M. A. (1995). 'Body-Mass Index Reference Curves for the Uk, 1990'. *Archives of disease in childhood,* 73 (1)*,* 25-29.
Conger, R. D., Conger, K. J. and Martin, M. J. (2010). 'Socioeconomic Status, Family Processes, and Individual Development'. *Journal of Marriage and Family,* 72 (3)*,* 685-704.
Conger, R. D., Ge, X. J., Elder, G. H., Lorenz, F. O. and Simons, R. L. (1994). 'Economic-Stress, Coercive Family Process, and Developmental Problems of Adolescents'. *Child Development,* 65 (2)*,* 541-561.
Connelly, R. (2011). Drivers of Unhealthy Weight in Childhood: Analysis of The Millennium Cohort Study. Edinburgh: Queens Printers of Scotland.
Cooper, K. and Stewart, K. (2013). 'Does money affect children's outcomes?'. *A systematic review. Joseph Rowntree Foundation*.
Costa-Font, J. and Gil, J. (2013). 'Intergenerational and socioeconomic gradients of child obesity'. *Social science & medicine,* 93*,* 29-37.
Coudouel, A., Hentschel, J. S. and Wodon, Q. T. (2002). 'Poverty measurement and analysis'. *A Sourcebook for poverty reduction strategies,* 1*,* 27-74.
Cox, M. J., Owen, M. T., Lewis, J. M. and Henderson, V. K. (1989). 'Marriage, Adult Adjustment, and Early Parenting'. *Child Dev,* 60 (5)*,* 1015-1024.
Craig, L. (2006). 'Does father care mean fathers share? A comparison of how mothers and fathers in intact families spend time with children'. *Gender & Society,* 20 (2)*,* 259-281.
Daniels, S. R. (2006). 'The consequences of childhood overweight and obesity'. *Future of Children,* 16 (1)*,* 47-67.
Davis, J. E. and Perkins, W. E. (1996). *Fathers' Care: A Review of the Literature*. Philadelphia: Univeristy of Pennsylvania.
De Vito, E., La Torre, G., Langiano, E., Berardi, D. and Ricciardi, G. (1999). 'Overweight and obesity among secondary school children in Central Italy'. *European Journal of Epidemiology,* 15 (7)*,* 649-654.

Department of Health (2012). *An update on the government's approach to tackling obesity*. National Audit Office.
Devaux, M., Sassi, F., Church, J., Cecchini, M. and Borgonovi, F. (2011). 'Exploring the Relationship Between Education and Obesity'. *OECD Journal: Economic Studies,* 2011 (1), 5.
Dietz, W. H. (1998). 'Health consequences of obesity in youth: Childhood predictors of adult disease'. *PEDIATRICS,* 101 (3), 518-525.
Donkin, A., Goldblatt, P. and Lynch, K. (2002). 'Inequalities in life expectancy by social class, 1972–1999'. *Health Statistics Quarterly,* 15, 5-15.
Dubois, L., Farmer, A., Girard, M., Burnier, D. and Porcherie, M. (2011). 'Demographic and socio-economic factors related to food intake and adherence to nutritional recommendations in a cohort of pre-school children'. *Public Health Nutr,* 14 (6), 1096-104.
Durantauleria, E., Rona, R. J. and Chinn, S. (1995). 'Factors Associated with Weight-for-Height and Skinfold Thickness in British Children'. *Journal of Epidemiology and Community Health,* 49 (5), 466-473.
Durnin, J. and Rahaman, M. M. (1967). 'The assessment of the amount of fat in the human body from measurements of skinfold thickness'. *British Journal of Nutrition,* 21 (03), 681-689.
Eagle, T. F., Sheetz, A., Gurm, R., Woodward, A. C., Kline-Rogers, E., Leibowitz, R., DuRussel-Weston, J., Palma-Davis, L., Aaronson, S., Fitzgerald, C. M., Mitchell, L. R., Rogers, B., Bruenger, P., Skala, K. A., Goldberg, C., Jackson, E. A., Erickson, S. R. and Eagle, K. A. (2012). 'Understanding childhood obesity in America: Linkages between household income, community resources, and children's behaviors'. *American Heart Journal,* 163 (5), 836-843.
Edwards, P., Roberts, I., Green, J. and Lutchmun, S. (2006). 'Deaths from injury in children and employment status in family: analysis of trends in class specific death rates'. *BMJ,* 333 (7559), 119.
Egger, G. and Swinburn, B. (1997). 'An "ecological" approach to the obesity pandemic'. *BMJ,* 315 (7106), 477-80.
Eidsdottir, S. P., Kristjansson, A. L., Sigfusdottir, I. D., Garber, C. E. and Allegrante, J. P. (2013). 'Secular trends in overweight and obesity among Icelandic adolescents: Do parental education levels and family structure play a part?'. *Scandinavian Journal of Public Health,* 41 (4), 384-391.
El-Sayed, A. M., Scarborough, P. and Galea, S. (2012a). 'Socioeconomic inequalities in childhood obesity in the United Kingdom: a systematic review of the literature'. *Obes Facts,* 5 (5), 671-92.
El-Sayed, A. M., Scarborough, P. and Galea, S. (2012b). 'Unevenly distributed: a systematic review of the health literature about socioeconomic inequalities in adult obesity in the United Kingdom'. *BMC Public Health,* 12 (18).
Enders, C. K. (2010). *Applied missing data analysis*. New York ; London: Guilford.
Enders, C. K. and Bandalos, D. L. (2001). 'The Relative Performance of Full Information Maximum Likelihood Estimation for Missing Data in Structural Equation Models'. *Structural Equation Modeling-a Multidisciplinary Journal,* 8 (3), 430-457.
Erikson, R. and Goldthorpe, J. H. (1992). *The constant flux : a study of class mobility in industrial societies*: Clarendon.
Erikson, R. and Goldthorpe, J. H. (2009). *Income and Class Mobility Between Generations in Great Britain: The Problem of Divergent Findings from the Data-sets of Birth Cohort Studies*: Swedish Institute for Social Research.
Erikson, R. and Goldthorpe, J. H. (2010). 'Has social mobility in Britain decreased? Reconciling divergent findings on income and class mobility'. *British Journal of Sociology,* 61 (2), 211-230.

Fairclough, S. J., Boddy, L. M., Hackett, A. F. and Stratton, G. (2009). 'Associations between children's socioeconomic status, weight status, and sex, with screen‐based sedentary behaviours and sport participation'. *International journal of pediatric obesity,* 4 (4)*,* 299-305.

Fernandez-Alvira, J. M., Mouratidou, T., Bammann, K., Hebestreit, A., Barba, G., Sieri, S., Reisch, L., Eiben, G., Hadjigeorgiou, C., Kovacs, E., Huybrechts, I. and Moreno, L. A. (2013). 'Parental education and frequency of food consumption in European children: the IDEFICS study'. *Public Health Nutr,* 16 (3)*,* 487-98.

Field, A. P. (2009). *Discovering statistics using SPSS : (and sex and drugs and rock 'n' roll)*. (3rd ed. ed.). London: SAGE.

Finch, N. (2002). 'Demographic Trends in the UK'. *Social Policy Research Unit, University of York, York*.

Flegal, K. M., Ogden, C. L., Yanovski, J. A., Freedman, D. S., Shepherd, J. A., Graubard, B. I. and Borrud, L. G. (2010). 'High adiposity and high body mass index-for-age in US children and adolescents overall and by race-ethnic group'. *American Journal of Clinical Nutrition,* 91 (4)*,* 1020-1026.

Foresight. (2007). *Tackling obesities: future choices. Project report*. [London]: Dept. of Innovation, Universities and Skills.

Freedman, D. S., Khan, L. K., Serdula, M. K., Srinivasan, S. R. and Berenson, G. S. (2000). 'Secular trends in height among children during 2 decades - The Bogalusa heart study'. *Archives of Pediatrics & Adolescent Medicine,* 154 (2)*,* 155-161.

Freedman, D. S., Ogden, C. L., Flegal, K. M., Khan, L. K., Serdula, M. K. and Dietz, W. H. (2007). 'Childhood overweight and family income'. *MedGenMed,* 9 (2)*,* 26.

Freedman, D. S. and Sherry, B. (2009). 'The Validity of BMI as an Indicator of Body Fatness and Risk Among Children'. *Pediatrics,* 124*,* S23-S34.

Freedman, D. S., Thornton, J. C., Mei, Z. G., Wang, J., Dietz, W. H., Pierson, R. N. and Horlick, M. (2004). 'Height and adiposity among children'. *Obesity Research,* 12 (5)*,* 846-853.

Freedman, D. S., Wang, J., Maynard, L. M., Thornton, J. C., Mei, Z., Pierson, R. N., Dietz, W. H. and Horlick, M. (2005). 'Relation of BMI to fat and fat-free mass among children and adolescents'. *International Journal of Obesity,* 29 (1)*,* 1-8.

Freeman, E., Fletcher, R., Collins, C. E., Morgan, P. J., Burrows, T. and Callister, R. (2012). 'Preventing and treating childhood obesity: time to target fathers'. *Int J Obes (Lond),* 36 (1)*,* 12-5.

Friedman, M. (1957). *A Theory of the Consumption Function, etc*. Princeton: Princeton University Press.

Fryers, T., Melzer, D. and Jenkins, R. (2003). 'Social inequalities and the common mental disorders - A systematic review of the evidence'. *Social Psychiatry and Psychiatric Epidemiology,* 38 (5)*,* 229-237.

Gable, S. and Lutz, S. (2000). 'Household, parent, and child contributions to childhood obesity'. *Family Relations,* 49 (3)*,* 293-300.

Gadermann, A. M., Guhn, M. and Zumbo, B. D. (2012). 'Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide'. *Practical Assessment, Research & Evaluation,* 17 (3)*,* 1-13.

Gallagher, J. (2013). *Tax fizzy drinks and ban junk food ads, say doctors* [Online]. [Last accessed 24th March].

Gallie, D. (2001). Skill change and the labour market: gender, class and unemployment, *In National Institute for Economic and Social Research conference Disadvantage in the Labour Market: Diversity and Commonality in Causes, Consequences and Redress.*

Garn, S. M., Bailey, S. M. and Cole, P. E. (1976). 'Similarities between Parents and Their Adopted-Children'. *American Journal of Physical Anthropology,* 45 (3)*,* 539-543.

Garrow, J. S. (1978). *Energy balance and obesity in man*. (2nd ed. completely revised. ed.). Amsterdam ; Oxford: Elsevier : North-Holland Biomedical Press.

Gatineau, M. and Mathrani, S. (2011). *Obesity and ethnicity*. Oxford: National Obesity Observatory.

Geyer, S., Hemstrom, O., Peter, R. and Vagero, D. (2006). 'Education, income, and occupational class cannot be used interchangeably in social epidemiology. Empirical evidence against a common practice'. *Journal of Epidemiology and Community Health,* 60 (9)*,* 804-810.

GfK NOP. 'Millennium Cohort Study Sweep 2: Technical Report'. [Online]. Available at: http://www.cls.ioe.ac.uk/studies.asp?section=00010002000100040006.

Giampietro, O., Virgone, E., Carneglia, L., Griesi, E., Calvi, D. and Matteucci, E. (2002). 'Anthropometric indices of school children and familiar risk factors'. *Prev Med,* 35 (5)*,* 492-498.

Gigante, D. P., Victora, C. G., Matijasevich, A., Horta, B. L. and Barros, F. C. (2013). 'Association of family income with BMI from childhood to adult life: a birth cohort study'. *Public Health Nutrition,* 16 (2)*,* 233-239.

Goldin, C. (1991). 'The role of World War II in the rise of women's work'. *The American Economic Review,* 81 (4)*,* 741-756.

Goldman, R. F. and Buskirk, E. R. (1961). 'Body volume measurement by underwater weighing: description of a method'. *Techniques for measuring body composition,* 78-89.

Goldthorpe, J. H. (2012). 'Back to Class and Status: Or Why a Sociological View of Social Inequality Should Be Reasserted'. *Revista Espanola De Investigaciones Sociologicas,* (137)*,* 43-58.

Goldthorpe, J. H. and McKnight, A. (2004). *The economic basis of social class*. London: Centre for Analysis of Social Exclusion.

Goodman, A. and Butler, N. R. (1986). *BCS70 - The 1970 British Cohort Study: The Sixteen-year Follow-up*. London: Social Statistics Research Unit, City University.

Graham, J. W. (2009). 'Missing Data Analysis: Making It Work in the Real World'. *Annual Review of Psychology,* 60*,* 549-576.

Graham, J. W., Olchowski, A. E. and Gilreath, T. D. (2007). 'How many imputations are really needed? Some practical clarifications of multiple imputation theory'. *Prevention Science,* 8 (3)*,* 206-213.

Grant, K. E., O'Koon, J. H., Davis, T. H., Roache, N. A., Poindexter, L. M., Armstrong, M. L., Minden, J. A. and McIntosh, J. M. (2000). 'Protective factors affecting low-income urban African American youth exposed to stress'. *Journal of Early Adolescence,* 20 (4)*,* 388-417.

Gray, J. C., Gatenby, R. and Huang, Y. (2010). *Millennium Cohort Study Sweep 4: Technical Report*. London: NatCen.

Gray, J. C., Gatenby, R., Simmonds, N. and Huang, Y. (2010). *Millennium Cohort Study Sweep 4: Technical Report*: NatCen.

Greenaway, D. and Haynes, M. (2003). 'Funding higher education in the UK: The role of fees and loans'. *Economic Journal,* 113 (485)*,* F150-F166.

Greenleaf, C., Petrie, T. A. and Martin, S. B. (2014). 'Relationship of Weight ‐ Based Teasing and Adolescents' Psychological Well ‐ Being and Physical Health'. *Journal of School Health,* 84 (1)*,* 49-55.

Greenlund, K. J., Liu, K., Dyer, A. R., Kiefe, C. I., Burke, G. L. and Yunis, C. (1996). 'Body mass index in young adults: Associations with parental body size and education in the CARDIA study'. *American Journal of Public Health,* 86 (4)*,* 480-485.

Griffiths, L., Hawkins, S., Cole, T., Law, C. and Dezateux, C. (2010a). Childhood overweight and obesity. Millennium Cohort Study Briefing 14. University of London Institute of Education.

Griffiths, L. J., Cole, T., Law, C. and Dezateux, C. (2010b). 'Childhood Overweight and Obesity'. In K. Hansen, H. Joshi and S. Dex (Eds), *Children of the 21st century : the first five years*. Bristol: Policy Press.

Guh, D. P., Zhang, W., Bansback, N., Amarsi, Z., Birmingham, C. L. and Anis, A. H. (2009). 'The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis'. *BMC Public Health,* 9*,* 88.

Gutierrez, R. G., Linhart, J. M. and Pitblado, J. S. (2003). 'From the help desk: Local polynomial regression and Stata plugins.'. *Stata Journal,* 3*,* 412-419.

Haines, J., Hannan, P. J., den Berg, P., Eisenberg, M. E. and Neumark‐Sztainer, D. (2013). 'Weight‐related teasing from adolescence to young adulthood: Longitudinal and secular trends between 1999 and 2010'. *Obesity,* 21 (9)*,* E428-E434.

Hannon, T. S., Rao, G. and Arslanian, S. A. (2005). 'Childhood obesity and type 2 diabetes mellitus'. *PEDIATRICS,* 116 (2)*,* 473-480.

Hansen, K., Johnson, J., Joshi, H., Calderwood, L., Jones, E., McDonald, J., Shepherd, P. and Smith, K. (2010). *Millennium Cohort Study First, Second, Third and Fourth Surveys: A Guide to the Datasets*. London: CLS, Institute of Education.

Haralambos, M. and Holborn, M. (2008). *Sociology : themes and perspectives*. (7th ed. ed.). London: Collins.

Harper, B. (2000). 'Beauty, Stature and the Labour Market: A British Cohort Study'. *Oxford Bulletin of Economics and Statistics,* 62*,* 771-800.

Hausman, J. A. (1978). 'Specification Tests in Econometrics'. *Econometrica,* 46 (6)*,* 1251-1271.

Hawkins, S. S., Cole, T. J., Law, C. and Hlth, M. C. S. C. (2009). 'An ecological systems approach to examining risk factors for early childhood overweight: findings from the UK Millennium Cohort Study'. *Journal of Epidemiology and Community Health,* 63 (2)*,* 147-155.

Health and Social Care Information Centre (2013a). *National Child Measurement Programme: England, 2012/13 school year*. London: Public Health England.

Health and Social Care Information Centre (2013b). *Statistics on Obesity,Physical Activity and Diet: England, 2013*: Public Helath England.

Herman, K. M., Craig, C. L., Gauvin, L. and Katzmarzyk, P. T. (2009). 'Tracking of obesity and physical activity from childhood to adulthood: the Physical Activity Longitudinal Study'. *Int J Pediatr Obes,* 4 (4)*,* 281-8.

Hill, J. O. and Peters, J. C. (1998). 'Environmental contributions to the obesity epidemic'. *Science,* 280 (5368)*,* 1371-1374.

HM Government. (2011). Healthy Lives, Healthy People: our strategy for public health in England. In D. o. Health (ed). London: COI.

Hollar, D., Messiah, S. E., Lopez-Mitnik, G., Hollar, T. L., Almon, M. and Agatston, A. S. (2010). 'Effect of a two-year obesity prevention intervention on percentile changes in body mass index and academic performance in low-income elementary school children'. *American Journal of Public Health,* 100 (4)*,* 646-53.

Hood, J. C. (1993). *Men, work, and family*. Newbury Park, Calif. ; London: Sage Publications.

Hooper, D., Coughlan, J. and Mullen, M. R. (2008). 'Structural Equation Modelling: Guidelines for Determining Model Fit'. *Electronic Journal of Business Research Methods,* 6 (1)*,* 53-60.

Horodynski, M. A., Baker, S., Coleman, G., Auld, G. and Lindau, J. (2011). 'The Healthy Toddlers Trial Protocol: an intervention to reduce risk factors for childhood obesity in economically and educationally disadvantaged populations'. *BMC Public Health,* 11*,* 581.

Howell, D. C. (2012). *Treatment of Missing Data--Part 1*. [Online]. Available at: http://www.uvm.edu/~dhowell/StatPages/More_Stuff/Missing_Data/Missing.html. [Last accessed 3rd January].

Hu, L. and Bentler, P. M. (1999). 'Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives'. *Struct Equ Modeling,* 6 (1)*,* 1-55.

Hurley, A. E., Scandura, T. A., Schriesheim, C. A., Brannick, M. T., Seers, A., Vandenberg, R. J. and Williams, L. J. (1997). 'Exploratory and confirmatory factor analysis: guidelines, issues, and alternatives'. *Journal of Organizational Behavior,* 18 (6)*,* 667-683.

Huybrechts, I., De Bacquer, D., Van Trimpont, I., De Backer, G. and De Henauw, S. (2006). 'Validity of parentally reported weight and height for preschool-aged children in Belgium and its impact on classification into body mass index categories'. *PEDIATRICS,* 118 (5)*,* 2109-18.

Hyde, T. M. and Miselis, R. R. (1983). 'Effects of area postrema/caudal medial nucleus of solitary tract lesions on food intake and body weight'. *Am J Physiol,* 244 (4)*,* R577-87.

Hyslop, D. R. and Imbens, G. W. (2001). 'Bias from classical and other forms of measurement error'. *Journal of Business & Economic Statistics,* 19 (4)*,* 475-481.

Jackson-Leach, R. and Lobstein, T. (2006). 'Estimated burden of paediatric obesity and co‐morbidities in Europe. Part 1. The increase in the prevalence of child obesity in Europe is itself increasing'. *International Journal of Pediatric Obesity,* 1 (1)*,* 26-32.

James, P. T., Leach, R., Kalamara, E. and Shayeghi, M. (2001). 'The worldwide obesity epidemic'. *Obesity Research,* 9*,* 228s-233s.

Janssen, I., Boyce, W. F., Simpson, K. and Pickett, W. (2006). 'Influence of individual- and area-level measures of socioeconomic status on obesity, unhealthy eating, and physical inactivity in Canadian adolescents'. *Am J Clin Nutr,* 83 (1)*,* 139-45.

Janssen, I., Craig, W. M., Boyce, W. F. and Pickett, W. (2004). 'Associations between overweight and obesity with bullying behaviors in school-aged children'. *PEDIATRICS,* 113 (5)*,* 1187-1194.

Jebb, S. A., Rennie, K. L. and Cole, T. J. (2004). 'Prevalence of overweight and obesity among young people in Great Britain'. *Public Health Nutr 2004,* 7*,* 461-5.

Jenkins, S. P. and Schluter, C. (2002a). The Effect of Family Income During Childhood on Later-Life Attainment: Evidence from Germany, *IZA Discussion Papers*: Institute for the Study of Labor (IZA) 604.

Jenkins, S. P. and Schluter, C. (2002b). *The effect of family income during childhood on later-life attainment: evidence from Germany*: IZA Discussion paper series.

Jo, Y. (2012). What Money Can Buy: Family Income and Child Obesity (4th Spetember 2012 ed.). http://www.artsci.wustl.edu/~gradconf/conf2012/Papers2012/YoungJo.pdf.

Jones, E. M. and Ketende, S. C. (2010). *Millennium Cohort Study: User Guide to Analysing MCS Data Using SPSS*. London: Centre for Longitudinal Studies, Institute of Education.

Jones, I. G. and Cameron, D. (1984). 'Social class analysis--an embarrassment to epidemiology'. *Community Med,* 6 (1)*,* 37-46.

Jotangia, D. (2005). *Obesity among young children under 11 [electronic resource]*. [Online].

Jurkowski, J. M., Lawson, H. A., Green Mills, L. L., Wilner, P. G., 3rd and Davison, K. K. (2014). 'The empowerment of low-income parents engaged in a childhood obesity intervention'. *Fam Community Health,* 37 (2)*,* 104-18.

Kaiser, H. F. (1960). 'The Application of Electronic-Computers to Factor-Analysis'. *Educational and Psychological Measurement,* 20 (1)*,* 141-151.

Kaiser, H. F. (1974). 'An index of factor simplicity'. *Psychometrika,* 39*,* 31-36.

Karlson, K. B., Holm, A. and Breen, R. (2012). 'Comparing Regression Coefficients Between Same-sample Nested Models Using Logit and Probit A New Method'. *Sociological Methodology,* 42*,* 286–313.

Kemptner, D., Jürges, H. and Reinhold, S. (2011). 'Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany'. *Journal of Health Economics,* 30 (2)*,* 340-354.

Kennedy, A. P., Shea, J. L. and Sun, G. (2009). 'Comparison of the Classification of Obesity by BMI vs. Dual‐energy X‐ray Absorptiometry in the Newfoundland Population'. *Obesity,* 17 (11)*,* 2094-2099.

Khaleque, A. and Rohner, R. P. (2012). 'Transnational Relations Between Perceived Parental Acceptance and Personality Dispositions of Children and Adults: A Meta-Analytic Review'. *Personality and Social Psychology Review,* 16 (2)*,* 103-115.

Klein-Platat, C., Wagner, A., Haan, M. C., Arveiler, D., Schlienger, J. L. and Simon, C. (2003). 'Prevalence and sociodemographic determinants of overweight in young French adolescents'. *Diabetes-Metabolism Research and Reviews,* 19 (2), 153-158.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. (3rd ed. ed.). New York ; London: Guilford.

Korupp, S. E., Ganzeboom, H. B. G. and Lippe, T. V. D. (2002). 'Do Mothers Matter? A Comparison of Models of the Influence of Mothers' and Fathers' Educational and Occupational Status on Children's Educational Attainment'. *Quality and Quantity,* 36 (1), 17-42.

Kuha, J. and Goldthorpe, J. H. (2010). 'Path analysis for discrete variables: the role of education in social mobility'. *Journal of the Royal Statistical Society Series a-Statistics in Society,* 173, 351-369.

Kullback, S. and Leibler, R. A. (1951). 'On Information and Sufficiency'. 79-86.

Lahelma, E., Martikainen, P., Laaksonen, M. and Aittomaki, A. (2004). 'Pathways between socioeconomic determinants of health'. *Journal of Epidemiology and Community Health,* 58 (4), 327-332.

Lakerveld, J., Verstrate, L., Bot, S. D., Kroon, A., Baan, C. A., Brug, J., Jansen, A. P., Droomers, M., Abma, T., Stronks, K. and Nijpels, G. (2013). 'Environmental interventions in low-SES neighbourhoods to promote healthy behaviour: enhancing and impeding factors'. *Eur J Public Health*.

Lamb, M. E. (1977a). 'The development of mother-infant and father-infant attachments in the second year of life'. *Developmental Psychology,* 13 (6), 637-648.

Lamb, M. E. (1977b). 'The development of parental preferences in the first two years of life'. *Sex Roles,* 3 (5), 495-497.

Lamb, M. E. (1977c). 'Father-Infant and Mother-Infant Interaction in the First Year of Life'. *Child Dev,* 48 (1), 167-181.

Lamerz, A., Kuepper-Nybelen, J., Wehle, C., Bruning, N., Trost-Brinkhues, G., Brenner, H., Hebebrand, J. and Herpertz-Dahlmann, B. (2005). 'Social class, parental education, and obesity prevalence in a study of six-year-old children in Germany'. *Int J Obes (Lond),* 29 (4), 373-380.

Lampard, A. M., Maclehose, R. F., Eisenberg, M. E., Neumark-Sztainer, D. and Davison, K. K. (2014). 'Weight-Related Teasing in the School Environment: Associations with Psychosocial Health and Weight Control Practices Among Adolescent Boys and Girls'. *J Youth Adolesc*.

Latner, J. D. and Stunkard, A. J. (2003). 'Getting worse: the stigmatization of obese children'. *Obes Res,* 11 (3), 452-6.

Leal, D. B., da Costa, F. F. and Altenburg de Assis, M. A. (2013). 'Sensitivity and specificity of body mass index-based classification systems for overweight in children 7-10 years old'. *RDCDH,* 15 (3), 267-275.

Lean, M. E. J., Han, T. S. and Morrison, C. E. (1995). 'Waist circumference as a measure for indicating need for weight management'. *Bmj,* 311 (6998), 158-161.

Leary, S., Davey Smith, G. and Ness, A. (2010). 'No evidence of large differences in mother-daughter and father-son body mass index concordance in a large UK birth cohort'. *Int J Obes (Lond),* 34 (7), 1191-2.

Leinonen, T., Martikainen, P. and Lahelma, E. (2012). 'Interrelationships between education, occupational social class, and income as determinants of disability retirement'. *Scandinavian Journal of Public Health,* 40 (2), 157-166.

Lindsay, A. C., Sussner, K. M., Kim, J. and Gortmaker, S. (2006). 'The role of parents in preventing childhood obesity'. *Future Child,* 16 (1), 169-86.

Lindsay, R. S., Hanson, R. L., Roumain, J., Ravussin, E., Knowler, W. C. and Tataranni, P. A. (2001). 'Body mass index as a measure of adiposity in children and adolescents: Relationship to adiposity by dual energy X-ray absorptiometry and to cardiovascular risk factors'. *Journal of Clinical Endocrinology & Metabolism,* 86 (9), 4061-4067.

Little, T. D., Slegers, D. W. and Card, N. A. (2006). 'A Non-arbitrary Method of Identifying and Scaling Latent Variables in SEM and MACS Models'. *Structural Equation Modeling,* 13 (1)*,* 59-72.

Lombardo, C., Battagliese, G., Lucidi, F. and Frost, R. O. (2012). 'Body dissatisfaction among pre-adolescent girls is predicted by their involvement in aesthetic sports and by personal characteristics of their mothers'. *Eating and Weight Disorders-Studies on Anorexia Bulimia and Obesity,* 17 (2)*,* E116-E127.

Loureiro, M. L., Sanz-de-Galdeano, A. and Vuri, D. (2010). 'Smoking Habits: Like Father, Like Son, Like Mother, Like Daughter?*'. *Oxford Bulletin of Economics and Statistics,* 72 (6)*,* 717-743.

Lovejoy, M. and Stone, P. (2012). 'Opting Back In: The Influence of Time at Home on Professional Women's Career Redirection after Opting Out'. *Gender Work and Organization,* 19 (6)*,* 631-653.

Lubans, D. R., Morgan, P. J. and Callister, R. (2012). 'Potential moderators and mediators of intervention effects in an obesity prevention program for adolescent boys from disadvantaged schools'. *J Sci Med Sport,* 15 (6)*,* 519-25.

Lukaski, H. C., Johnson, P. E., Bolonchuk, W. W. and Lykken, G. I. (1985). 'Assessment of fat-free mass using bioelectrical impedance measurements of the human body'. *The American journal of clinical nutrition,* 41 (4)*,* 810-817.

MacCallum, R. C., Browne, M. W. and Sugawara, H., M. (1996). 'Power Analysis and Determination of Sample Size for Covariance Structure Modeling'. *Psychological Methods,* 1 (2)*,* 130-49.

Mackenbach, J. P. and Kunst, A. E. (1997). 'Measuring the magnitude of socioeconomic inequalities in health: an overview of available measures illustrated with two examples from Europe'. *Social science and medicine,* 55*,* 757 - 771.

Mackenbach, J. P., Stirbu, I., Roskam, A.-J. R., Schaap, M. M., Menvielle, G., Leinsalu, M. and Kunst, A. E. (2008a). 'Socioeconomic inequalities in health in 22 European countries'. *New England Journal of Medicine,* 358 (23)*,* 2468-2481.

Mackenbach, J. P., Stirbu, I., Roskam, A. J. R., Schaap, M. M., Menvielle, G., Leinsalu, M., Kunst, A. E. and Socioec, E. U. W. G. (2008b). 'Socioeconomic inequalities in health in 22 European countries'. *New England Journal of Medicine,* 358 (23)*,* 2468-2481.

Maddala, G. S. and Lahiri, K. (2009). 'Chapter 15 Panel Data Analysis', *Introduction to econometrics* (4th ed. ed., pp. 583-589). New York ; Chichester: Wiley.

Maddox, G. L., Back, K. W. and Liederman, V. R. (1968). 'Overweight as a social deviance and disability'. *Journal of Health and Social Behavior,* 9*,* 287-298.

Mare, R. D. (1991). 'Five Decades of Educational Assortative Mating'. *American Sociological Review,* 56 (1)*,* 15-32.

Marmot, M. and Bell, R. (2012). 'Fair society, healthy lives'. *Public Health,* 126*,* S4-S10.

Marmot, M., Ryff, C. D., Bumpass, L. L., Shipley, M. and Marks, N. F. (1997). 'Social inequalities in health: Next questions and converging evidence'. *Soc Sci Med,* 44 (6)*,* 901-910.

Matijasevich, A., Victora, C. G., Golding, J., Barros, F. C., Menezes, A. M., Araujo, C. L. and Smith, G. D. (2009). 'Socioeconomic position and overweight among adolescents: data from birth cohort studies in Brazil and the UK'. *BMC Public Health,* 9.

Mayer, S. (2010). 'Revisiting an old question: How much does parental income affect child outcomes'. *Focus,* 27 (2)*,* 21-26.

Mayer, S. E. (1997). *What money can't buy: Family income and children's life chances*: Harvard University Press.

Mayer, S. E. (2002). *The influence of parental income on children's outcomes*: Knowledge Management Group, Ministry of Social Development Wellington,, New Zealand.

Mazess, R. B., Barden, H. S., Bisek, J. P. and Hanson, J. (1990). 'Dual-energy x-ray absorptiometry for total-body and regional bone-mineral and soft-tissue composition'. *The American journal of clinical nutrition,* 51 (6)*,* 1106-1112.

McIntosh, A., Kubena, K. S., Tolle, G., Dean, W., Kim, M. J., Jan, J. S. and Anding, J. (2011). 'Determinants of Children's Use of and Time Spent in Fast-food and Full-service Restaurants'. *Journal of Nutrition Education and Behavior,* 43 (3)*,* 142-149.

McPherson, K., Marsh, T. and Brown, M. (2007). foresight: Tackling Obesities: Future Choices – Modelling Future Trends in Obesity & Their Impact on Health. In G. O. f. Science (ed) (2nd ed.). London: Department of Innovation Universities and Skills.

Meara, E. (2001). *Why is health related to socioeconomic status?* : National Bureau of Economic Research.

Mela, C. F. and Kopalle, P. K. (2002). 'The impact of collinearity on regression analysis: the asymmetric effect of negative and positive correlations'. *Applied Economics,* 34 (6)*,* 667-677.

Metcalf, B. S., Hosking, J., Fremeaux, A. E., Jeffery, A. N., Voss, L. D. and Wilkin, T. J. (2011). 'BMI was right all along: taller children really are fatter (implications of making childhood BMI independent of height) EarlyBird 48'. *International Journal of Obesity,* 35 (4)*,* 541-547.

Meyer, D. R. and Garasky, S. (1993). 'Custodial Fathers: Myths, Realities, and Child Support Policy'. *Journal of Marriage and Family,* 55 (1)*,* 73-89.

Monasta, L., Lobstein, T., Cole, T. J., Vignerova, J. and Cattaneo, A. (2011). 'Defining overweight and obesity in pre-school children: IOTF reference or WHO standard?'. *Obes Rev,* 12 (4)*,* 295-300.

Mood, C. (2010). 'Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It'. *European Sociological Review,* 26 (1)*,* 67-82.

Mostazir, M., Jeffery, A., Voss, L. and Wilkin, T. (2013). 'Gender-assortative waist circumference in mother-daughter and father-son pairs, and its implications. An 11-year longitudinal study in children (EarlyBird 59)'. *Pediatr Obes*.

Musher-Eizenman, D. R., Holub, S. C., Miller, A. B., Goldstein, S. E. and Edwards-Leeper, L. (2004). 'Body size stigmatization in preschool children: The role of control attributions'. *J Pediatr Psychol,* 29 (8)*,* 613-620.

Muthén, L. K. and Muthén, B. O. (1998-2012). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.

Nardone, P., Lamberti, A., Baglio, G. and Spinelli, A. (2010). 'Maternal education and prevalence of obesity among children'. *European Journal of Public Health,* 20*,* 248-249.

National Obesity Observatory. (2009). 'Body Mass Index as a measure of obesity'. [Online]. Available at: http://www.noo.org.uk/NOO_pub/briefing_papers. [Last accessed 7th March 2014].

Nazroo, J. Y. and Karlsen, S. (2001). *Ethnic inequalities in health: social class, racism and identity*: Health Variations Programme.

Neumark-Sztainer, D., Story, M. and Harris, T. (1999). 'Beliefs and Attitudes about Obesity among Teachers and School Health Care Providers Working with Adolescents'. *Journal of Nutrition Education,* 31 (1)*,* 3-9.

Newman, D. A. (2003). 'Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques'. *Organizational Research Methods,* 6 (3)*,* 328-362.

NHS (2008). *Help stop childhood obesity before it starts*. www.sirc.org.

Nisbett, R. E. (1972). 'Eating behavior and obesity in men and animals'. *Adv Psychosom Med,* 7*,* 173-93.

O'Neill, J. L., McCarthy, S. N., Burke, S. J., Hannon, E. M., Kiely, M., Flynn, A., Flynn, M. A. T. and Gibney, M. J. (2007). 'Prevalence of overweight and obesity in Irish school children, using four different definitions'. *European Journal of Clinical Nutrition,* 61 (6)*,* 743-751.

Oakes, J. M. and Rossi, P. H. (2003). 'The measurement of SES in health research: current practice and steps toward a new approach'. *Soc Sci Med,* 56 (4)*,* 769-84.

Obesity Steering Group (2013). *Measuring Up: The medical profession's prescription to the obesity crisis*: Academy of Medical Royal Colleges

Observatory, N. O. (2012). 'NOO Data factsheet: Child Obesity and Socioeconomic Status'. [Online]. Available at: www.noo.org.uk.
OECD (2012). *Health at a Glance: Europe 2012,* . OECD Publishing http://dx.doi.org/10.1787/9789264183896-en.
OECD (2013). *Health at a Glance 2013: OECD INDICATORS*. OECD Publishing. http://dx.doi.org/10.1787/health_glance-2013-en.
Olvera, N., Dempsey, A., Gonzalez, E. and Abrahamson, C. (2013). 'Weight-related teasing, emotional eating, and weight control behaviors in Hispanic and African American girls'. *Eating behaviors,* 14 (4), 513-517.
Onis, M. d., Onyango, A. W., Borghi, E., Siyam, A., Nishida, C. and Siekmann, J. (2007). 'Development of a WHO growth reference for school-aged children and adolescents'. *Bulletin of the World Health Organization,* 85 (9), 660-667.
Oswal, A. and Yeo, G. S. (2007). 'The leptin melanocortin pathway and the control of body weight: lessons from human and murine genetics'. *Obes Rev,* 8 (4), 293-306.
Parikh, R., Mathai, A., Parikh, S., Sekhar, G. C. and Thomas, R. (2008). 'Understanding and using sensitivity, specificity and predictive values - Reply'. *Indian Journal of Ophthalmology,* 56 (4), 341-341.
Park, M. H., Falconer, C., Viner, R. M. and Kinra, S. (2012). 'The impact of childhood obesity on morbidity and mortality in adulthood: a systematic review'. *Obes Rev,* 13 (11), 985-1000.
Patrick, H. and Nicklas, T. A. (2005). 'A review of family and social determinants of children's eating patterns and diet quality'. *J Am Coll Nutr,* 24 (2), 83-92.
Peck, S. (23rd Jan, 2013). 'Tax junk food and subsidise vegetables to fight ‘poor obesity crisis’'. *The Telegraph*.
Perez-Pastor, E. M., Metcalf, B. S., Hosking, J., Jeffery, A. N., Voss, L. D. and Wilkin, T. J. (2009). 'Assortative weight gain in mother-daughter and father-son pairs: an emerging source of childhood obesity. Longitudinal study of trios (EarlyBird 43)'. *Int J Obes (Lond),* 33 (7), 727-35.
Peterson, J. L., Puhl, R. M. and Luedicke, J. (2012). 'An Experimental Assessment of Physical Educators' Expectations and Attitudes: The Importance of Student Weight and Gender'. *Journal of School Health,* 82 (9), 432-440.
Pevalin, D. J. and Rose, D. F. (2002). *A researcher's guide to the national statistics socio-economic classification*. London: SAGE.
Platt, L. (2011). *Inequality within ethnic groups.* York: JRF.
Plewis, I. (2007a). *The Millennium Cohort Study: Technical Report on Sampling*. London: Centre for Longitudinal Studies, Institute of Education, University of London.
Plewis, I. (2007b). 'Non-Response in a Birth Cohort Study: The Case of the Millennium Cohort Study'. *International Journal of Social Research Methodology,* 10 (5), 325-334.
Plewis, I., Calderwood, L., Hawkes, D. and Nathan, G. (2004). *National Child Development Study and 1970 British Cohort Study Technical Report: Changes in the NCDS and BCS70 Populations and Samples over Time*. CLS: Institute of Education.
Power, C. and Elliott, J. (2006). 'Cohort profile: 1958 British Birth Cohort (National Child Development Study)'. *International Journal of Epidemiology,* 35 (1), 34-41.
Power, C. and Moynihan, C. (1988). 'Social class and changes in weight-for-height between childhood and early adulthood'. *Int J Obes,* 12 (5), 445-53.
Prentice, A. M. and Jebb, S. A. (2001). 'Beyond body mass index'. *Obes Rev,* 2 (3), 141-7.
Press Association. (4th October, 2011). 'UK could introduce 'fat tax', says David Cameron'. *The Guardian*.
Public Health England. (2013). *Health risks of childhood obesity*. [Online]. Available at: http://www.noo.org.uk/NOO_about_obesity/obesity_and_health/health_risk_child. [Last accessed 25th July 2013].
Puhl, R. M. and Heuer, C. A. (2009). 'The Stigma of Obesity: A Review and Update'. *Obesity,* 17 (5), 941-964.

Puhl, R. M. and Heuer, C. A. (2010). 'Obesity Stigma: Important Considerations for Public Health'. *American Journal of Public Health,* 100 (6)*,* 1019-1028.
Puhl, R. M., Heuer, C. A. and Brownell, K. D. (2010). 'Stigma and Social Consequences of Obesity', *Clinical Obesity in Adults and Children* (pp. 25-40): Wiley-Blackwell. Available [Online] at: http://dx.doi.org/10.1002/9781444307627.ch3.
Puhl, R. M. and Latner, J. D. (2007). 'Stigma, Obesity, and the Health of the Nation's Children '. *Psychol Bull.,* 133 (4)*,* 557-580.
Radin, N. (1994). 'Primary caregiving fathers in intact families'. In A.E.Gottfried and A.W.Gottfried (Eds), *Redefining Families: Implications for Childrens Development* (pp. 55-97). New York: Plenum.
Rees, R., Oliver, K., Woodman, J. and Thomas, J. (2009). *Children's views about obesity, body size, shape and weight: A systematic review*. London: Institute of Education.
Reilly, J. J., Armstrong, J., Dorosty, A. R., Emmett, P. M., Ness, A., Rogers, I., Steer, C. and Sherriff, A. (2005). 'Early life risk factors for obesity in childhood: cohort study'. *BMJ,* 330 (7504)*,* 1357.
Reilly, J. J., Dorosty, A. R. and Emmett, P. M. (2000). 'Identification of the obese child: adequacy of the body mass index for clinical practice and epidemiology'. *Int J Obes Relat Metab Disord,* 24 (12)*,* 1623-7.
Reilly, J. J. and Kelly, J. (2011). 'Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: systematic review'. *Int J Obes (Lond),* 35 (7)*,* 891-8.
Richardson, S. A., Goodman, N., Hastorf, A. H. and Dornbusch, S. M. (1961). 'Cultural Uniformity in Reaction to Physical-Disabilities'. *American Sociological Review,* 26 (2)*,* 241-247.
Riphahn, R. T. and Serfling, O. (2005). 'Item non-response on income and wealth questions'. *Empirical Economics,* 30 (2)*,* 521-538.
Robinson, C., Craig, R. and Bridges, S. (2013). *Tracking the health of the nation*. [Online]. Available at: http://www.natcen.ac.uk/series/health-survey-for-england. [Last accessed 12th June].
Rohner, R. P. and Veneziano, R. A. (2001). 'The Importance of Father Love: History and Contemporary Evidence'. *Review of General Psychology,* 5 (4)*,* 382-405.
Rolland-Cachera, M.-F., Deheeger, M., Bellisle, F., Sempe, M., Guilloud-Bataille, M. and Patois, E. (1984). 'Adiposity rebound in children: a simple indicator for predicting obesity'. *The American journal of clinical nutrition,* 39 (1)*,* 129-135.
Rona, R. J. (1995). 'The National Study of Health and Growth (Nshg) - 23 Years on the Road'. *International Journal of Epidemiology,* 24*,* S69-S74.
Rona, R. J. and Chinn, S. (1982). 'National study of health and growth: social and family factors and obesity in primary schoolchildren'. *Ann Hum Biol,* 9 (2)*,* 131-45.
Rose, D. (2005). Socio-economic Classifications: Classes and Scales, Measurement and Theories, *European Association for Survey Research Conference*.
Rose, D., Pevalin, D. J. and O'Reilly, K. (2005). *The National Statistics Socio-economic Classification: Origins, Development and Use*. University of Essex: Institute for Social and Economic Research.
Rosenberg, J. and Wilcox, W. B. (2006). *The Importance of Fathers in the Healthy Development of Children*: U.S. Department of Health and Human Services.
Rosenberg, R. (2012). *MCS4: Guide to Derived Variables*. London: Centre for Longitudinal Studies.
Royston, P. (2009). 'Multiple imputation of missing values: Further update of ice, with an emphasis on categorical variables'. *Stata Journal,* 9 (3)*,* 466-477.
Royston, P. and White, I. R. (2011). 'Multiple imputation by chained equations (MICE): implementation in Stata'. *Journal of Statistical Software,* 45 (4)*,* 1-20.
Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* . New York, USA: John Willey & Sons.

Rudolf, M. (2010). *Tackling obesity through the Healthy Child Programme: a framework for action*. Leeds: University of Leeds, Leeds Community Healthcare NHS Trust.

Sarkadi, A., Kristiansson, R., Oberklaid, F. and Bremberg, S. (2008). 'Fathers' involvement and children's developmental outcomes: a systematic review of longitudinal studies'. *Acta Paediatrica,* 97*,* 153-158.

Savage, J. S., Fisher, J. O. and Birch, L. L. (2007). 'Parental influence on eating behavior: Conception to adolescence'. *Journal of Law Medicine & Ethics,* 35 (1)*,* 22-34.

Sayer, L. C., Bianchi, S. M. and Robinson, J. P. (2004). 'Are Parents Investing Less in Children? Trends in Mothers' and Fathers' Time with Children'. *AJS,* 110 (1)*,* 1-43.

Scaglioni, S., Arrizza, C., Vecchi, F. and Tedeschi, S. (2011). 'Determinants of children's eating behavior'. *Am J Clin Nutr,* 94 (6 Suppl)*,* 2006S-2011S.

Scarborough, P., Bhatnagar, P., Wickramasinghe, K. K., Allender, S., Foster, C. and Rayner, M. (2011). 'The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: an update to 2006-07 NHS costs'. *J Public Health (Oxf),* 33 (4)*,* 527-35.

Schroer, N. A. (1985). *Perceptions of In-service Teachers and Pre-service Teachers Toward Obese and Normal-weight Children*. Texas A&M University, Dissertation Abstracts International.

Schwartz, M. B., Vartanian, L. R., Nosek, B. A. and Brownell, K. D. (2006). 'The influence of one's own body weight on implicit and explicit anti-fat bias'. *Obesity (Silver Spring),* 14 (3)*,* 440-7.

Scott, K. M., Bruffaerts, R., Simon, G. E., Alonso, J., Angermeyer, M., de Girolamo, G., Demyttenaere, K., Gasquet, I., Haro, J. M., Karam, E., Kessler, R. C., Levinson, D., Medina Mora, M. E., Oakley Browne, M. A., Ormel, J., Villa, J. P., Uda, H. and Von Korff, M. (2007). 'Obesity and mental disorders in the general population: results from the world mental health surveys'. *International Journal of Obesity,* 32 (1)*,* 192-200.

Serdula, M. K., Ivery, D., Coates, R. J., Freedman, D. S., Williamson, D. F. and Byers, T. (1993). 'Do obese children become obese adults? A review of the literature'. *Prev Med,* 22 (2)*,* 167-77.

Shackleton, N. L. and Campbell, T. (2013). 'Are teachers' judgements of pupils' ability influenced by body shape?'. *Int J Obes (Lond)*.

Shafer, M. H. and Ferraro, K. H. (2011). 'The stigma of obesity: does percieved weight discrimination affect identity and physical health'. *Social Psychology Quarterly,* 74 (1)*,* 76-97.

Shankardass, K., McConnell, R., Jerrett, M., Lam, C., Wolch, J., Milam, J., Gilliland, F. and Berhane, K. (2013). 'Parental stress increases body mass index trajectory in pre-adolescents'. *Pediatr Obes*.

Shavers, V. L. (2007). 'Measurement of socioeconomic status in health disparities research'. *J Natl Med Assoc,* 99 (9)*,* 1013-23.

Shepherd, P. (1995). *The National Child Development Study (NCDS): An Introduction to the Origins of the Study and the Methods of Data Collection*

Centre for Longitudinal Studies.

Shrewsbury, V. and Wardle, J. (2008). 'Socioeconomic status and adiposity in childhood: a systematic review of cross-sectional studies 1990-2005'. *Obesity (Silver Spring),* 16 (2)*,* 275-84.

Singh, A. S., Mulder, C., Twisk, J. W. R., Mechelen, W. v. and Chinapaw, M. J. M. (2008). 'Tracking of childhood overweight into adulthood: a systematic review of the literature'. *obesity reviews,* 9*,* 474 - 488.

Singh, G. K., Siahpush, M. and Kogan, M. D. (2010). 'Rising social inequalities in US childhood obesity, 2003-2007'. *Ann Epidemiol,* 20 (1)*,* 40-52.

Solon, G. (1992). 'Intergenerational Income Mobility in the United-States'. *American Economic Review,* 82 (3)*,* 393-408.

Stalsberg, R. and Pedersen, A. V. (2010). 'Effects of socioeconomic status on the physical activity in adolescents: a systematic review of the evidence'. *Scand J Med Sci Sports,* 20 (3)*,* 368-83.

Stamatakis, E., Primatesta, P., Chinn, S., Rona, R. and Falascheti, E. (2005). 'Overweight and obesity trends from 1974 to 2003 in English children: what is the role of socioeconomic factors?'. *Arch Dis Child,* 90 (10)*,* 999-1004.

Stamatakis, E., Wardle, J. and Cole, T. J. (2010). 'Childhood obesity and overweight prevalence trends in England: evidence for growing socioeconomic disparities'. *Int J Obes (Lond),* 34 (1)*,* 41-7.

Stamatakis, E., Zaninotto, P., Falaschetti, E., Mindell, J. and Head, J. (2010). 'Time trends in childhood and adolescent obesity in England from 1995 to 2007 and projections of prevalence to 2015'. *J Epidemiol Community Health,* 64 (2)*,* 167-74.

Starrels, M. E. (1994). 'Gender differences in parent-child relations'. *Journal of family Issues,* 15 (1)*,* 148-165.

StataCorp. (1985-2013). Stata Survey Data Reference Manual (Vol. 13). Texas: Stata Press.

StataCorp. (2013). Stata Statistical Software: Release 13. College Station, TX: StataCorp LP.

Steinberger, J., Jacobs, D. R., Raatz, S., Moran, A., Hong, C. P. and Sinaiko, A. R. (2005). 'Comparison of body fatness measurements by BMI and skinfolds vs dual energy X-ray absorptiometry and their relation to cardiovascular risk factors in adolescents'. *Int J Obes (Lond),* 29 (11)*,* 1346-52.

Stenhammar, C., Olsson, G., Bahmanyar, S., Hulting, A. L., Wettergren, B., Edlund, B. and Montgomery, S. (2010). 'Family stress and BMI in young children'. *Acta Paediatr,* 99 (8)*,* 1205-12.

Strauss, R. S. and Pollack, H. A. (2003). 'Social marginalization of overweight children'. *Archives of Pediatrics & Adolescent Medicine,* 157 (8)*,* 746-752.

Sullivan, O. (2010). 'Changing Differences by Educational Attainment in Fathers' Domestic Labour and Child Care'. *Sociology-the Journal of the British Sociological Association,* 44 (4)*,* 716-733.

Swinburn, B., Egger, G. and Raza, F. (1999). 'Dissecting obesogenic environments: The development and application of a framework for identifying and prioritizing environmental interventions for obesity'. *Preventive Medicine,* 29 (6)*,* 563-570.

Szreter, S. R. S. (1984). 'The Genesis of the Registrar-Generals Social Classification of Occupations'. *British Journal of Sociology,* 35 (4)*,* 522-546.

Tampieri, A. (2010). *Sex and the Uni: Higher Education Effects in Job and Marital Satisfaction*.

Thomas, D. (1994). 'Like Father, like Son; Like Mother, like Daughter: Parental Resources and Child Height'. *The Journal of Human Resources,* 29 (4)*,* 950-988.

Thompson, L. and Walker, A. J. (1989). 'Gender in Families - Women and Men in Marriage, Work, and Parenthood'. *Journal of Marriage and the Family,* 51 (4)*,* 845-871.

Torssande, J. and Erikson, R. (2010). 'Stratification and Mortality—A Comparison of Education, Class, Status, and Income'. *Eur Sociol Rev,* 26 (4)*,* 465-474.

Tyler, D. O. and Horner, S. D. (2008). 'Collaborating with low-income families and their overweight children to improve weight-related behaviors: An intervention process evaluation'. *Journal for Specialists in Pediatric Nursing,* 13 (4)*,* 263-274.

UNESCO-UIS (2006). *International Standard Classification of Education: ISCED 1997*.

Van Buuren, S., Boshuizen, H. C. and Knook, D. L. (1999). 'Multiple imputation of missing blood pressure covariates in survival analysis'. *Statistics in medicine,* 18 (6)*,* 681-694.

Van den Pol, A. N. (1982). 'Lateral hypothalamic damage and body weight regulation: role of gender, diet, and lesion placement'. *Am J Physiol,* 242 (3)*,* R265-74.

Vandenburg, S. G. (1972). 'Assortative mating, or who marries whom?'. *Behav Genet,* 2 (2)*,* 127-57.

Veneziano, R. (2004). Parental Roles. In C. Ember and M. Ember (eds), *Encyclopedia of Sex and Gender – Men and Women in the World's Cultures*. Berlin Heidelberg: Springer-Verlag

Vize, R. (2011). 'Reality of the NHS budget squeeze'. *BMJ,* 343.

von Hippel, P. T. (2007). 'Regression with Missing Ys: An Improved Strategy for Analyzing Multiply Imputed Data'. *Sociological Methodology 2007, Vol 37,* 37*,* 83-117.

Wake, M., Hesketh, K. and Waters, E. (2003). 'Television, computer use and body mass index in Australian primary school children'. *Journal of Paediatrics and Child Health,* 39 (2)*,* 130-134.

Walley, A. J., Asher, J. E. and Froguel, P. (2009). 'The genetic contribution to non-syndromic human obesity'. *Nat Rev Genet,* 10 (7)*,* 431-42.

Walque, d. (2004). Education, Information and smoking decisions: Evidence from smoking histories 1940-2000 Washington: World Bank.

Wang, Y. and Zhang, Q. (2006). 'Are American children and adolescents of low socioeconomic status at increased risk of obesity? Changes in the association between overweight and family income between 1971 and 2002'. *The American journal of clinical nutrition,* 84 (4)*,* 707-716.

Wang, Y. C., McPherson, K., Marsh, T., Gortmaker, S. L. and Brown, M. (2011). 'Health and economic burden of the projected obesity trends in the USA and the UK'. *Lancet,* 378 (9793)*,* 815-25.

Wardle, J., Brodersen, N. H., Cole, T. J., Jarvis, M. J. and Boniface, D. R. (2006). 'Development of adiposity in adolescence: five year longitudinal study of an ethnically and socioeconomically diverse sample of young people in Britain'. *BMJ,* 332 (7550)*,* 1130-5.

Washbrook, E., Gregg, P. and Propper, C. (2013). A decomposition analysis of the relationship between parental income and multiple child outcomes, *CMPO Working Paper Series*: University of Bristol.

Wayman, J. C. (2003). *Multiple imputation for missing data: What is it and how can I use it*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Weightman, A. L., Morgan, H. E., Shepherd, M. A., Kitcher, H., Roberts, C. and Dunstan, F. D. (2012). 'Social inequality and infant health in the UK: systematic review and meta-analyses'. *BMJ Open,* 2 (3).

Wertheim, E. H., Martin, G., Prior, M., Sanson, A. and Smart, D. (2002). 'Parent influences in the transmission of eating and weight related values and behaviors'. *Eat Disord,* 10 (4)*,* 321-34.

White, M., Adamson, A., Chadwick, T., Dezateux, C., Griffiths, L., Howel, D., Kelly, S., Law, C., Li, L., Conte, R. L., Power, C. and Stamp, E. (2007). *The changing social patterning of obesity: an analysis to inform practice and policy development*. Newcastle: Public Health Research Consortium.

Wickramasinghe, V. P., Lamabadusuriya, S. P., Cleghorn, G. J. and Davies, P. S. W. (2009). 'Validity of currently used cutoff values of body mass index as a measure of obesity in Sri Lankan Children'. *The Ceylon medical journal* 54 (4)*,* 114-119.

Wijtzes, A. I., Jansen, W., Jansen, P. W., Jaddoe, V. W., Hofman, A. and Raat, H. (2013). 'Maternal educational level and preschool children's consumption of high-calorie snacks and sugar-containing beverages: Mediation by the family food environment'. *Prev Med*.

Williams, L. K., Andrianopoulos, N., Cleland, V., Crawford, D. and Ball, K. (2013). 'Associations Between Education and Personal Income With Body Mass Index Among Australian Women Residing in Disadvantaged Neighborhoods'. *American Journal of Health Promotion,* 28 (1)*,* 59-65.

Williams, Z. (14th December, 2011). 'Obesity is about poverty and cheap food, not a lack of moral fibre'*. The Guardian*.

Wills, M. (2004). 'Orthopedic complications of childhood obesity'. *Pediatr Phys Ther,* 16 (4)*,* 230-5.

Winkleby, M. A., Jatulis, D. E., Frank, E. and Fortmann, S. P. (1992). 'Socioeconomic-Status and Health - How Education, Income, and Occupation Contribute to Risk-Factors for Cardiovascular-Disease'. *American Journal of Public Health,* 82 (6)*,* 816-820.

World Health Organisation. (2013). *Global Strategy on Diet, Physical Activity and Health*. [Online]. [Last accessed 16th august].

World Health Organisation. (2014). *Obesity and overweight*. [Online]. Available at: http://www.who.int/mediacentre/factsheets/fs311/en/. [Last accessed 17th March].

Wyatt, I. D. and Hecker, D. E. (2006). 'Occupational changes during the 20th century'. *Monthly Labor Review,* 129 (3)*,* 35-57.

Zeller, M. H., Reiter-Purtill, J. and Ramey, C. (2008). 'Negative peer perceptions of obese children in the classroom environment'. *Obesity (Silver Spring),* 16 (4)*,* 755-62.

Zimmermann, M. B., Gubeli, C., Puntener, C. and Molinari, L. (2004). 'Detection of overweight and obesity in a national sample of 6-12-y-old Swiss children: accuracy and validity of reference values for body mass index from the US Centers for Disease Control and Prevention and the International Obesity Task Force'. *Am J Clin Nutr,* 79 (5)*,* 838-43.

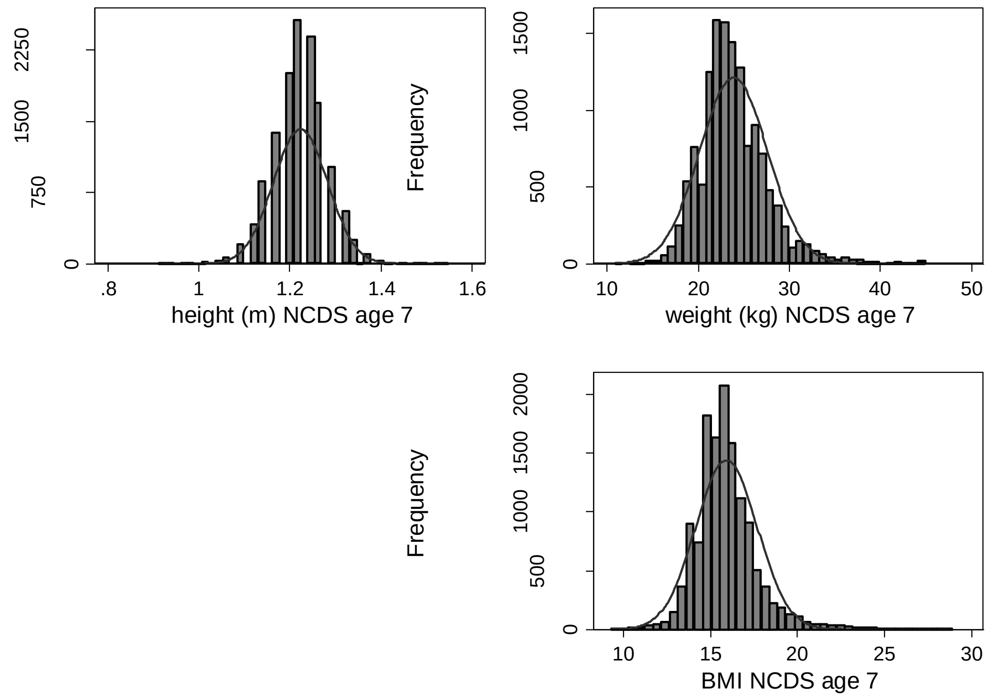
Appendices

Appendix A: The distribution of Height, Weight and BMI

In this appendix I show the distribution of height, weight and BMI for children aged between ages 7 and 10 in the NCDS, BCS and MCS cohorts. Children between ages 7 and 10 are those that make up the sample for the main analysis in chapter 4.

National Child Development Study

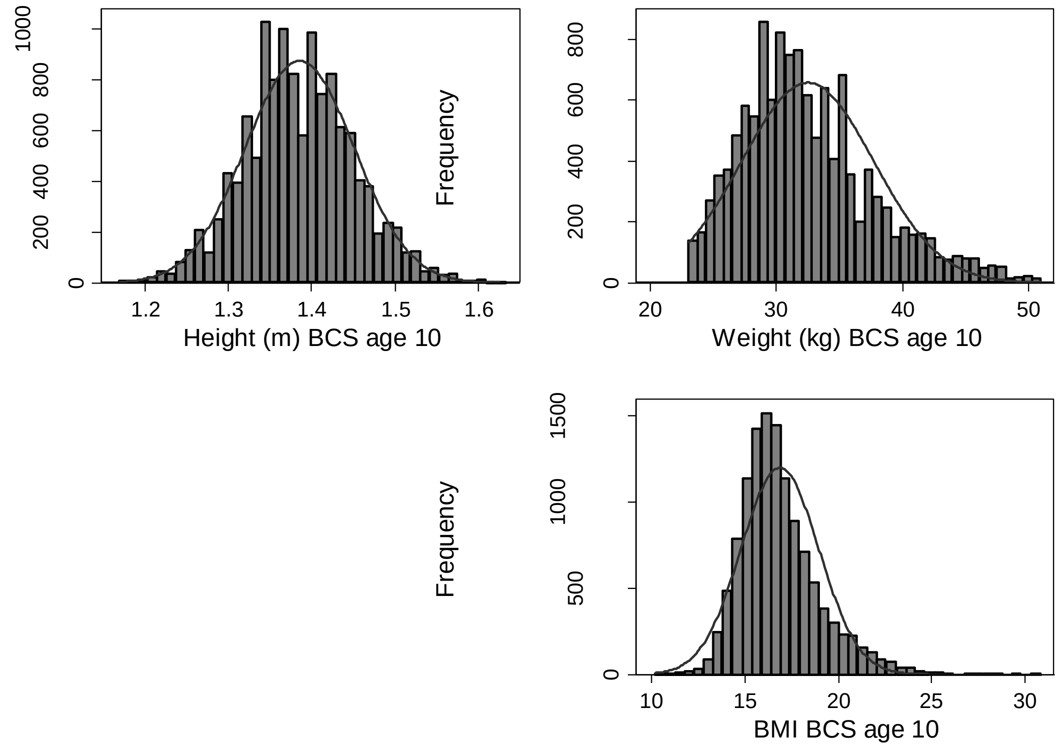
Figure A.1. The distribution of Height, Weight and BMI in the NCDS cohort



curved lines represent appropriately scaled normal density curve

British Cohort Study

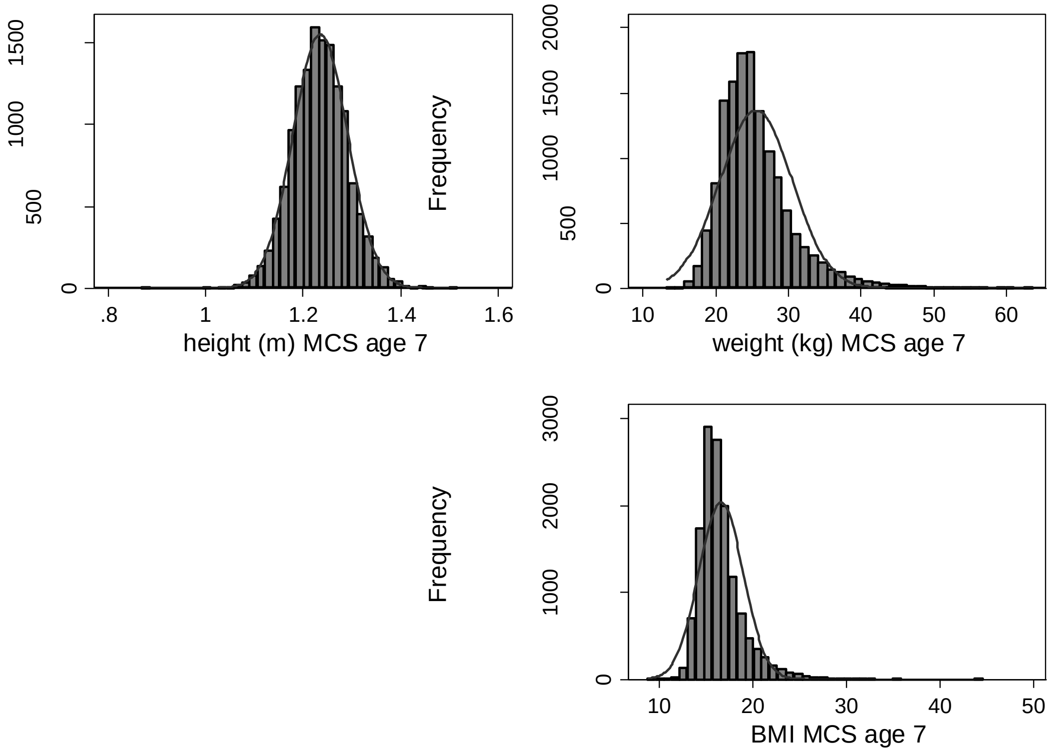
Figure A.2. The distribution of Height, Weight and BMI in the BCS cohort.



curved lines represent appropriately scaled normal density curve

Millennium Cohort Study

Figure A.3. The distribution of Height, Weight and BMI in the MCS cohort.



curved lines represent appropriately scaled normal density curve

Appendix B: Fitting the latent variable model and model fit assessment

There are several criteria for assessing whether a latent variable approach is appropriate, but the first step is to consider the correlation coefficients between observed variables. The correlations between the indicators are shown in table A.1. Within the framework for CFA, these observed variables are expected to be caused by the same underlying construct, therefore changes in one indicator should have a relationship with changes in another. The magnitude of the correlation coefficients show that these variables are indeed related (Field, 2009), with the correlation coefficients showing a similar pattern within each birth cohort data set.

Table A.1. correlation matrix for the indicators of SES in the NCDS, BCS and MCS.

NCDS	RGSC	NS-SEC	Mum school	Dad school	Home owner	Crowding
RGSC	1					
NS-SEC	0.66	1				
Mum school	0.45	0.46	1			
Dad school	0.59	0.60	0.54	1		
Home owner	0.45	0.47	0.44	0.46	1	
Crowding	0.33	0.35	0.35	0.40	0.45	1
BCS70	RGSC	NS-SEC	Mum school	Dad school	Home owner	Crowding
RGSC	1					
NS-SEC	0.77	1				
Mum school	0.43	0.44	1			
Dad school	0.59	0.59	0.64	1		
Home owner	0.46	0.44	0.44	0.51	1	
Crowding	0.34	0.36	0.32	0.35	0.30	1
MCS	RGSC	NS-SEC	Mum school	Dad school	Home owner	Crowding
RGSC	1					
NS-SEC	0.71	1				
Mum school	0.46	0.51	1			
Dad school	0.43	0.46	0.48	1		
Home owner	0.52	0.60	0.43	0.30	1	
Crowding	0.39	0.43	0.28	0.24	0.59	1

Notes: The 'polychoric' command was used in stata. This command takes into account the level of measurement of the data and calculates the appropriate type of correlation. When both variables have 10 or fewer observed values, a polychoric correlation is calculated, when only one of the variables takes on 10 or fewer values (i.e., one variable is continuous and the other categorical) a polyserial correlation is calculated, and if both variables take on more than 10 values a Pearson's correlation is calculated

There were two measures of social class available, the RGSC (which may be considered as closer to a measure of occupational status than class (Bartley, 2004)) and the NS-SEC. Several model fit criteria were used to asses model fit in three competing models for the measurement of the latent factor socioeconomic status. Model fit statistics are presented in table A.2. Each competing model included mother's education, father's education, home ownership and the ratio of rooms per person in the household. Indicators of social class changed between each model specification: 1) RGSC, 2) RGSC & NS-SEC, and 3) NS-SEC.

The log-likelihood value, the Akaike Information Criterion (AIC) and the Bayes Information Criterion (BIC) criteria indicate that model specification 3, the model containing just the NS-SEC provides the best fitting model. Likelihood values represent the likelihood of the parameter estimates given the observed data. As these values are transformed using the natural logarithm, values closer to 0 represent larger likelihoods. The AIC and BIC are largely based on the log-likelihood values, but the AIC takes into account the complexity (the number of parameters) of the model, and the BIC takes into account the sample size and the complexity of the model. The AIC and BIC are referred to as information criterion. These measures arose from information theory (Burnham and Anderson, 2002; Kullback and Leibler, 1951). When two models are fit on the same data, the model with the smaller value of the information criterion is considered better. As can be seen in equation A.1 and A.2 below, the AIC and BIC are similar measures and only differ in the amount by which they penalize large models. The AIC and BIC values are consistently lower for the model containing just NS-SEC in all three birth cohorts.

Equation A.1. The AIC criteria.

$$AIC = -2(\ln(\text{likelihood})) + 2*K$$

Where likelihood is the probability of the data given the model and K equals the number of free parameters.

Equation A.2. The BIC criteria.
BIC= -2(ln(likelihood)) + ln(N)*K

Where likelihood is the probability of the data given the model and K equals the number of free parameters and N is the number of observations.

Table A.2 Model fit criteria for different specifications of confirmatory factor analysis.

		Log likelihood	df	AIC	BIC	n
NCDS						
	RGSC	-55911	15	111852	111969	17887
	RGSC & NS-SEC	-67307	18	134650	134790	18111
	NS-SEC	-47064	15	94159	94276	18105
BCS						
	RGSC	-45574	15	91178	91293	15437
	RGSC & NS-SEC	-57578	18	115191	115328	15437
	NS-SEC	-45260	15	90550	90664	15437
MCS						
	RGSC	-44022	15	88073	88186	14043
	RGSC & NS-SEC	-55249	18	110533	110669	14043
	NS-SEC	-41032	15	82094	82208	14043

When deriving the latent variable, the level of measurement of the variables was taken into account using logit regression for binary variables and ordinal logit for ordinal variables. Table A.3 shows the standardised coefficients or “factor loadings”. The strength of these loadings demonstrates the variance in the indicator that is explained by the latent variable. Higher loadings suggest that the indicator variable is measuring the underlying latent variable with higher precision. Indeed the loadings are all reasonable (Field, 2009).

Global fit statistics can also be used to asses model fit (Byrne, 2011). These fit statistics tell us different pieces of information referring to how well the model fits the data, therefore Several fit indices should be used to determine model fit (Hooper, Coughlan and Mullen, 2008). The chi squared statistic, the root mean square error of approximation (RMSEA), and the Comparative Fit Index (CFI) are presented as the measures of global model fit and are described below.

Table A.3. Standardised factor loadings for the CFA with and without covariance between ratio of rooms and home ownership in the NCDS, BCS70 and MCS.

	NCDS		**BCS**		**MCS**	
	CFA	CFA with covariance	CFA	CFA with covariance	CFA	CFA with covariance
NS-SEC	0.67	0.73	0.70	0.70	0.73	0.86
Mothers Education	0.65	0.68	0.71	0.71	0.60	0.67
Fathers Education	0.75	0.80	0.84	0.84	0.52	0.57
Home ownership	0.73	0.62	0.63	0.62	0.88	0.68
House crowding	0.59	0.50	0.47	0.47	0.46	0.38
Correlation Home ownership with crowding		0.21		0.01		0.24
Ordinal Alpha[44]	0.80		0.80		0.77	

Model fit statistics

How well the hypothesis model describes the data is measured via model fit. The evaluation of model should be based upon the model as a whole (global model fit), as well as the individual parameters (Byrne, 2011) and should be based upon several model fit criteria (Hooper, Coughlan and Mullen, 2008). In contrast to how a null hypothesis is normally conceptualised in social science, within the SEM framework the null hypothesis is that the specified model holds in the population. The primary focus of the estimation process in SEM is to yield parameter estimates that minimise the discrepancy (the residual) between the sample covariance matrix and population covariance matrix implied by the model (Hu and Bentler, 1999; Little, Slegers and Card, 2006). This objective is achieved by minimizing a discrepancy function (Fmin), where the discrepancy between the sample covariance matrix and the population covariance matrix is least.

To reiterate the above algebraically, we take ‘s’ to represent the sample covariance matrix, ‘∑’ to represent the population covariance matrix, and ‘θ’ to represent a vector of the model parameters, so that ‘∑θ’ represents the covariance matrix implied by the model. The null hypothesis is therefore ‘∑=∑θ’. Fmin reflects the point in the

[44] Ordinal alpha provides a measure of internal consistency based on the polychoric correlation matrix. Ordinal alpha was calculated using the formula provided by Gaderman, Guhn & Zimbo (2012) (Gadermann, Guhn and Zumbo, 2012). Ordinal alpha is conceptually identical to Cronbach’s alpha, but is based on the polychoric correlation matrix. Alpha values above 0.7 indicate a good amount of internal consistency amongst the items (Bland and Altman, 1997). The ordinal alpha values shown in table 4.5 indicate a good amount of internal consistency in the indicators, suggesting a high degree of interrelatedness amongst the indicators (Gadermann, Guhn and Zumbo, 2012).

estimation where S- ∑θ = minimum. Fmin therefore measures the extent to which 's' differs from '∑θ'. This value is used to calculate Chi Square statistic χ^2, one measure of global model fit (Byrne, 2011; Hu and Bentler, 1999).

Chi Square

The Chi Square statistic represents the discrepancy between the sample covariance matrix, s, and the restricted covariance matrix ∑θ. The formula for the Chi square statistic is shown in equation A.3.

Equation A.3. The chi square statistic.

$$\chi^2 = (N-1)*F_{min}$$

Where χ^2 is the chi square statistic, N is the number observations[45] and F_{min} is the discrepancy function (Hu and Bentler, 1999; Kline, 2011).

Lower values of the chi square statistic indicate a smaller amount of discrepancy between the observed and fitted values. One of the most widely noted disadvantages of the chi square statistic is its sensitivity to sample size (Byrne, 2011a; Hooper, Coughlan and Mullen, 2008; Hu and Bentler, 1999; Kline, 2011). Even very small discrepancies between the sample covariance matrix and the restricted covariance matrix can become highly significant with large sample sizes. However large sample sizes are required in the analysis of covariance structures, and because all models are approximations there will always be some discrepancy.

The chi square statistic is useful for comparing the fit of nested models and it is used in the computation of many other model fit indices, many designed specifically to deal with the problem of sensitivity to sample size. The descriptions and calculations for many of these fit indices are provided by Byrne (2011), in her invaluable introductory text to SEM. Here I focus on the Root Mean Square Error of Approximation (RMSEA) and the Comparative Fit Index (CFI).

RMSEA

[45] In the Mplus software package, the calculation of chi square only N (the sample size) is used in the calculation of chi square, rather than the tradition N-1 (Byrne, 2011).

The RMSEA is a measure of approximate fit rather than “exact” or perfect fit. The RMSEA is also based on the residual or discrepancy between the hypothesized model and the observed data, But the RMSEA is sensitive to the complexity of the model and the sample size. The formula for the RMSEA is shown in equation A.4. RMSEA values can range from 0 to 1, with values closer to 0 representing smaller discrepancy given the complexity of the model and the sample size. Different cut off criteria for the RMSEA have been recommended but these criteria appear to be the general rule: Values less than 0.05 indicate good fit, values less than 0.08 indicate reasonable fit, values ranging from 0.08-0.10 indicate mediocre fit and values greater than 0.10 indicate poor fit (MacCallum, Browne and Sugawara, 1996). The distribution of RMSEA is known, therefore it is possible to calculate confidence intervals around the predicted value. If the lower value of these intervals is close to 0 and the upper value less than 0.08 we can be reasonably confident the model fits the data well (Hooper, Coughlan and Mullen, 2008)

Equation A.4 The RMSEA formula

$$\frac{\sqrt{(\chi 2\ -\ \mathrm{df})}}{\sqrt{[\mathrm{df}(\mathrm{N}\ -\ 1)]}}$$

Where $\chi 2$ is the chi squared statistic, df is the degrees of freedom in the hypothesized model and N is the sample size.

The comparative Fit Index

The CFI compares the fit of the proposed model with a null or ‘baseline’ model, in which all variables are assumed to be orthogonal. The measure is expressed as a proportionate improvement in model fit. Values for the CFI are normed and range between 0 and 1, with values close to 1 being indicative of a well-fitting model, indeed a cut of value on 0.95 was advised by Hu and Bentler (1999) on the basis of simulation studies. The CFI is computed as shown in equation A.5.

Equation A.5. The Comparative fit index formula

$$\mathrm{CFI} = 1 - [(\chi^2\mathrm{H}\text{-}\mathrm{dfH})/(\chi^2\mathrm{B}\text{-}\mathrm{dfB})$$

Where χ^2 is the chi squared value, H is the hypothesized model, B is the baseline model, and df refers to degrees of freedom.

As can be seen in table A.4, the chi square values and the RMSEA values for the NCDS and MCS are slightly high, but allowing the residuals of home ownership and ratio of rooms per person (crowding) to co-vary substantially improves model fit. The covariance between the residuals of home ownership and crowding likely represents an urban/rural divide, whereby we might expect the size of houses and home ownership vary depending upon other things than socioeconomic status, including the area/region in which they live. The covariance was not included in the BCS in the final model, as it did not result in better fit. Correlations between factor scores generated including and not including this covariance were very high (>.97 in all cohorts). Therefore it is unlikely that the choice to include the residual covariance in the NCDS and MCS has a large influence on the results.

Table A.4. Model fit statistics for CFA with and without covariance between ratio of rooms and home ownership included in the NCDS, BCS and MCS.

	NCDS		**BCS**		**MCS**	
Fit statistics	CFA	CFA with covariance	CFA	CFA with covariance	CFA	CFA with covariance
Chi2	311 (5df)	32 (4df)	111 (5df)	116 (4df)	558 (5df)	183 (4df)
RMSEA (95% C.I)	0.058 (0.053-0.064)	0.020 (0.014 – 0.026)	0.037 (0.031-0.043)	0.043 (0.036 – 0.049)	0.076 (0.071 – 0.081)	0.048 (0.042-0.054)
CFI	0.98	0.99	0.99	0.99	0.97	0.99

The fit statistics reported here were produced in Mplus version 7.1. (Muthén and Muthén, 1998-2012). Stata 13 will not provide these fit statistics for non-continuous data (StataCorp 2013).

Appendix C: Coefficients for the interactions between SES, time and gender

NS-SEC

Table A.5. The interaction between NS-SEC, time period dummies and gender.

		OR	S.E	z	p	95% confidence Intervals	
NS-SEC+							
	2	0.99	0.25	-0.03	0.97	0.60	1.64
	3	0.91	0.25	-0.33	0.74	0.53	1.56
	4	1.14	0.28	0.54	0.59	0.71	1.85
	5	1.29	0.30	1.11	0.27	0.82	2.03
	6	1.14	0.27	0.56	0.58	0.72	1.80
	7	1.15	0.26	0.63	0.53	0.74	1.79
Time period^							
	1980	0.72	0.19	-1.23	0.22	0.43	1.22
	2007	2.02	0.45	3.15	0.00	1.30	3.13
NS-SEC#Time period							
2 1980		1.35	0.47	0.88	0.38	0.69	2.66
2 2007		1.36	0.39	1.08	0.28	0.78	2.38
3 1980		0.75	0.30	-0.71	0.48	0.34	1.66
3 2007		1.81	0.57	1.88	0.06	0.98	3.35
4 1980		1.34	0.45	0.87	0.38	0.69	2.59
4 2007		1.26	0.37	0.79	0.43	0.71	2.23
5 1980		1.01	0.33	0.04	0.97	0.54	1.91
5 2007		0.95	0.28	-0.18	0.85	0.53	1.69
6 1980		1.17	0.39	0.48	0.63	0.61	2.23
6 2007		1.55	0.44	1.56	0.12	0.89	2.69
7 1980		1.27	0.40	0.77	0.44	0.69	2.34
7 2007		1.36	0.37	1.11	0.27	0.79	2.32
female							
Female = 1		1.55	0.41	1.67	0.10	0.93	2.59
NS-SEC#female							
2#female		0.93	0.32	-0.21	0.84	0.48	1.82
3#female		0.89	0.33	-0.31	0.75	0.43	1.83
4#female		1.07	0.35	0.20	0.84	0.56	2.03
5#female		0.71	0.22	-1.08	0.28	0.38	1.32
6#female		0.75	0.24	-0.90	0.37	0.40	1.40
7#female		0.91	0.27	-0.33	0.74	0.50	1.63
Time period#female							
1980#female		1.17	0.41	0.45	0.65	0.59	2.33
2007#female		0.88	0.26	-0.43	0.67	0.49	1.58
NS-SEC#Time period#female							
2#1980#female		0.81	0.37	-0.46	0.64	0.33	1.97
2#2007#female		1.03	0.39	0.07	0.94	0.49	2.17
3#1980#female		1.29	0.68	0.49	0.62	0.46	3.64
3#2007#female		0.83	0.35	-0.45	0.65	0.36	1.90
4#1980#female		0.75	0.33	-0.66	0.51	0.31	1.78
4#2007#female		0.91	0.35	-0.24	0.81	0.42	1.95
5#1980#female		1.32	0.56	0.64	0.52	0.57	3.05

5#2007#female	1.88	0.76	1.58	0.12	0.86	4.14
6#1980#female	1.19	0.52	0.40	0.69	0.51	2.79
6#2007#female	1.38	0.52	0.86	0.39	0.66	2.90
7#1980#female	1.09	0.45	0.22	0.83	0.49	2.43
7#2007#female	1.27	0.47	0.65	0.52	0.62	2.60
constant	0.07	0.01	-13.24	0.00	0.05	0.11
n	32 972					
Fit statistics	R2 [a]	log pseudolikelihood [b]			BIC	
	0.0406	-12340.95			-317340.739	
Test of interaction	chi(12)=10.03, p=0.61					

[a] pseudo R2

[b] pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.

Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.

[+]NS-SEC 1 (Higher managerial and professional occupations) is the reference category

[^] 1965 is the reference category

RGSC

Table A.6. The interaction between RGSC, time period dummies and gender.

	OR	S.E	z	p	95% confidence Intervals	
RGSC+						
IV	1.12	0.24	0.51	0.61	0.73	1.72
III	1.09	0.22	0.44	0.66	0.74	1.61
II	1.32	0.29	1.25	0.21	0.86	2.02
I	0.92	0.26	-0.28	0.78	0.53	1.61
Time period						
1980	1.20	0.37	0.59	0.56	0.66	2.18
2007	2.94	0.85	3.71	0.00	1.66	5.19
RGSC#Time period						
IV 1980	0.68	0.25	-1.07	0.29	0.34	1.38
IV 2007	1.00	0.32	-0.01	0.99	0.53	1.89
III 1980	0.70	0.22	-1.13	0.26	0.37	1.30
III 2007	0.90	0.27	-0.35	0.73	0.50	1.63
II 1980	0.58	0.20	-1.58	0.12	0.30	1.14
II 2007	0.67	0.21	-1.24	0.22	0.36	1.26
I 1980	0.51	0.23	-1.49	0.14	0.21	1.24
I 2007	0.71	0.28	-0.87	0.38	0.33	1.53
Female						
female =1	1.29	0.34	0.98	0.33	0.77	2.15
RGSC#female						
2#female	1.18	0.35	0.56	0.57	0.66	2.12
3#female	1.06	0.29	0.23	0.82	0.62	1.82
4#female	1.00	0.30	-0.01	0.99	0.55	1.80
5#female	1.18	0.45	0.42	0.67	0.55	2.50
Time period#female						
1980#female	0.88	0.38	-0.29	0.77	0.38	2.06
2007#female	1.08	0.45	0.19	0.85	0.48	2.45
rgsc#time period#female						
IV#1980#female	1.29	0.64	0.50	0.61	0.48	3.41
IV#2007#female	0.85	0.39	-0.35	0.72	0.34	2.10
III#1980#female	1.39	0.63	0.73	0.47	0.57	3.36
III#2007#female	0.95	0.41	-0.11	0.91	0.41	2.23
II#1980#female	1.34	0.64	0.60	0.55	0.52	3.43
II#2007#female	0.88	0.40	-0.27	0.79	0.36	2.15
I#1980#female	1.72	1.05	0.88	0.38	0.52	5.70
I#2007#female	0.71	0.39	-0.61	0.54	0.24	2.09
constant	0.08	0.01	-13.61	0.00	0.05	0.11
n	36 128					
Fit statistics	R2	log pseudolikelihood			BIC	
	0.0372	-13318.755			-351763.845	
Test of interaction		chi(8)=2.65, p=0.95				

[a] *pseudo R2*
[b] *pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.*
Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.
[+]*RGSC V is the reference category*
[^] *1965 is the reference category*

Parental education

Table A.7. The interaction between whether or not mother was in school past minimum school leaving age (stayed/ did not stay), time period dummies and gender.

	OR	S.E	z	p	95% confidence Intervals	
Mumschool - stayed[+]	0.74	0.08	-2.59	0.01	0.59	0.93
Time period[^]						
1980	0.86	0.07	-1.83	0.07	0.73	1.01
2007	2.58	0.19	12.58	0.00	2.23	2.99
Mum school # Time period						
stayed # 1980	1.01	0.17	0.07	0.94	0.72	1.42
stayed # 2007	1.13	0.16	0.91	0.37	0.87	1.48
Female						
female =1	1.38	0.10	4.58	0.00	1.20	1.58
mumschool#female						
stayed # female	1.18	0.18	1.09	0.28	0.88	1.58
Time period # female						
1980 # female	1.21	0.13	1.76	0.08	0.98	1.50
2007 # female	1.04	0.11	0.36	0.72	0.85	1.27
mumschool#time period#female						
stayed#1980#female	0.88	0.20	-0.57	0.57	0.57	1.36
stayed#2007#female	0.74	0.14	-1.64	0.10	0.52	1.06
constant	0.09	0.00	-46.79	0.00	0.08	0.10
n	36 181					
Fit statistics	R2 [a]	log pseudolikelihood [b]			BIC	
	0.0363	-13516.511			-352355.358	
Test of interaction	chi(2)=2.87, p=0.24					

[a] pseudo R2

[b] pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.

Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.

[+] did not stay at school past the minimum leaving age is the reference category

[^] 1965 is the reference category

Table A.8. The interaction between whether or not father was in school past minimum school leaving age (stayed/ did not stay), time period dummies and gender.

	OR	S.E	z	p	95% confidence Intervals	
Dad school - stayed	0.76	0.09	-2.34	0.02	0.61	0.96
Time period						
1980	0.91	0.08	-1.15	0.25	0.77	1.07
2007	2.58	0.20	12.29	0.00	2.22	3.00
Dad school # Time period						
stayed # 1980	0.84	0.15	-0.96	0.34	0.59	1.20
stayed # 2007	0.94	0.14	-0.40	0.69	0.71	1.26
Female						
female =1	1.38	0.10	4.58	0.00	1.20	1.58
Dadschool#female						
stayed # female	1.22	0.19	1.33	0.19	0.91	1.65
Time period # female						
1980 # female	1.17	0.13	1.44	0.15	0.94	1.45
2007 # female	0.97	0.10	-0.30	0.77	0.79	1.19
Dadschool#time period#female						
stayed#1980#female	1.00	0.23	-0.01	0.99	0.63	1.57
stayed#2007#female	0.86	0.17	-0.79	0.43	0.58	1.26
constant	0.09	0.00	-47.22	0.00	0.08	0.10
n	32 467					
Fit statistics	R2	log pseudolikelihood			BIC	
	0.0328	-11956.969			-312978.624	
Test of interaction	chi(2)=0.84, p=0.66					

[a] *pseudo R2*
[b] *pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.*
Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.
[+] *did not stay at school past the minimum leaving age is the reference category*
[^] *1965 is the reference category*

Latent measure of SES

Table A.9. The interaction between the quintiles of the latent measure of SES, time period dummies and gender.

	OR	S.E	z	p	95% confidence Intervals	
SES quintiles[+]						
2	1.30	0.19	1.84	0.07	0.98	1.73
3	1.17	0.17	1.06	0.29	0.88	1.55
4	1.22	0.17	1.43	0.15	0.93	1.59
5	1.05	0.15	0.35	0.73	0.79	1.40
Time period[^]						
1980	0.94	0.14	-0.44	0.66	0.70	1.25
2007	3.13	0.41	8.62	0.00	2.41	4.06
SES quintile # Time period						
2 1980	0.85	0.18	-0.75	0.46	0.56	1.30
2 2007	0.80	0.15	-1.16	0.25	0.55	1.17
3 1980	1.01	0.22	0.06	0.95	0.67	1.54
3 2007	0.87	0.16	-0.75	0.46	0.60	1.26
4 1980	0.79	0.17	-1.10	0.27	0.53	1.20
4 2007	0.77	0.14	-1.47	0.14	0.54	1.09
5 1980	0.75	0.17	-1.32	0.19	0.48	1.15
5 2007	0.64	0.12	-2.31	0.02	0.44	0.94
Female						
female = 1	1.47	0.20	2.90	0.00	1.13	1.91
SES quintiles # female						
2#female	0.90	0.17	-0.54	0.59	0.62	1.32
3#female	0.92	0.18	-0.41	0.68	0.63	1.35
4#female	0.95	0.17	-0.29	0.77	0.66	1.36
5#female	1.04	0.20	0.19	0.85	0.71	1.52
Time period#female						
1980#Male	1.17	0.23	0.82	0.41	0.80	1.72
2007#Male	0.97	0.17	-0.15	0.88	0.69	1.38
LogitSES#Time period#female						
2#1980#female	1.11	0.31	0.38	0.70	0.64	1.94
2#2007#female	1.28	0.33	0.95	0.34	0.77	2.11
3#1980#female	0.94	0.27	-0.21	0.83	0.54	1.64
3#2007#female	0.95	0.24	-0.19	0.85	0.58	1.57
4#1980#female	1.00	0.28	-0.01	0.99	0.58	1.71
4#2007#female	0.89	0.21	-0.48	0.63	0.56	1.43
5#1980#female	0.98	0.28	-0.07	0.95	0.56	1.73
5#2007#female	0.85	0.22	-0.65	0.52	0.51	1.40
_cons	0.07	0.01	-25.80	0.00	0.06	0.09
n	32 972					
Fit statistics	R2 [a]	log pseudolikelihood [b]			BIC	
	0.0388	-14186.828			-380347.511	
Test of interaction	chi(8)=3.32, p=0.91					

[a] pseudo R2
[b] pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.
Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.
[+] lowest quintile of SES is the reference category
[^] 1965 is the reference category

Appendix D: Principal Components Analysis and Results

Principal components analysis

As a robustness check Principal Components Analysis (PCA) will also be used to generate a measure of socioeconomic status using the same variables as in the confirmatory factor analysis. PCA is a very similar technique to factor analysis, with the main difference being largely conceptual. PCA is a descriptive data reduction technique that converts a set of correlated variables into a set of uncorrelated variables called principal components. The first component is always the component which describes the largest variance in the data. This procedure is related to factor analysis, but rather than assuming that the observed variables are predicted by the latent variable, the latent variables (the components) are predicted by the observed variables. PCA creates the latent variables based on a linear combination of the observed variables, therefore changing the observed variables will the interpretation of the latent construct. This is in direct contrast to factor analysis.

Visual inspection of the correlation coefficients, as well as Bartlett's tests of sphericity, were used to determine whether the data were appropriate for PCA analysis (Bartlett, 1937). Also the Kaiser-Mayer-Olkin (KMO) test was used to test whether the variables have enough in common to warrant a factor analysis, the high KMO values shown in table A.10 suggest that they do (Kaiser, 1974). The amount of variance accounted for by one principal component is shown in Table A.10. At least half of all the variance in the 5 indicators was accounted for by a single component solution. The appropriateness of a one component solution was investigated using scree plots (Cattell, 1983), shown in Figure A.4. An Eigen value cut off of 1.0 (Kaiser, 1960) was also employed in each data set, and as can be seen in figure A.4, only one component in each data set had an Eigen Value greater than 1.

Figure A.4. Scree plots taken from a polychoric principal component analysis of parental education, social class, home ownership and ratio of rooms in the NCDS, BCS and MCS.

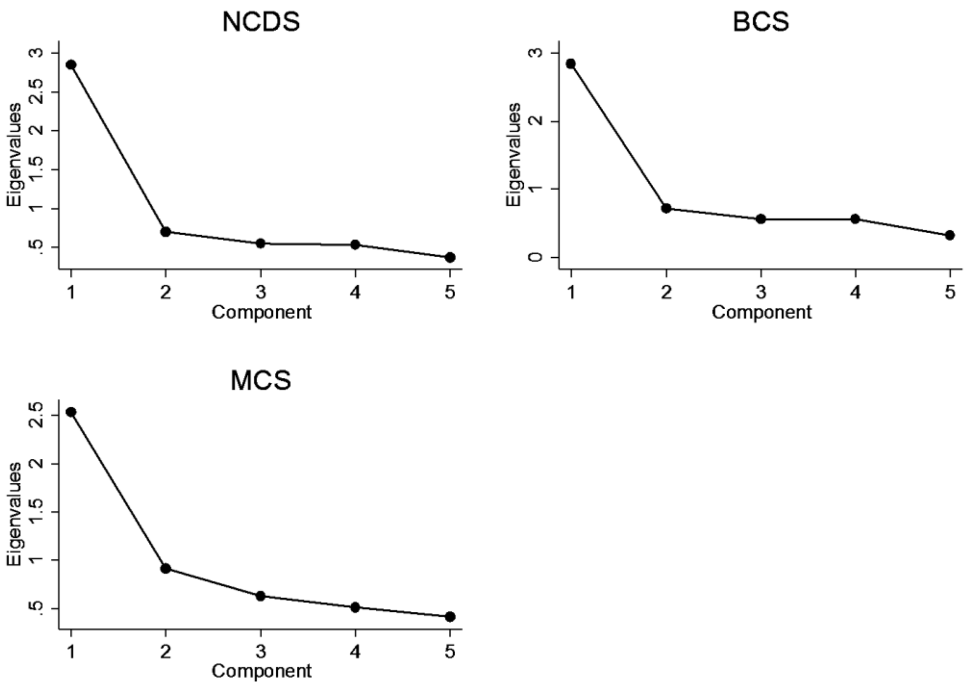


Table A.10. The Kaiser-Mayer-Olkin (KMO) values and proportion of variation explained by a one component solution from a polychoric principal component analysis of parental education, social class, home ownership and ratio of rooms in the NCDS, BCS and MCS.

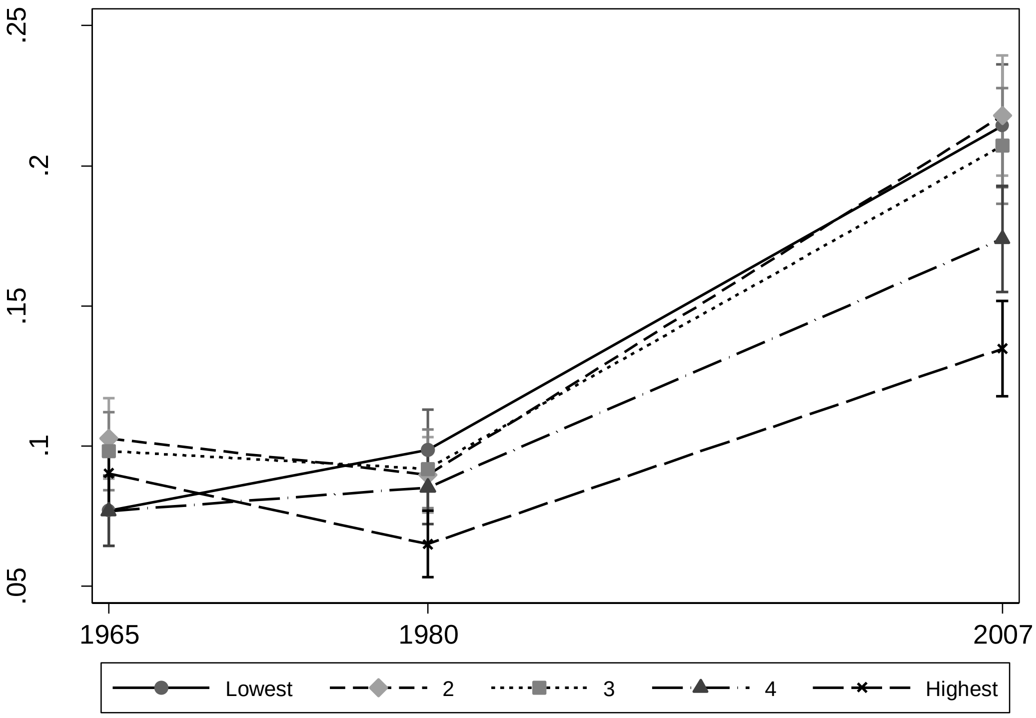
	NCDS	BCS	MCS
n	8717	8413	9948
KMO	0.8188	0.8096	0.77
Variance	57%	56.91%	50.65%

The results of the analysis using PCA instead of a latent measure of SES are presented in table A.11. PCA uses listwise deletion, therefore the sample size is reduced to 27078 children, whose parents had responses to all five variables. As can be seen in table A.11 and figure A.5, the results are substantively no different to using the latent measure of SES, there is a significant interaction between time and SES quintile and it appears that this interaction is being driven by the low prevalence rates of the highest SES group.

Table A.11. The interaction between the quintiles of the SES derived through PCA and time period dummies.

	OR	S.E	z	p	95% confidence Intervals	
PCA SES						
2	1.37	0.16	2.65	0.01	1.09	1.74
3	1.31	0.16	2.22	0.03	1.03	1.65
4	1.00	0.13	-0.01	0.99	0.78	1.28
5	1.19	0.15	1.42	0.16	0.94	1.51
Time period						
1980	1.31	0.16	2.23	0.03	1.03	1.67
2007	3.27	0.36	10.63	0.00	2.63	4.07
PCA SES#Time period						
2 1980	0.66	0.11	-2.51	0.01	0.47	0.91
2 2007	0.74	0.11	-1.96	0.05	0.55	1.00
3 1980	0.71	0.12	-2.04	0.04	0.51	0.99
3 2007	0.73	0.11	-2.04	0.04	0.54	0.99
4 1980	0.85	0.15	-0.91	0.36	0.61	1.20
4 2007	0.77	0.12	-1.63	0.10	0.57	1.05
5 1980	0.53	0.10	-3.52	0.00	0.38	0.76
5 2007	0.48	0.08	-4.65	0.00	0.35	0.65
constant	0.08	0.01	-27.73	0.00	0.07	0.10
n	27 078					
Fit statistics	R2	log pseudolikelihood			BIC	
	0.0321	-10290.71			-255544.606	
Test of interaction	chi(8)=27.02, p<0.001					

Figure A.5. The proportion of children ages 7-10 who were overweight by SES quintile as measured using PCA in 1965, 1980 and 2007.



Appendix E: Comparing different specification of the latent model

Several different specifications for deriving a latent measure of socioeconomic status were run to assess whether using missing data methods, accounting for the variables' level of measurement, including covariances between error terms and taking into account the complex sampling in the MCS would produce very different results. In SES1, 2 & 3 the variables are treated as though they are continuous. In SES 4 & 6 the level of measurement is taken into account, this is done using logit and ordinal logit regression in Stata (SES 4), and probit and ordinal probit regression in Mplus (SES 6). The specification SES 5 shows the results of the principal components analysis which was discussed in Appendix D. Missing data methods are used in specifications 1, 3, 4 & 6, whereas list wise deletion is applied in SES 2 & SES 5.

The correlations between the different specifications are shown in table A.12. The correlations between the different specifications are very high. This means that regardless of the method used to create the composite measure of SES, the results are near identical.

Table A.12. Correlations between different specifications on the measurement of SES.

NCDS	SES1	SES2	SES3	SES4	SES5	SES6
SES1	1					
SES2	0.9998	1				
SES3	0.9937	0.9943	1			
SES4	0.9764	0.9764	0.9528	1		
SES5	0.9963	0.9954	0.9866	0.9788	1	
SES6	0.9832	0.9838	0.9678	0.9961	0.9831	1
BCS	SES1	SES2	SES3	SES4	SES5	SES6
SES1	1					
SES2	0.9992	1				
SES3	0.9991	0.997	1			
SES4	0.991	0.9933	0.9865	1		
SES5	0.9827	0.9866	0.9775	0.9824	1	
SES6	0.9866	0.9901	0.9811	0.9985	0.9833	1
MCS	SES1	SES2	SES3	SES4	SES5	SES6
SES1	1					
SES2	0.9967	1				
SES3	0.9961	0.9912	1			
SES4	0.995	0.9924	0.9901	1		
SES5	0.9958	0.9947	0.9962	0.9934	1	
SES6	0.9899	0.9927	0.9893	0.9944	0.9928	1

SES 1: All variables were treated as continuous, MLMV estimation.
SES 2: All variables treated as continuous, ML estimation.
SES 3: All variables treated as continuous, MLMV estimation, Covariance between home ownership and the ratio of rooms to people fitted.
SES 4: Logit and ordered logit regression used for binary and ordered categorical variables, MLMV estimation.
SES 5: Principal components analysis of the correlation matrix, Polyserial and polychoric correlations were utilised. Listwse deletion was applied.
SES 6: Factor scores were generated in Mplus. Probit and ordered probit regressions were used to account for categorical data using the WLSMV estimator. For MCS cohort stratification, clustering and weighting was applied in Mplus.

Appendix F: Looking at socioeconomic inequalities in the trend of child overweight and obesity in a white only sample

The white only sample was made up of 10496 responses in 1965, 11651 in 1980 and 11241 in 2007. Only 2% of respondents in the NCDS and 4% of respondents in the BCS cohort were classified as an ethnic group other than white. Ethnicity was not well reported in the NCDS cohort and was not measured at the age 7 sweep of data collection. Therefore responses from age 11 were utilised. Still ethnicity was not well recorded, it was recorded during the medical examination at age11 & age16, when the medical officer was asked to decide the child's ethnic background (Euro Caucasian/ African Negroid/ Indian Pakistani/ Other Asian/ Other) based on their features. In the BCS and MCS ethnic categorisation was based on parental self-report of child's ethnicity. To create the white only sample the 'Euro Caucasian' group of the NCDS, the 'English etc, Irish & Other European' in the BCS, and the 'White' group in the MCS were selected. Given the different methods for classifying children's ethnicities and the different categorisations of ethnicity within each cohort it is not definitive that this white only sample represents a homogenous group of people. However, any potential impact of ethnicity on the results will be substantially reduced in this analysis.

Results

There results with the NS-SEC, shown in table A.13 and depicted in figure A.6, were very similar to those presented in the main paper. The interaction between NS-SEC and time was statistically significant (chi2(8) =23.14, $p<0.05$, n=29769). This suggested that differences in the proportion of children within each NS-SEC group classified as overweight changed over time. The depiction of the interaction, figure A.6, shows that, as in the main analysis, children in the highest NS-SEC group had very low increases in overweight between 1980 and 2007. The patterning of the results for the RGSC measure were also very similar to those in the main analysis. The results are shown in table A.14. However, the interaction of RGSC groups with time was not statistically significant (chi2(8)= 12.70, $p>0.05$, n=31363), although as can be seen in figure A.6 the RGSC groups still show the same patterns as in the main analysis, with RGSC II-V showing similar increases in child overweight over time and a very low increase in the prevalence of overweight for the RGSC I group. The non-significance of the interaction may be due to a loss of statistical power due to the smaller sample size.

The results for parental education, shown in table A.15 (mothers) and A.16 (fathers) are very similar to those presented in the main text, there are no significant interactions between mother's education and child overweight, or father's education and child overweight, suggesting that the association between parental education and child overweight has not changed over time. However, the interaction for father's education was approaching significance ($p<0.09$), suggesting that the influence of father's education was becoming more important in the most recent cohort. The latent measure of SES showed very similar patterns to those presented in the main analysis, as can be seen in table A.17. The interaction between the latent SES quintiles and time was statistically significant ($chi2(8)=23.14$, $P<0.05$, $n= 33386$), suggesting that proportion of children classified as overweight within the SES quintiles varied with time. Figure A.6 shows that this difference over time is likely driven by the highest SES quintile and not the lowest.

Table A.13. The interaction between NS-SEC and time period dummies for the white only sample

	OR	S.E	z	p	95% confidence Intervals	
NS-SEC[+]						
2	1.01	0.18	0.05	0.96	0.71	1.43
3	0.83	0.16	-0.95	0.34	0.56	1.22
4	1.23	0.21	1.19	0.23	0.88	1.73
5	1.12	0.19	0.68	0.50	0.81	1.56
6	1.02	0.17	0.10	0.92	0.73	1.42
7	1.24	0.20	1.35	0.18	0.91	1.69
Time period[^]						
1980	0.81	0.15	-1.14	0.26	0.57	1.16
2007	1.87	0.30	3.95	0.00	1.37	2.56
NS-SEC#Time period						
2 1980	1.15	0.27	0.62	0.54	0.73	1.82
2 2007	1.34	0.27	1.46	0.14	0.90	1.99
3 1980	0.92	0.25	-0.29	0.77	0.54	1.57
3 2007	1.75	0.40	2.47	0.01	1.12	2.73
4 1980	1.13	0.26	0.52	0.61	0.72	1.76
4 2007	1.15	0.24	0.66	0.51	0.76	1.72
5 1980	1.19	0.26	0.79	0.43	0.77	1.83
5 2007	1.33	0.28	1.32	0.19	0.87	2.02
6 1980	1.27	0.29	1.05	0.29	0.81	1.97
6 2007	1.90	0.39	3.17	0.00	1.28	2.83
7 1980	1.24	0.26	1.02	0.31	0.82	1.87
7 2007	1.45	0.28	1.90	0.06	0.99	2.13
constant	0.09	0.01	-17.37	0.00	0.07	0.12
n	27 078					
Fit statistics	R2 [a]	log pseudolikelihood [b]			BIC	
	0.0339	-11225.217			-283877.03	
Test of interaction	chi(12)=23.16, p=0.03					

[a] pseudo R2

[b] pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.

Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.

[+]NS-SEC 1 (Higher managerial and professional occupations) is the reference category

[^] 1965 is the reference category

Table A.14. The interaction between RGSC and time period dummies for the white only sample

	OR	S.E	z	p	95% confidence Intervals	
RGSC[+]						
IV	1.05	0.17	0.30	0.76	0.76	1.45
III	1.02	0.15	0.15	0.88	0.76	1.37
III	1.14	0.19	0.76	0.45	0.82	1.57
I	0.92	0.20	-0.38	0.70	0.61	1.40
Time Points[^]						
1980	1.14	0.26	0.59	0.56	0.74	1.77
2007	2.80	0.63	4.53	0.00	1.79	4.36
RGSC#Time points						
IV 1980	0.88	0.23	-0.48	0.63	0.53	1.47
IV 2007	1.07	0.27	0.27	0.79	0.65	1.76
III 1980	0.84	0.20	-0.74	0.46	0.53	1.33
III 2007	0.96	0.23	-0.18	0.85	0.60	1.52
II 1980	0.70	0.18	-1.39	0.16	0.43	1.15
II 2007	0.72	0.18	-1.35	0.18	0.44	1.16
I 1980	0.66	0.21	-1.29	0.20	0.36	1.24
I 2007	0.63	0.19	-1.52	0.13	0.35	1.14
constant	0.09	0.01	-16.70	0.00	0.07	0.12
n	31363					
Fit statistics	R2 [a]	log pseudolikelihood [b]			BIC	
	0.0319	-11655.214			-301154.278	
Test of interaction	chi(8)=12.70, p=0.1224					

[a] pseudo R2
[b] pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.
Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.
[+]RGSC V is the reference category
[^] 1965 is the reference category

Table A.15. The interaction between whether or not the mother stayed at school past the minimum age and time period dummies for the white only sample

	OR	S.E	z	p	95% confidence Intervals	
Mum School - stayed[+]	0.77	0.07	-3.08	0.00	0.65	0.91
Time Point[^]						
1980	1.00	0.06	-0.07	0.94	0.89	1.12
2007	2.67	0.15	17.53	0.00	2.40	2.99
mumschool# Time period						
stayed # 1980	0.99	0.12	-0.05	0.96	0.79	1.25
stayed # 2007	0.96	0.10	-0.35	0.72	0.79	1.18
constant	0.10	0.00	-57.52	0.00	0.09	0.11
n	31203					
Fit statistics	$R2^{a}$	log pseudolikelihood [b]			BIC	
	0.0311	-11766.962			-299238.95	
Test of interaction		chi(2)=0.17, p=0.92				

[a] pseudo R2
[b] pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.
Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.
[+] did not stay at school past the minimum school leaving age is the reference category
[^] 1965 is the reference category

Table A.16. The interaction between whether or not the father stayed at school past the minimum age and time period dummies for the white only sample

	OR	S.E	z	p	95% confidence Intervals	
Dad School - stayed[+]	0.86	0.07	-1.82	0.07	0.72	1.01
Time Point[^]						
1980	1.03	0.06	0.53	0.60	0.92	1.15
2007	2.61	0.15	16.95	0.00	2.34	2.92
Dadschool# Time period						
stayed # 1980	0.85	0.10	-1.31	0.19	0.68	1.08
stayed # 2007	0.79	0.09	-2.17	0.03	0.64	0.98
constant	0.10	0.00	-58.21	0.00	0.09	0.11
n	28236					
Fit statistics	$R2^{a}$	log pseudolikelihood [b]			BIC	
	0.0279	-10521.327			-268206.862	
Test of interaction		chi(2)=4.71, p=0.09				

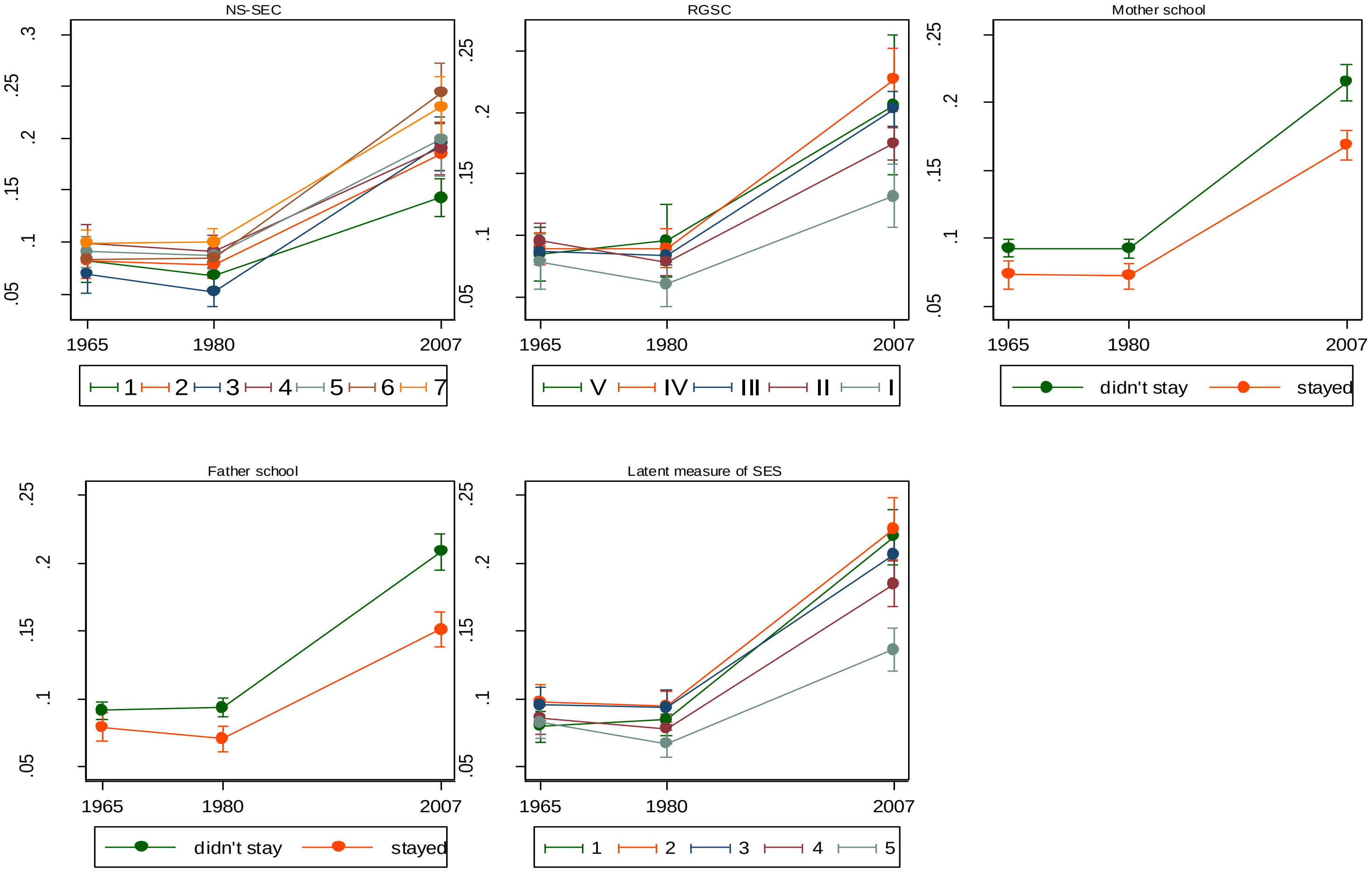
[a] pseudo R2
[b] pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.
Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.
[+] did not stay at school past the minimum school leaving age is the reference category
[^] 1965 is the reference category

Table A.17. The interaction between the quintiles of the latent measures of SES and time period dummies for the white only sample

	OR	S.E	z	p	95% confidence Intervals	
SES quintiles[+]						
2	1.25	0.14	2.02	0.04	1.01	1.55
3	1.23	0.13	1.97	0.05	1.00	1.52
4	1.09	0.12	0.79	0.43	0.88	1.36
5	1.05	0.12	0.41	0.68	0.84	1.30
Time Period[^]						
1980	1.07	0.12	0.64	0.52	0.87	1.33
2007	3.24	0.32	11.88	0.00	2.67	3.94
SES quintiles# Time period						
2 1980	0.90	0.13	-0.69	0.49	0.67	1.21
2 2007	0.83	0.12	-1.30	0.19	0.63	1.10
3 1980	0.91	0.14	-0.64	0.52	0.68	1.22
3 2007	0.75	0.10	-2.07	0.04	0.58	0.98
4 1980	0.84	0.13	-1.14	0.25	0.62	1.13
4 2007	0.74	0.10	-2.15	0.03	0.57	0.97
5 1980	0.74	0.12	-1.89	0.06	0.55	1.01
5 2007	0.54	0.08	-4.33	0.00	0.41	0.71
constant	0.09	0.01	-31.40	0.00	0.07	0.10
n	33386					
Fit statistics	R2 [a]	log pseudolikelihood [b]			BIC	
	0.0334	-12335.484			-322824.018	
Test of interaction	chi(8)=20.34, p=0.009					

[a] pseudo R2
[b] pseudo likelihoods used for parameter estimates when survey data are weighted. Likelihood ratio tests are not valise for pseudo likelihoods.
Probability weights are held at a constant (1) for NCDS and BCS cohorts and given the values of 'dovwt2' for the MCS cohort.
[+]SES quintile 1 (the lowest 20% of the SES distribution) is the reference category
[^] 1965 is the reference category

Figure A.6. Depictions of the interactions between different measures of SES and time in the proportion of children classified as overweight.



Appendix G: Social class for the household

Emily Beller (2009) provides a thorough discussion on the merits of including social class information from both parents in the household, not just the father figure. She uses indices of model fit to demonstrates that in cases where information from both mother's and father's social class is provided in intergenerational mobility models, model fit is improved. Whilst the research presented here is not intergenerational mobility research, it does concern intergenerational relationships. The increasing rates of labour force participation among mother as well as increasing diversity in family forms mean that it may no longer be adequate, or even possible, to classify a household's social position based on the father's social class.

In this appendix I assess model fit for using a highest household social class (also known as high class dominance), using an average of the mother's and father's social class, or using mother's and father's social class as independent covariates. The NS-SEC 7 category version and three category version were used to make a measure of highest household social class. The three category version of the NS-SEC can be conceptualised as ordinal, therefore mothers and fathers level of social class were compared on this measure. Father's 7 category NS-SEC was used to generate an initial value for highest household social class. In instances where mothers had a higher social class (as defined by a lower figure on the three category version of the NS-SEC), or where there was no father figure in the household, highest household social class was replaced with the mother's social class. In order to make a measure of average social class, the average social class was crudely derived from two parent households by adding social class together and dividing by two. The social class of the parent in lone parent households was included after this averaging.

As in the main text, a logistic regression model is estimated, with overweight and obese/ not overweight and obese as the outcome measure. The results are presented in table A.18. Specification 1 refers to a model where main and partner respondent's education are entered into the model as separate covariates. Highest household social class is used in specification 2, and average social class sin specification 3. Income is included as a covariate in the models, as this is the main variable of interest in chapter

5. The coefficients and model fit are estimated for a model containing just income and the different measures of social class (a), and with demographic characteristics of the parents and their education levels (b).

Table A.18. Different specifications of social class. Presented as odds ratios.

	Spec 1a	Spec1b	Spec2a	Spec2b	Spec3a	Spec3b
TA income quint 1	Reference					
TA income quint 2	1.16 (0.09)+	1.21 (0.10)*	1.14 (0.09)+	1.22 (0.10)*	1.13 (0.09)	1.21 (0.10)*
TA income quint 3	1.10 (0.09)	1.18 (0.11)+	1.05 (0.09)	1.18 (0.11)+	1.05 (0.09)	1.18 (0.11)+
TA income quint 4	1.04 (0.09)	1.15 (0.12)	0.97 (0.09)	1.15 (0.12)	0.99 (0.09)	1.17 (0.12)
TA income quint 5	0.90 (0.10)	1.03 (0.12)	0.81 (0.09)*	1.02 (0.12)	0.84 (0.09)	1.05 (0.12)
Main NS-SEC 1	Reference					
Main NS-SEC 2	1.08 (0.13)	1.06 (0.13)				
Main NS-SEC 3	1.04 (0.13)	1.00 (0.14)				
Main NS-SEC 4	0.99 (0.15)	0.96 (0.16)				
Main NS-SEC 5	1.45 (0.22)*	1.40 (0.23)*				
Main NS-SEC 6	1.19 (0.16)	1.17 (0.16)				
Main NS-SEC 7	1.07 (0.15)	1.05 (0.16)				
Main class missing	1.20 (0.19)	1.11 (0.19)				
Part NS-SEC 1	Reference (0.00)	1.00 (0.00)				
Part NS-SEC 2	1.23 (0.12)*	1.20 (0.11)+				
Part NS-SEC 3	1.03 (0.16)	0.95 (0.14)				
Part NS-SEC 4	1.27 (0.12)*	1.15 (0.11)				
Part NS-SEC 5	1.08 (0.12)	0.98 (0.11)				
Part NS-SEC 6	1.35 (0.17)*	1.19 (0.16)				
Part NS-SEC 7	1.44 (0.17)**	1.25 (0.15)+				
Part class missing	1.43 (0.15)***	1.01 (0.19)				
HH NS-SEC 1			Reference			
HH NS-SEC 2			1.19 (0.09)*	1.17 (0.09)*		
HH NS-SEC 3			1.17 (0.12)	1.07 (0.11)		

HH NS-SEC 4			1.20	1.14		
			(0.14)	(0.14)		
HH NS-SEC 5			1.60	1.44		
			(0.25)**	(0.23)*		
HH NS-SEC 6			1.42	1.31		
			(0.15)**	(0.14)*		
HH NS-SEC 7			1.17	1.07		
			(0.14)	(0.13)		
HH class missing			1.57	1.31		
			(0.28)*	(0.23)		
Average 1					Reference	
Average 1.5					1.11	1.17
					(0.21)	(0.23)
Average 2.0					1.18	1.12
					(0.19)	(0.19)
Average 2.5					1.11	1.12
					(0.22)	(0.23)
Average 3.0					1.28	1.19
					(0.22)	(0.21)
Average 3.5					1.17	1.16
					(0.21)	(0.22)
Average 4					1.13	1.09
					(0.20)	(0.20)
Average 4.5					1.24	1.23
					(0.23)	(0.23)
Average 5					1.69	1.62
					(0.32)**	(0.32)*
Average 5.5					1.27	1.27
					(0.26)	(0.27)
Average 6					1.46	1.33
					(0.27)*	(0.26)
Average 6.5					1.70	1.64
					(0.36)*	(0.36)*
Average 7					1.33	1.21
					(0.25)	(0.24)
Average missing					1.44	1.25
					(0.26)*	(0.24)
N	13,799	13,799	13,799	13,799	13,799	13,799
Log pseudolikelihood	-6885.916	-6816.186	-6895.018	-6821.917	-6892.107	-6818.476
Pseudo R2	0.0070	0.0171	0.0057	0.0162	0.0061	0.0167
BIC	-117555	-117285	-117613	-117350	-117572	-117309

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Spec 1 = main and partner respondents social class entered separately
Spec 2 = highest household social class
Spec 3 = Average social class
a = Time averaged income + social class
b = spec a+ Ethnicity, longstanding illness/disability, region, parental age, Main respondent education, Partner respondent education.

Model fit statistics suggested that specification 2 results in a better fitting model and more parsimonious model. The difference in BIC values between specification 1 and 2 of -58.054 provides strong support for specification 2 over specification 1. The difference between specification 2 and 3 of -41.840, also provides support for specification 2 over specification 3. Social class is not the main variable of interest in this particular chapter, but it does play an important role in the model and so it is essential that it is measured well and captures information at the household level. The smaller number of categories in the highest social class variable provides additional power to detect statistically significant differences between the social classes. Where mothers and social class are both controlled for in the model (specification 1), the interpretation of the coefficients can become confusing. For example the coefficients on main respondent's social class are the differences in the odds where partner respondent's social class is held at an average value. Where the average social class position is taken, the social class groupings of people may not represent a heterogeneous group of people. A value of 3.5 for example can be achieved where one parent has a value of 1 and the other a value of 7, equally it can be achieved when one parent has a value of 4 and the other a value of 3. Whilst the latter case is more likely, the former is still possible.

Given the findings from this investigation into different ways of classifying social class in the household, the highest household social class is used in chapter 5.

Appendix H: Complete case analysis

The use of ‘ad-hoc’ methods for dealing with missing data have been criticised. Complete case analysis has been suggested by some as better than using ad-hoc methods for dealing with missing data. Therefore the results are presented here for those who have responses to all the covariates.

However, as I state in the main text of this thesis, complete case analysis will likely result in a biased sample. Table A.19 shows the characteristics of the sample used in the complete case analysis, and those used in the main analysis. The sample of people who responded to all the covariates have on average higher equivalised incomes (about ¼ of a SD higher), a higher level of education, and a higher proportion are in managerial and professional occupations. None of the cases included in the complete case analysis are lone parents.

The results, presented in table A.20 for males and females separately, show that for this sample the lowest income quintile does not have the highest proportion of overweight. For both males and females, children in income quintile 2 have the highest odds of overweight. For females there are significant differences in the odds of being overweight in income quintile 2 compared to income quintile 1, after accounting for parental education, and this difference remains after adjustment for social class. For males no statistically significant differences in the odds remains after adjustment for parental education.

These results confirm the main analysis in that there is no evidence that children in the lowest income quintile have higher odds of being overweight. However for females those in the second income quintile have an increased risk of being overweight independent of demographic characteristics, parental education and social class. In the main analysis it was males that showed this pattern. The influence of parental employment patterns was investigated in the main analysis, with no clear findings that this was driving the results. It is investigated here as well.

Table A.19. Comparisons of the complete case sample and the sample used in the main analysis

	Sample in main analysis	Complete case sample
N	**13799**	**9198**
Overweight	20%	19%
Equivalised Income Mean (SD)	346 (197)	400 (198)
Quintiles of Income	Mean (range)	Mean (range)
Quint 1	134 (40 - 172)	161 (40 - 214)
Quint 2	214 (172 - 258)	262 (214 - 308)
Quint 3	307 (258 - 360)	357 (308 - 407)
Quint 4	424 (360 - 502)	467 (407 - 543)
Quint 5	666 (502 - 1258)	703 (543 -1207)
Main Respondent Education	%	%
NVQ 1	8	6
NVQ 2	27	26
NVQ 3	15	15
NVQ 4	29	34
NVQ 5	6	7
Overseas	3	3
None	11	8
Social Class	%	%
High Manag/prof	16	16
Lower Manag/prof	28	40
Intermediate	13	12
Small emp/self employed	9	11
Lower Sup & Technical	7	3
Semi Routine	13	11
Routine	7	7
No Classification	8	
Longstanding Illness/disability	%	%
yes	19	18
Ethnicity	%	%
White	86	88
Other Ethnic group	3	2
Black/Black British	2	2
Pakistani/Bangladeshi	4	4
Indian	3	1
Mixed	1	1
Sex	%	%
Female	49	49
Male	51	51

Table A.20. Results for full case analysis. Results presented as odds ratios.

Female	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Tran Income Q2	1.04 (0.15)					
Tran Income Q3	1.03 (0.16)					
Tran Income Q4	0.96 (0.13)					
Tran Income Q5	0.80 (0.12)					
TA Income Q2		1.25 (0.17)+	1.24 (0.17)	1.29 (0.18)+	1.33 (0.19)*	1.33 (0.19)*
TA Income Q3		1.04 (0.13)	1.04 (0.14)	1.11 (0.16)	1.17 (0.17)	1.18 (0.18)
TA Income Q4		0.88 (0.12)	0.90 (0.13)	0.97 (0.15)	1.07 (0.17)	1.08 (0.18)
TA Income Q5		0.77 (0.10)+	0.79 (0.12)	0.88 (0.15)	1.03 (0.17)	1.03 (0.18)
N	4,547	4,549	4,549	4,549	4,548	4,548
Male	**Model 0**	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Tran Income Q2	1.18 (0.18)					
Tran Income Q3	1.26 (0.20)					
Tran Income Q4	0.97 (0.16)					
Tran Income Q5	0.84 (0.12)					
TA Income Q2		1.21 (0.18)	1.20 (0.19)	1.21 (0.20)	1.24 (0.21)	1.23 (0.21)
TA Income Q3		0.95 (0.15)	0.94 (0.16)	0.96 (0.17)	1.03 (0.19)	1.00 (0.19)
TA Income Q4		0.83 (0.12)	0.79 (0.13)	0.84 (0.15)	0.91 (0.17)	0.88 (0.17)
TA Income Q5		0.70 (0.10)*	0.64 (0.11)**	0.73 (0.13)+	0.82 (0.16)	0.81 (0.16)
N	4,647	4,649	4,649	4,649	4,649	4,649

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 0 = Transitory Income; Model 1 = Time averaged Income; Model 2 = Model1 + Ethnicity, longstanding illness/disability, region, parental age; Model 3 = Model 2 + Main respondent education; Model 4 = Model 3 + Partner respondent education; Model 5 = Model 4 + Highest social class of parents.
Estimates weighted using dovwt2 survey design & attrition weight.

Whether or not the main and partner respondents were in employment was attained from the variables “dmpjob00” (main respondent) and “dppjob00” (partner respondent). These are self-reported to responses to whether or not he main and partner respondents are in work, including part-time work. These variables were entered into the model sequentially, and the results are presented in table A.21. After controlling for whether the main respondent was in employment, the difference in the odds between the lowest and second lowest income quintile was reduced (OR=1.26) and the difference was no longer statistically significant. Where the main respondent was not in employment the odds of the child being overweight were lower (0.83). This suggests that whilst having a working mother results in a higher income, it is also associated with increased odds in the child being overweight. The inclusion of partner respondents employment status further reduced the difference in the odds between income quintile 1 and 2 (OR=1.21).

Table A.21. Employment status of parents and parental income.

Females	Model 6	Model 7
TA income Q1	reference	
TA Income Q2	1.26 (0.19)	1.21 (0.19)
TA Income Q3	1.11 (0.18)	1.06 (0.17)
TA Income Q4	1.01 (0.17)	0.96 (0.17)
TA Income Q5	0.97 (0.17)	0.92 (0.17)
Main in employment	Reference	
Main not in employment	0.83 (0.08)+	0.84 (0.08)+
Part in employment		Reference
Part not in employment		0.81 (0.14)
	4,548	4,548

+ $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

Appendix I: Full list of coefficients for chapter 5

Table A.22. Full list of coefficients for male children regarding the relationship between income and child overweight status.

Male	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Trans[a] Income Quintile 1	Reference					
Trans Income Quintile 2	1.20 (0.13)+					
Trans Income Quintile 3	1.23 (0.14)+					
Trans Income Quintile 4	0.97 (0.11)					
Trans Income Quintile 5	0.83 (0.10)					
Trans Income Miss	1.45 (1.76)					
TA[b] Income Quintile 1		Reference				
TA Income Quintile 2		1.22 (0.14)+	1.29 (0.15)*	1.31 (0.17)*	1.33 (0.17)*	1.36 (0.18)*
TA Income Quintile 3		1.09 (0.13)	1.22 (0.17)	1.26 (0.19)	1.33 (0.20)+	1.38 (0.22)*
TA Income Quintile 4		0.87 (0.09)	0.98 (0.13)	1.06 (0.16)	1.15 (0.18)	1.23 (0.20)
TA Income Quintile 5		0.75 (0.09)*	0.83 (0.11)	0.95 (0.14)	1.09 (0.17)	1.20 (0.22)
White			Reference			
Other Ethnic			0.79 (0.26)	0.80 (0.26)	0.82 (0.27)	0.80 (0.26)

Black/Black British	1.58	1.67	1.70	1.66
	(0.29)*	(0.31)**	(0.32)**	(0.32)**
Pakistani/Bangladeshi	1.17	1.20	1.18	1.16
	(0.20)	(0.21)	(0.21)	(0.22)
Indian	1.31	1.35	1.33	1.33
	(0.34)	(0.35)	(0.35)	(0.35)
Mixed	1.79	1.85	1.91	1.85
	(0.36)**	(0.37)**	(0.39)**	(0.37)**
Child illness NO	Reference			
Child Illness Missing	1.01	1.13	1.21	1.19
	(0.88)	(1.00)	(1.06)	(1.05)
Child Illness YES	1.10	1.10	1.10	1.10
	(0.11)	(0.11)	(0.11)	(0.11)
Partner Illness NO	Reference			
Partner Illness YES	0.97	0.97	0.97	0.97
	(0.11)	(0.11)	(0.11)	(0.11)
Partner illness Missing	1.03	1.03	0.88	0.87
	(0.12)	(0.12)	(0.12)	(0.12)
Main Illness NO	Reference			
Min Illness YES	1.25	1.24	1.24	1.23
	(0.10)**	(0.10)**	(0.10)**	(0.10)*
Main Illness Missing	1.07	0.99	0.94	0.98
	(0.91)	(0.85)	(0.81)	(0.84)
London	Reference			
North East	0.86	0.89	0.92	0.89
	(0.15)	(0.15)	(0.15)	(0.14)
North West	0.74	0.75	0.77	0.76
	(0.13)+	(0.13)	(0.13)	(0.13)+
Yorkshire	0.88	0.90	0.91	0.91
	(0.12)	(0.12)	(0.13)	(0.13)
East Midlands	0.80	0.81	0.82	0.83

	(0.14)	(0.14)	(0.14)	(0.15)
West Midlands	0.88	0.91	0.92	0.88
	(0.14)	(0.15)	(0.15)	(0.15)
East of England	0.76	0.75	0.76	0.75
	(0.12)+	(0.12)+	(0.12)+	(0.12)+
South East	0.78	0.79	0.81	0.80
	(0.11)+	(0.11)+	(0.11)	(0.11)
South West	0.78	0.80	0.83	0.83
	(0.13)	(0.13)	(0.13)	(0.13)
Wales	0.97	1.00	1.02	1.02
	(0.14)	(0.14)	(0.14)	(0.14)
Scotland	0.83	0.85	0.88	0.86
	(0.12)	(0.12)	(0.12)	(0.12)
Northern Ireland	1.13	1.18	1.20	1.17
	(0.17)	(0.18)	(0.18)	(0.18)
Partner age	1.01	1.01	1.01	1.01
	(0.01)	(0.01)	(0.01)	(0.01)
Main age	1.00	1.01	1.01	1.01
	(0.01)	(0.01)	(0.01)	(0.01)
Missing age dummy	1.08	1.09	0.88	0.90
	(0.15)	(0.15)	(0.19)	(0.20)
Main NVQ level 1		Reference		
Main NVQ level 2		1.19	1.19	1.21
		(0.17)	(0.17)	(0.17)
Main NVQ level 3		1.03	1.05	1.05
		(0.16)	(0.16)	(0.16)
Main NVQ level 4		0.86	0.90	0.90
		(0.14)	(0.14)	(0.15)
Main NVQ level 5		0.76	0.82	0.84
		(0.16)	(0.18)	(0.19)
Main Overseas Quals[c]		1.35	1.36	1.35

	(0.29)	(0.29)	(0.29)
Main No Quals	0.96	0.94	0.93
	(0.17)	(0.17)	(0.17)
Partner NVQ level 1		Reference	
Part NVQ level 2		0.71	0.73
		(0.12)*	(0.12)+
Part NVQ level 3		0.65	0.69
		(0.12)*	(0.13)*
Part NVQ level 4		0.66	0.72
		(0.12)*	(0.13)+
Part NVQ level 5		0.54	0.60
		(0.12)**	(0.14)*
Part Overseas Quals		0.77	0.77
		(0.21)	(0.21)
Part No Quals		0.98	0.96
		(0.19)	(0.19)
Part Ed Missing		1.08	1.21
		(0.28)	(0.37)
HH[d] NS-SEC 1			Reference
HH NS-SEC 2			1.19
			(0.13)
HH NS-SEC 3			1.23
			(0.19)
HH NS-SEC 4			1.10
			(0.19)
HH NS-SEC 5			1.36
			(0.34)
HH NS-SEC 6			1.27
			(0.19)
HH NS-SEC 7			1.02

						(0.18)
HH NS-SEC missing						1.69 (0.43)*
Constant	0.206	0.219	0.125	0.110	0.134	0.107
N	6,969	6,969	6,969	6,969	6,969	6,969

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

[a] Transitory income, income measured at one time point.
[b] time averaged income, income averaged over several time points.
[c] qualifications
[d] Household social class

Table A.23. Full list of coefficients for male children regarding the relationship between income and child overweight status. Results presented as odds ratios.

Female	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Trans[a] Income Quintile 1	Reference					
Trans Income Quintile 2	0.98					
	(0.10)					
Trans Income Quintile 3	0.96					
	(0.10)					
Trans Income Quintile 4	0.82					
	(0.08)+					
Trans Income Quintile 5	0.73					
	(0.08)**					
Trans Income Miss	3.72					
	(2.94)+					
TA[b] Income Quintile 1		Reference				
TA Income Quintile 2		1.00	1.06	1.09	1.09	1.10
		(0.11)	(0.12)	(0.12)	(0.12)	(0.13)
TA Income Quintile 3		0.89	0.95	1.00	1.02	1.06
		(0.09)	(0.10)	(0.12)	(0.12)	(0.13)
TA Income Quintile 4		0.84	0.91	0.99	1.04	1.11
		(0.09)	(0.11)	(0.13)	(0.13)	(0.15)
TA Income Quintile 5		0.64	0.69	0.76	0.85	0.92
		(0.07)***	(0.09)**	(0.11)+	(0.12)	(0.14)
White			Reference			
Other Ethnic			0.64	0.65	0.67	0.65

	(0.20)	(0.21)	(0.21)	(0.21)
Black/Black British	2.08	2.14	2.21	2.23
	(0.43)***	(0.44)***	(0.46)***	(0.47)***
Pakistani/Bangladeshi	0.98	0.99	1.02	1.02
	(0.15)	(0.16)	(0.16)	(0.16)
Indian	1.09	1.11	1.19	1.20
	(0.27)	(0.28)	(0.31)	(0.31)
Mixed	0.97	1.01	1.04	1.03
	(0.20)	(0.21)	(0.22)	(0.21)
Child illness NO	Reference			
Child Illness Missing	0.58	0.57	0.58	0.57
	(0.77)	(0.78)	(0.79)	(0.79)
Child Illness YES	1.15	1.15	1.15	1.15
	(0.10)	(0.10)	(0.10)	(0.10)
Partner Illness NO	Reference			
Partner Illness YES	0.96	0.96	0.96	0.96
	(0.10)	(0.10)	(0.10)	(0.10)
Partner illness Missing	1.12	1.12	1.04	1.04
	(0.12)	(0.12)	(0.12)	(0.12)
Main Illness NO	Reference			
Min Illness YES	1.10	1.10	1.09	1.09
	(0.09)	(0.09)	(0.09)	(0.09)
Main Illness Missing	1.87	1.94	1.93	1.92
	(2.28)	(2.44)	(2.40)	(2.38)
London	Reference			
North East	0.93	0.93	0.94	0.92
	(0.20)	(0.20)	(0.20)	(0.20)
North West	0.93	0.93	0.93	0.92
	(0.14)	(0.14)	(0.14)	(0.14)

Yorkshire	0.91	0.91	0.92	0.91
	(0.14)	(0.14)	(0.14)	(0.14)
East Midlands	0.76	0.76	0.77	0.75
	(0.12)+	(0.12)+	(0.12)+	(0.12)+
West Midlands	0.81	0.82	0.81	0.80
	(0.13)	(0.13)	(0.13)	(0.13)
East of England	0.77	0.77	0.76	0.75
	(0.14)	(0.14)	(0.14)	(0.14)
South East	0.77	0.77	0.77	0.76
	(0.13)	(0.13)	(0.13)	(0.13)
South West	0.67	0.68	0.69	0.68
	(0.11)*	(0.12)*	(0.12)*	(0.12)*
Wales	1.09	1.10	1.12	1.10
	(0.17)	(0.17)	(0.18)	(0.18)
Scotland	0.79	0.81	0.81	0.79
	(0.13)	(0.13)	(0.13)	(0.13)
Northern Ireland	1.02	1.03	1.04	1.01
	(0.16)	(0.16)	(0.16)	(0.16)
Partner age	1.00	1.00	1.00	1.00
	(0.01)	(0.01)	(0.01)	(0.01)
Main age	1.01	1.01	1.01	1.01
	(0.01)	(0.01)	(0.01)	(0.01)+
Missing age dummy	0.98	0.99	0.90	0.90
	(0.11)	(0.11)	(0.16)	(0.16)
Main NVQ level 1		Reference		
Main NVQ level 2		0.84	0.84	0.85
		(0.12)	(0.12)	(0.12)
Main NVQ level 3		0.75	0.76	0.77
		(0.12)+	(0.13)	(0.13)
Main NVQ level 4		0.79	0.83	0.84

	(0.12)	(0.13)	(0.14)
Main NVQ level 5	0.68	0.74	0.74
	(0.15)+	(0.16)	(0.17)
Main Overseas Quals[c]	0.84	0.87	0.87
	(0.22)	(0.23)	(0.23)
Main No Quals	0.88	0.86	0.86
	(0.15)	(0.14)	(0.15)
Partner NVQ level 1		Reference	
Part NVQ level 2		0.95	0.96
		(0.16)	(0.17)
Part NVQ level 3		1.00	1.01
		(0.21)	(0.21)
Part NVQ level 4		0.77	0.79
		(0.13)	(0.14)
Part NVQ level 5		0.67	0.69
		(0.13)*	(0.14)+
Part Overseas Quals		0.65	0.65
		(0.15)+	(0.16)+
Part No Quals		1.08	1.08
		(0.22)	(0.21)
Part Ed Missing		1.09	1.10
		(0.28)	(0.28)
HH[d] NS-SEC 1			Reference
HH NS-SEC 2			1.16
			(0.13)
HH NS-SEC 3			0.93
			(0.14)
HH NS-SEC 4			1.18
			(0.19)

HH NS-SEC 5						1.54 (0.33)*
HH NS-SEC 6						1.36 (0.22)+
HH NS-SEC 7						1.13 (0.19)
HH NS-SEC missing						1.07 (0.25)
Constant	0.328	0.337	0.251	0.282	0.281	0.225
N	6,969	6,969	6,969	6,969	6,969	6,969

+ $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

[a] Transitory income, income measured at one time point.
[b] time averaged income, income averaged over several time points.
[c] qualifications
[d] Household social class

Appendix J: Robustness check: Considering the frequency of income poverty and its relationship with child overweight and obesity

This appendix considers the relationship between measures of poverty and child overweight and obesity. As stated in the main text poverty in this instance refers to 'relative income poverty', whereby people who fall below a certain point on the equivalised income distribution, they are deemed to have sufficient income to fully participate in normal life within a given society (Bradshaw *et al.*, 2008).

In the main text I present the results for those who are in persistent poverty, i.e those who fall below the poverty line on all four measurement occasions, compared to those who do not. The families who are classified as not being in poverty at all four time occasions are likely a diverse group of people, including those who were in poverty on three out of four occasions, and those who were never classified as in poverty. Therefore to look at the relationship between poverty and child overweight and obesity in more detail, here I consider different frequencies of being in poverty.

The frequency of poverty is coded as '0' for those who were never in poverty, "1" for those who were in poverty on one occasion, "2" for those in poverty on two occasions and so on. Families were excluded from the analysis where they did not report the information sufficient to calculate poverty status at any sweep. If persistent poverty is associated with child overweight it would be expected that families who are in poverty in all four or three out of the four sweeps would have higher odds of their child being overweight than families who were never in poverty.

The results from this analysis are presented in table A.24, separately for males and females. The unadjusted relationship (model 1), suggests that for males, being in poverty on 1,2, 3 and 4 occasions results in higher odds of being overweight than children who are never in poverty. The differences in the odds are statistically significant for male children in families who were in poverty on 1 or 3 occasions. After adjustment for demographic characteristics, the differences in the odds are reduced, shown by the odds ratios moving closer to 1. Male children whose families are in poverty at 1 time point have significantly higher odds of being overweight than children who are never in poverty. The differences in the odds are no longer statistically significant after adjustment for main respondent and partner respondent's education.

Table A.24. The relationship between the frequency of being categorised as in poverty and child overweight

Males	Model 1	Model 2	Model 3	Model 4	Model 5
Poverty 0 times	reference				
Poverty 1 times	1.33 (0.16)*	1.30 (0.16)*	1.25 (0.16)+	1.17 (0.15)	1.17 (0.15)
Poverty 2 times	1.16 (0.17)	1.09 (0.17)	1.04 (0.16)	0.97 (0.15)	0.95 (0.15)
Poverty 3 times	1.35 (0.19)*	1.24 (0.20)	1.18 (0.20)	1.07 (0.18)	1.07 (0.19)
Poverty 4 times	1.18 (0.14)	1.04 (0.15)	0.98 (0.15)	0.89 (0.13)	0.87 (0.14)
N	5,884	5,884	5,884	5,884	5,884
Females					
Poverty 0 times	reference				
Poverty 1 times	1.07 (0.11)	1.06 (0.12)	1.00 (0.11)	0.96 (0.11)	0.92 (0.10)
Poverty 2 times	1.47 (0.19)**	1.43 (0.18)**	1.33 (0.17)*	1.24 (0.16)+	1.14 (0.15)
Poverty 3 times	1.32 (0.18)*	1.27 (0.19)	1.15 (0.18)	1.07 (0.17)	0.98 (0.16)
Poverty 4 times	1.28 (0.15)*	1.15 (0.17)	1.01 (0.16)	0.94 (0.14)	0.86 (0.14)
N	5,830	5,830	5,830	5,830	5,830

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 1 = Frequency of poverty; Model 2 = Model1 + Ethnicity, longstanding illness/disability, region, parental age; Model 3 = Model 2 + Main respondent education; Model 4 = Model 3 + Partner respondent education; Model 5 = Model 4 + Highest social class of parents.
Estimates weighted using dovwt2 survey design & attrition weight

For female children there were higher odds of being classified as overweight when families were classified as in poverty on 2, 3 or 4 occasions compared to families that were never classified as in poverty. After adjustment for demographic characteristics (model 2), the differences in the odds were reduced, but there were still significant differences in the odds of being overweight for children from families who were in poverty on 2 occasions compared to families that were never in poverty. After adjustment for main respondent education (model 3), the differences in the odds were reduced further, and the estimated odds ratios suggest no differences in the odds of being overweight when the family was in poverty on 1 or all four occasions compared to families that were never in poverty. Further adjusting for partner respondent's education (model 4) further reduces the differences in the odds for families who were classified as in poverty on 2 or 3 occasions. Furthermore in this specification of the model, none of the differences in the odds are statistically significant, however the difference between families who were classified as in poverty on 2 occasions and those who were never in poverty is approaching significance. Controlling for social class further attenuates the difference in the odds for families who were in poverty on 2 occasions.

Appendix K: Robustness check: Applying Karlson, Holm & Breen's method for making logit coefficients comparable

Directly comparing coefficients between nested logit or probit models is erroneous (Karlson, Holm & Breen, 2012). The difference in the magnitude of coefficients between nested models may not be the result of confounding or mediation, but may be due to rescaling of the model. They suggest that if the added covariate 'z' is correlated to the outcome variable 'y', it will change the magnitude of the coefficients on the variable of interest 'x' even if the newly added covariate and the variable of interest are not correlated. This is because the 'z' covariate will reduce the variance in the error term.

Karlson, Holm & Breen (KHB) (2012) introduce a novel solution which decomposes the change in the coefficients into a rescaling component and a confounding component, as well as creating a method for formally testing the degree of confounding. The method is to compare the full model with a reduced model that substitutes some z variables with the residuals of the z variables from a regression of the z variables on the variable of interest 'x', please refer to Karlson, Holm and Breen (2012) for full explanation and details.

The results using the KHB method are shown in table A.25. In table A.25 the reduced model in model 1 refers to the model with only income quintiles in, and the full model refers to the model with demographic characteristics included (parental longstanding illness/disability, parental age, region of residence, whether child has longstanding illness/disability, and parental report of child's ethnicity). In model 2 the full model includes everything in the full model 1 and main respondent's education. In model 3 the full model includes everything in model 2 and partner respondent's education. In Model 4 the full model includes everything in model 3 and highest household social class. The difference test formally tests the hypothesis that the effect of confounding, net of rescaling, is zero.

Table A.25. Coefficients of the reduced (just income) and full model specification, and a t-test of the difference between the full and reduced coefficients.

Female	OR	Model 1	Model 2	Model 3	Model 4
Income Quintile 2	Reduced	1.00	1.00	1.00	1.00
		(0.11)	(0.11)	(0.11)	(0.11)
	Full	1.06	1.09	1.09	1.10
		(0.12)	(0.12)	(0.12)	(0.13)
	Diff	0.95	0.92	0.92	0.91
		(0.03)	(0.04)+	(0.04)+	(0.05)+
Income Quintile 3	Reduced	0.89	0.89	0.89	0.89
		(0.10)	(0.10)	(0.10)	(0.10)
	Full	0.94	0.99	1.02	1.07
		(0.11)	(0.12)	(0.13)	(0.15)
	Diff	0.95	0.90	0.88	0.83
		(0.05)	(0.06)	(0.06)+	(0.07)*
Income Quintile 4	Reduced	0.85	0.85	0.85	0.85
		(0.09)	(0.09)	(0.09)	(0.09)
	Full	0.90	0.97	1.03	1.12
		(0.11)	(0.13)	(0.14)	(0.17)
	Diff	0.95	0.88	0.82	0.75
		(0.06)	(0.07)	(0.07)*	(0.08)**
Income Quintile 5	Reduced	0.64	0.64	0.64	0.64
		(0.07)***	(0.07)***	(0.07)***	(0.07)***
	Full	0.67	0.75	0.84	0.94
		(0.09)**	(0.11)*	(0.13)	(0.15)
	Diff	0.96	0.86	0.76	0.68
		(0.07)	(0.09)	(0.08)*	(0.09)**
N		6,830	6,830	6,830	6,830

Male	OR	Model 1	Model 2	Model 3	Model 4
Income Quintile 2	Reduced	1.22	1.22	1.22	1.23
		(0.14)+	(0.15)+	(0.15)+	(0.15)+
	Full	1.29	1.31	1.33	1.34
		(0.16)*	(0.17)*	(0.17)*	(0.17)*
	Diff	0.95	0.93	0.92	0.92
		(0.04)	(0.05)	(0.05)	(0.05)
Income Quintile 3	Reduced	1.10	1.10	1.10	1.10
		(0.13)	(0.13)	(0.13)	(0.13)
	Full	1.20	1.24	1.31	1.33
		(0.15)	(0.17)	(0.18)+	(0.20)+
	Diff	0.91	0.89	0.84	0.83
		(0.05)	(0.07)	(0.07)*	(0.08)*
Income Quintile 4	Reduced	0.88	0.88	0.87	0.88
		(0.10)	(0.10)	(0.10)	(0.10)
	Full	0.95	1.03	1.12	1.16
		(0.13)	(0.16)	(0.17)	(0.19)
	Diff	0.92	0.85	0.78	0.75
		(0.07)	(0.08)+	(0.08)*	(0.09)*
Income Quintile 5	Reduced	0.76	0.76	0.76	0.76
		(0.09)*	(0.09)*	(0.09)*	(0.09)*
	Full	0.80	0.92	1.06	1.14
		(0.11)	(0.15)	(0.18)	(0.21)
	Diff	0.95	0.82	0.71	0.67
		(0.08)	(0.09)+	(0.09)**	(0.09)**
N		6,969	6,969	6,969	6,969

Appendix L: Robustness check: Different numbers of income groups and income as a continuous variable

Different groupings of income

In the main analysis I do not find evidence for an association between income and child overweight, after controlling for demographic characteristics and parental education. Even the unadjusted relationship between income and child overweight and obesity is small. However it is possible that the lack of a relationship could be a statistical artefact caused by the use of income quintiles, rather than some other grouping of income.

The choice of using income quintiles is somewhat arbitrary. Quintiles divides the equivalised income distribution into 5 equal groups, with twenty percent of the equivalised income distribution within each group. It may be that the cut off value for membership into the low income group is too high, so that an association between very low income and child overweight is being masked because those on the lowest incomes are grouped together with those with higher incomes. Alternatively it may be the case that the threshold is too stringent, and those on reasonably low income are not being included in the low income category. To some extent this problem is addressed with the measure of poverty, as the premise of this measure is that it should capture those who have insufficient incomes. However, this matter is investigated further by dividing the equivalised income distribution into tertiles (equal groups of three) and septiles (equal groups of 7). This will allow me to investigate whether the results are being driven by the grouping of income. The results for income septiles are presented in table A.26 and for income tertiles in A.27.

Table A.26. The relationship between septiles of equivalised income and child overweight. Results presented as odds ratios.

Female	Model 1	Model 2	Model 3	Model 4	Model 5
Income Septile 1	Reference				
Income Septile 2	1.00	1.06	1.09	1.09	1.09
	(0.13)	(0.14)	(0.14)	(0.14)	(0.15)
Income Septile 3	1.08	1.15	1.21	1.22	1.25
	(0.14)	(0.15)	(0.17)	(0.17)	(0.18)
Income Septile 4	0.90	0.96	1.02	1.05	1.12
	(0.11)	(0.12)	(0.14)	(0.15)	(0.17)
Income Septile 5	0.83	0.90	0.98	1.03	1.11
	(0.11)	(0.13)	(0.15)	(0.15)	(0.18)
Income Septile 6	0.75	0.82	0.91	0.99	1.09
	(0.10)*	(0.12)	(0.14)	(0.15)	(0.18)
Income Septile 7	0.67	0.72	0.82	0.94	1.05
	(0.09)**	(0.11)*	(0.13)	(0.15)	(0.18)
	6,830	6,830	6,830	6,830	6,830
Males	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Income Septile 1	Reference				
Income Septile 2	0.93	0.98	0.99	1.00	0.99
	(0.11)	(0.13)	(0.13)	(0.14)	(0.13)
Income Septile 3	1.17	1.26	1.27	1.31	1.31
	(0.16)	(0.18)	(0.20)	(0.20)+	(0.21)+
Income Septile 4	0.98	1.08	1.11	1.17	1.18
	(0.13)	(0.17)	(0.20)	(0.21)	(0.22)
Income Septile 5	0.93	1.01	1.07	1.16	1.18
	(0.12)	(0.16)	(0.19)	(0.20)	(0.21)
Income Septile 6	0.72	0.78	0.84	0.94	0.97
	(0.09)**	(0.12)	(0.15)	(0.17)	(0.19)
Income Septile 7	0.65	0.68	0.78	0.90	0.95
	(0.09)**	(0.11)*	(0.14)	(0.16)	(0.19)
N	6,969	6,969	6,969	6,969	6,969

+ $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

Model 1 = Time averaged Income; Model 2 = Model1 + Ethnicity, longstanding illness/disability, region, parental age; Model 3 = Model 2 + Main respondent education; Model 4 = Model 3 + Partner respondent education; Model 5 = Model 4 + Highest social class of parents.
Estimates weighted using dovwt2 survey design & attrition weight

Table A.27. The relationship between teriles of equivalised income and child overweight. Results presented as odds ratios.

Female	Model 1	Model 2	Model 3	Model 4	Model 5
Income Tertile 1	Reference				
Income Tertile 2	0.94	0.98	1.03	1.05	1.10
	(0.07)	(0.08)	(0.09)	(0.09)	(0.10)
Income Tertile 3	0.73	0.78	0.86	0.94	1.01
	(0.06)***	(0.08)*	(0.10)	(0.11)	(0.13)
	6,830	6,830	6,830	6,830	6,830
Male	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Income Tertile 1	Reference				
Income Tertile 2	1.09	1.22	1.25	1.30	1.31
	(0.10)	(0.13)+	(0.14)*	(0.14)*	(0.15)*
Income Tertile 3	0.76	0.84	0.94	1.04	1.07
	(0.07)**	(0.10)	(0.11)	(0.13)	(0.14)
	6,969	6,969	6,969	6,969	6,969

+ $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

Model 1 = Time averaged Income; Model 2 = Model1 + Ethnicity, longstanding illness/disability, region, parental age; Model 3 = Model 2 + Main respondent education; Model 4 = Model 3 + Partner respondent education; Model 5 = Model 4 + Highest social class of parents.
Estimates weighted using dovwt2 survey design & attrition weight

The unadjusted relationship between income septiles and child overweight status suggests that female children who are in the lower income groups (septile 1 and 2) have statistically significantly higher odds of being classified as overweight than children who are in the higher income groups (septile 6 and 7). Females in septile 3 also have significantly higher odds of being overweight than children in income septile 6 and 7. After adjustment for demographic characteristics, the difference between income septile 1 and income septile 6 is no longer statistically significant, but the difference between the two lowest income septiles, and the highest income septile remains. The difference in the odds between income septile 3, which consists of families between the 28th and 42nd percentile of the income distribution, and income septiles 6 and 7 is statistically significant. After adjustment for education of the main and partner respondents, female sin income septile 3 have the highest odds of being classified as overweight, but the difference in the odds is not statistically significantly different from any other income septile. The difference in the odds between the high income septiles and low income septiles are greatly reduced. After further adjustment for social class, females in the lowest income septile have the lowest estimated odds of being overweight.

For males the unadjusted relationship is very similar to females, with children in the highest income septiles having statistically significantly lower odds of being overweight, than children in the lowest income septiles. Also the highest odds of being overweight are in income septile 3. After adjustment for demographic characteristics, the difference between income septile 1 and income septile 6 is no longer statistically significant, but the difference between the two lowest income septiles, and the highest income septile remains. The difference in the odds between income septile 3 and income septiles 6 and 7 is statistically significant. After adjustment for main and partner respondent's education, there are no significant differences in the odds between the highest and lowest income septiles. The difference in odds between income septile 1 and 3 is approaching significance, with children in income septile 3 having increased odds of being overweight compared to the lowest income septile. This persists after adjustment for social class.

As with the main analysis then, the results for income septiles show that for male children there is evidence for higher odds of being overweight for children just below the middle of the equivalised income distribution. Whereas for female children there is no evidence of differences in the odds after adjustment for parental education. For neither males nor females is there evidence that children in the lowest income groups have the highest of being overweight after controlling for demographic characteristics. As might be expected the results with income tertiles show a similar finding but due to the larger group sizes, and hence more power, the differences in the odds for males are statistically significant even after full adjustment of the models. The results with income tertiles show that males in the second income tertile, that is families between the 33rd and 66th percentile of the equivalised income distribution have significantly higher odds of being overweight, whereas for females there is no differences in the odds of being overweight after adjustment for parental education.

The relationship between standardised log familial income and child overweight

In the main body of the paper income is divided into quintiles, due to non-linearity's for the males. The case can be made, however, that the deviations form log-linearity are small and that it is still acceptable to include log-linear income as a continuous variable in the model. Therefore the results are presented here in table A.28 for the relationship between log-linear income and child overweight.

Table A.28. The relationship between the standardised log income and overweight. Results are presented as odds ratios.

Male	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Log Transitory income	0.94 (0.03)+					
Log Permanent income		0.88 (0.03)***	0.89 (0.04)*	0.93 (0.05)	0.98 (0.05)	0.99 (0.05)
	6,969	6,969	6,969	6,969	6,969	6,830
Female						
Log Transitory income	0.90 (0.03)**					
Log Permanent income		0.87 (0.03)***	0.88 (0.04)**	0.91 (0.04)*	0.95 (0.05)	0.99 (0.05)
	6,830	6,830	6,830	6,830	6,830	6,830

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 0 = Transitory Income; Model 1 = Permanent Income; Model 2 = Model1 + Ethnicity, longstanding illness/disability, region, parental age; Model 3 = Model 2 + Main respondent education; Model 4 = Model 3 + Partner respondent education; Model 5 = Model 4 + Highest social class of parents.

As shown in table A.28, and as presented in the main results section, there is a stronger relationship between income and child overweight when permanent measures rather than transitory measures of income are used. For transitory measures of income, a standard deviation (SD) increase in the log of weekly equivalised income leads to a 6% decrease in the odds of being classified as overweight for males and 10% decrease in the likelihood for females. Whereas for permanent income a SD increase in log income is associated with a 12% and 13% decrease in the likelihood of overweight for males and females respectively.

The relationship between income and child overweight persists after adjustment for demographics without much attenuation. However after adjustment for main parental respondent education, the strength of the relationship is reduced and the relationship is no longer significant for males. After further adjustment for partner respondent's education, there relationship is no longer significant for females.

Income as a non linear

A quadratic term was tested for females and males. The results from the analysis including quadratic terms can be seen in table A.29. In the unadjusted models, the quadratic term is only significant for males. The quadratic term remains statistically significant for males throughout model adjustments, but the main effect is no longer statistically significant after adjustment for parental education. For females the quadratic term approaches significance after adjustment for demographic characteristics, but neither the quadratic, nor the main effect of income are statistically significant after adjustment for partner respondents education. The significant quadratic term for males is negative, as it results in a reduction in the odds ratios. This suggests a concave relationship between income and overweight for male children.

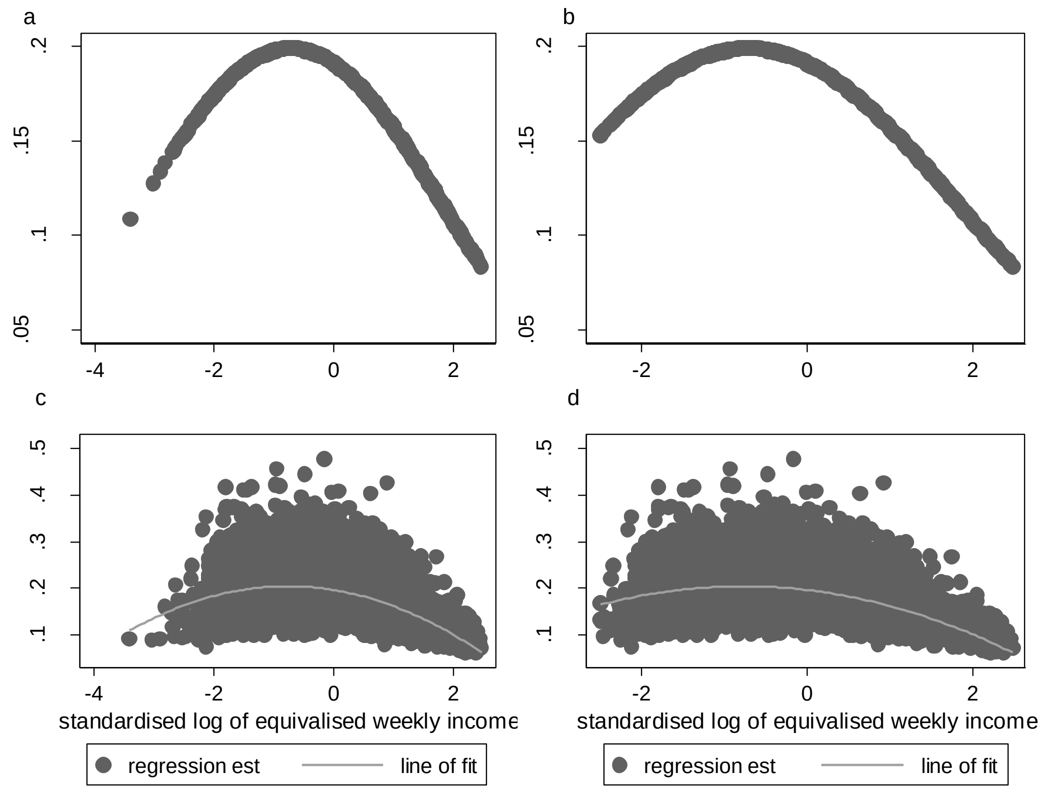
Table A.29. The relationship between the standardised log of equivalised income and overweight. Results are presented as odds ratios.

Male	Model 1	Model 2	Model 3	Model 4	Model 5
STD log income	0.87	0.88	0.93	0.97	0.99
	(0.03)***	(0.04)*	(0.05)	(0.05)	(0.06)
STD log income sq	0.91	0.88	0.90	0.91	0.91
	(0.04)*	(0.04)**	(0.04)*	(0.04)*	(0.04)*
	6,969	6,969	6,969	6,969	6,969
Female	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
STD log income	0.86	0.87	0.91	0.95	0.98
	(0.03)***	(0.04)**	(0.04)*	(0.05)	(0.05)
STD log income sq	0.96	0.94	0.94	0.96	0.96
	(0.03)	(0.03)+	(0.03)+	(0.03)	(0.03)
	6,830	6,830	6,830	6,830	6,830

Model 1 = Time averaged Income; Model 2 = Model1 + Ethnicity, longstanding illness/disability, region, parental age; Model 3 = Model 2 + Main respondent education; Model 4 = Model 3 + Partner respondent education; Model 5 = Model 4 + Highest social class of parents.
Estimates weighted using dovwt2 survey design & attrition weight

The relationship between the fitted quadratic term and income is shown in figure A.7. As can be seen in figure A.7, the probability of overweight increases to a point in the income distribution (approximately -1 SD of log equivalised income) before decreasing rapidly with increasing income thereafter. The relationship between income and child overweight for males was considered across the full range of incomes (as shown in in figure A.7 a & c), as well as only for those above -2.5 SD on the log of income (figure A.7 b & d). Including the values below -2.5SD appears to be exaggerating the curvature in the relationship between income and child overweight. There are very few data points below -2.5 but they are having a large influence on the estimates.

Figure A.7. The non-linear relationship between income and overweight for male children.



Figures a and b, show the unadjusted relationship between income and overweight, figures c and d show the relationship after full adjustment of the model. The two figures on the right hand side, figures b and d only show the distribution of equivalised log income from -2.5SD.

There is evidence that families at approximately -1SD on the log of equivalised weekly income, that is approximately 5.2 on the logarithmic scale, or £181 on the equivalised income scale, have the highest probability of having an overweight male child. Figure A.7d suggests that there doesn't seem to be much difference in the probability for male children with families between -2SD and the average income (0), after this point there is a sharp decrease in probability with increasing incomes.

Appendix M: Robustness check: Growing up in Scotland

The Growing up in Scotland survey (GuS), covers a similar time period and population to the MCS. The longitudinal sample includes 3477 children aged 5/6 from the birth cohort sample, which began in 2004/2005. The cross sectional survey design weight from sweep 6 (2010/2011) was applied throughout the analysis. Overweight was defined by the IOTF criteria. The variables included in the analysis were very similar to those included in the MCS. Time averaged income was measured using annual equivalised net income averaged across six sweeps of data collection. Transitory income was measured using responses to income at sweep 6 only. Demographic characteristics included parental ethnicity (child's not reported), parental age and longstanding illness/disability. Parental education was measured using highest NVQ equivalent qualification for each parent/carer. Social class was measured by the household NS-SEC. The same modelling strategy as the main analysis was applied.

As shown in table A.30, there was no statistically significant differences in the odds of overweight for males or females in any specification of the models, this could be due to the smaller sample size. The patterning of the odds for male and children appeared to be quite similar in the GuS cohort, with higher odds of being overweight for those in quintile 2 or 3. Therefore the analysis was also run on the whole sample, table A.31, to boost the power so that differences in the odds could be distinguished.

Even after combining the males and females, there are no statistically significant differences in the odds of being overweight by income quintile in the unadjusted models. After adjustment for demographic characteristics, the odds of being overweight for those in income quintile 3 and the lowest income quintile are approaching significance, with higher odds for those in income quintile 3. After consecutive adjustments of the model, all income quintiles have higher odds of being overweight than children in the lowest income quintiles. After controlling for social class the difference between income quintile 1 and 3 is statistically significant, and the difference between income quintile 5 and 1 is approaching significance, with higher odds of overweight in income quintile 5.

Table A.30. The relationship between income quintiles and child overweight in the Growing up in Scotland Survey for males and females separately. Results are presented as odds ratios.

Male	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Tran Income Q2	1.29					
	(0.27)					
Tran Income Q3	1.25					
	(0.27)					
Tran Income Q4	1.00					
	(0.22)					
Tran Income Q5	0.69					
	(0.17)					
TA Income Q2		0.98	1.20	1.19	1.17	1.22
		(0.23)	(0.31)	(0.31)	(0.31)	(0.32)
TA Income Q3		1.04	1.42	1.37	1.34	1.50
		(0.24)	(0.36)	(0.37)	(0.36)	(0.42)
TA Income Q4		0.80	1.13	1.13	1.17	1.31
		(0.18)	(0.29)	(0.32)	(0.33)	(0.39)
TA Income Q5		0.76	1.09	1.17	1.28	1.44
		(0.18)	(0.30)	(0.36)	(0.40)	(0.48)
N	1,770	1,770	1,770	1,769	1,769	1,769
Female						
Tran Income Q2	1.28					
	(0.24)					
Tran Income Q3	0.77					
	(0.16)					
Tran Income Q4	0.82					
	(0.16)					
Tran Income Q5	0.84					
	(0.16)					
TA *Income Q2*		1.05	1.25	1.29	1.25	1.31
		(0.22)	(0.28)	(0.30)	(0.29)	(0.32)
TA *Income Q3*		1.07	1.40	1.46	1.41	1.53
		(0.21)	(0.31)	(0.35)	(0.34)	(0.38)+
TA *Income Q4*		0.89	1.20	1.34	1.31	1.44
		(0.18)	(0.28)	(0.34)	(0.33)	(0.39)
TA *Income Q5*		0.80	1.11	1.31	1.42	1.57
		(0.16)	(0.26)	(0.34)	(0.37)	(0.44)
N	1,707	1,707	1,707	1,707	1,707	1,707

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 0 = Transitory Income; Model 1 = Time averaged income; Model 2 = Model1 + parental ethnicity, longstanding illness/disability, parental age; Model 3 = Model 2 + Main parental respondent education; Model 4 = Model 3 + Partner respondent education; Model 5 = Model 4 + Household social class of parents.
Estimates weighted using the sweep 6 cross sectional analysis weight (DfWTbrth).

Table A.31. The relationship between income quintiles and child overweight in the Growing up in Scotland Survey for all children. Results are presented as odds ratios.

ALL	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Tran Income Q2	1.28 (0.18)+					
Tran Income Q3	0.97 (0.14)					
Tran Income Q4	0.90 (0.13)					
Tran Income Q5	0.80 (0.12)					
TA *Income Q2*		1.00 (0.16)	1.18 (0.20)	1.20 (0.21)	1.18 (0.20)	1.24 (0.21)
TA *Income Q3*		1.06 (0.16)	1.36 (0.23)+	1.37 (0.24)+	1.34 (0.24)+	1.47 (0.27)*
TA *Income Q4*		0.85 (0.13)	1.13 (0.20)	1.21 (0.23)	1.22 (0.23)	1.35 (0.27)
TA *Income Q5*		0.79 (0.12)	1.08 (0.19)	1.23 (0.24)	1.34 (0.27)	1.49 (0.31)+
N	3,477	3,477	3,477	3,477	3,477	3,477

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 0 = Transitory Income; Model 1 = Time averaged income; Model 2 = Model1 + parental ethnicity, longstanding illness/disability, parental age; Model 3 = Model 2 + Main parental respondent education; Model 4 = Model 3 + Partner respondent education; Model 5 = Model 4 + Household social class of parents.
Estimates weighted using the sweep 6 cross sectional analysis weight (DfWTbrth).

Appendix N: Robustness check: Multinomial Logit

Multinomial logistic regressions are usually used when making comparisons between unrelated groups. The dependent variable in question is assumed to be nominal with more than two categories. This means that using this method assumes that the groups are not ordered in any meaningful way. Examples of when the multinomial logit are appropriate include comparisons between countries, or comparing children from three different types of schools. In this application the assumption that the groups are unrelated is violated because there is ordering of the dependent variable. Children who are underweight[46] have lower BMI's, given their age and sex, than children who are healthy weight. Children who are overweight have higher BMI's, given their age and sex, than children who are healthy weight.

However, the multinomial logit was deemed more suitable than the ordered logit. The ordered logit model would assume a linear relationship between income and child weight classification (under/ healthy/ normal) and would assume proportional odds[47]. Therefore the assumption would have to be made that the size and shape of the relationship between income and child overweight did not vary depending upon weight status classification. However, it is expected that low income will be associated with an increased risk of underweight as well as overweight, therefore the relationship is nonlinear, also it is expected that income will have more of an effect at the extreme ends of the BMI distribution (underweight and overweight) than for the healthy weight children (who make up the majority of the sample).

There are methods for relaxing some of the assumptions of the ordered logit through the use of generalised ordered logits. However, the results from these models can be difficult to interpret, due to the amount of parameters estimated. Therefore, whilst it is not the most efficient use of the information available, the underweight, healthy weight and overweight groups are compared as though they were unrelated and unordered groupings of children.

[46] Underweight was defined by Cole et al (2007), and was created in the same way as the IOTF criteria for overweight and is base don't he same sample of children.
[47] A Brant test confirmed that this assumption would be violated in an ordered logit.

The multinomial logit calculates the relative risk of being in a specific category, compared to a reference category. In this case "healthy weight" will be the reference category and both overweight and underweight groups will be compared to the healthy weight group. The assumption that these groups are unrelated means that when comparing the overweight to the healthy weight group, it is effectively the same as running the logistic regression models having deleted all those who were underweight. Therefore I am able to look at the "risk" of being in the overweight category, compared to the healthy weight category. The equation for the multinomial logit is shown below:

Equation A.6. Multinomial logistic regression

$$\text{Ln}\left(\frac{P\ underweight}{P\ Healthy\ weight}\right) = \alpha + \beta 1x1 + \beta 1x2 \ldots \ldots \beta 1Xn$$

$$\text{Ln}\left(\frac{P\ overweight}{P\ Healthy\ weight}\right) = \alpha + \beta 2x1 + \beta 2x2 \ldots \ldots \beta 2Xn$$

The multinomial regression relies on the assumption of independence of irrelevant alternatives. This assumption states that the relative odds between any two outcomes are independent of the number and nature of other outcomes being simultaneously considered. This means that none of the categories can act as substitutes for the others. This assumption holds in the current application, as a child is either classified as underweight, healthy weight or overweight. The categories are distinct and being classified as overweight eliminates the possibility of being classified as underweight or healthy weight. The results from the multinomial regression are shown in table A.32. As can be seen in table A.32, particularly for males, children in the lowest income quintiles are the most likely to be classified as overweight.

The focus of this analysis is on overweight, therefore the coefficients of most interest are those which compare the overweight group to the healthy weight group. For females there appears to be a linear relationship between income and the relative risk of being overweight compared to healthy weight, with higher incomes resulting in a lower risk. However after controlling for main and partner respondent's education, there are no statistically significant differences in the relative risks. The unadjusted relationship for males suggests that the relative risk of being overweight, compared to healthy weight, is highest for those in income quintile 2 and that there are significant differences between income quintile 1 and 5 and income quintile 2 and 5. After adjustment for demographic characteristics the difference between income quintile 1 and 2 is statistically significant, with male children in income quintile 2 at a higher risk of being overweight than children in income quintile 1. This difference persists after adjustment for parental education and social class.

The results using the multinomial logistic regression are very similar to those using the logistic regression models in the main analysis.

Table A.32. Multinomial logit considering the relationship between income and child overweight & obesity, and child underweight. Results presented as relative risk ratios.

Female	OR	Model 1	Model 2	Model 3	Model 4	Model 5
Underweight	Income Q2	0.82	0.93	0.94	0.91	0.96
		(0.14)	(0.17)	(0.17)	(0.17)	(0.17)
	Income Q3	0.84	1.09	1.09	1.04	1.09
		(0.15)	(0.21)	(0.21)	(0.21)	(0.22)
	Income Q4	0.71	0.90	0.88	0.82	0.81
		(0.13)+	(0.20)	(0.20)	(0.19)	(0.20)
	Income Q5	0.93	1.14	1.06	1.00	0.91
		(0.17)	(0.24)	(0.26)	(0.26)	(0.25)
Healthy	Reference group					
Overweight/Obese	Income Q 2	0.98	1.05	1.08	1.08	1.10
		(0.11)	(0.12)	(0.12)	(0.12)	(0.13)
	Income Q 3	0.88	0.94	1.00	1.02	1.08
		(0.09)	(0.10)	(0.12)	(0.12)	(0.14)
	Income Q 4	0.82	0.89	0.96	1.01	1.11
		(0.09)+	(0.11)	(0.13)	(0.13)	(0.16)
	Income Q 5	0.63	0.68	0.75	0.85	0.94
		(0.07)***	(0.09)**	(0.11)*	(0.12)	(0.14)
	N	6,830	6,830	6,830	6,830	6,830
Male	**OR**	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Underweight	Income Q 2	0.77	0.83	0.84	0.84	0.85
		(0.12)+	(0.13)	(0.14)	(0.14)	(0.15)
	Income Q 3	0.54	0.62	0.64	0.62	0.64
		(0.10)***	(0.12)*	(0.14)*	(0.13)*	(0.15)+
	Income Q 4	0.74	0.84	0.89	0.86	0.93
		(0.13)+	(0.18)	(0.21)	(0.22)	(0.24)
	Income Q 5	0.58	0.62	0.69	0.67	0.77
		(0.11)**	(0.14)*	(0.18)	(0.18)	(0.21)
Healthy	Reference Group					
Overweight_Obese	Income Q 2	1.19	1.27	1.29	1.31	1.32
		(0.14)	(0.15)*	(0.17)+	(0.17)*	(0.17)*
	Income Q 3	1.04	1.16	1.20	1.26	1.29
		(0.12)	(0.16)	(0.19)	(0.19)	(0.21)
	Income Q 4	0.85	0.94	1.02	1.10	1.15
		(0.09)	(0.13)	(0.16)	(0.17)	(0.19)
	Income Q5	0.72	0.77	0.89	1.03	1.11
		(0.09)**	(0.11)+	(0.13)	(0.16)	(0.19)
N		6,969	6,969	6,969	6,969	6,969

+ $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

Model 1 = Time averaged Income; Model 2 = Model1 + Ethnicity, longstanding illness/disability, region, parental age; Model 3 = Model 2 + Main respondent education; Model 4 = Model 3 + Partner respondent education; Model 5 = Model 4 + Highest social class of parents.
Estimates weighted using dovwt2 survey design & attrition weight

Appendix O: Coding the ISCED into the MCS

ISCED is an instrument for compiling internationally comparable education statistics. The ISCED 97 covers information pertaining to the level of education and the field of education. The level of education broadly reflects the degree of complexity of the content of the educational programme. The education levels correspond to the knowledge, skills and capabilities required of participants for completion. Within these levels programmes are further classified by their destination i.e whether they allow access to further education or to the labour market, and by the orientation of the programme, whether it is general, pre-vocational, or vocational. More detailed information about the ISCED classification can be found here.

There are seven levels of education in ISCED 97.

- Level 0: **Pre-primary education**
- Level 1: **Primary education**
- Level 2: **Lower secondary education**
- Level 3: **Upper secondary education**
- Level 4: **Post-secondary non-tertiary education**
- Level 5: **Tertiary education (first stage)**
- Level 6: **Tertiary education (second stage)**

Highest academic and vocational qualification to date of the mother and father was coded in the data by using information from every sweep of data collection, in which parents reported any new qualifications. Where qualification information was missing, the highest qualification reported at any sweep was utilised as the highest qualification.

Academic qualification categories reported in the MCS

- Higher Degree
- first Degree
- Diplomas in Higher Education
- A, AS, S Levels
- level, GCSE grades A-C
- GCSE grades D-G
- Other academic qualifications (including overseas)
- None of these qualifications.

Vocational qualifications categories reported in the MCS

- professional qualifications at degree level
- nursing, other medical qualifications
- NVQ, SVQ, GSVQ level 3
- NVQ, SVQ, GSVQ level 2
- NVQ, SVQ, GSVQ level 1
- other qualifications (including overseas)
- None of these qualifications.

Table A.33 shows how the OECD recommends coding the ISCED within the UK educational framework. Taken from OECD (1999). Classifying Educational Programmes. Manual for ISCED-97 Implementation in OECD Countries, http://www.oecd.org/dataoecd/41/42/1841854.pdf

ISCED	Programme orientation		
Level	A	B	C
0	Nursery schools, Playgroups, Reception classes; left school before age 11 (no qualification)		
1	Primary school, Adult literacy and numeracy courses; left school age 11–14 (no qualification)		
2	Left school after age 14 without qualification		
3	GCE A/AS Level, Higher Grade, CSYS (all general), GNVQ/GSVQ Advanced, NVQ Level 3 (vocational)		GCSEs, Standard Grade, GNVQ/GSVQ Foundation & Intermediate, NVQ Levels 1 (partly pre-vocational) & 2 (vocational)
4	HE Access Courses		
5	medium: BA long: MA, PGCE, PGDE	NVQ Levels 4 & 5, HNC, HND, CertHE, DipHE	
6	Doctorate/PhD		

The OECD have provided guidelines on classifying education programmes within the UK educational framework, these are shown in table A.33. These guidelines were used to code information on education from the parents of the MCS children into the ISCED classification of levels of education.

Where parents finished school before the age of 11 and they reported having 'none of these' to both the academic and vocational qualifications, they were coded with ISCED 0. For those who reported leaving school between the ages of 11-14 and reported 'none of these' to both the academic and vocational qualification they were coded with ISCED 1. Those who left school after the age of 14 but reported 'none of these' to both the academic and vocational qualifications were coded as ISCED 2. Those who reported obtaining GCSE qualifications (either A-C or D-G), or NVQ, SVQ, GNVQ at levels 1 or 2 were coded as ISCED 3C. Those who reported obtaining A, AS or S level qualifications, or NVQ, SVQ or GNVQ level 3 were coded as ISCED 3A.Those who reported having diplomas in higher education, professional qualifications at degree level, nursing and other medical qualifications were coded as ISCED 5B.

There was a problem coding ISCED 5A and ISCED 6. ISCED 5a should contain people with first degrees, PGCE qualifications and masters level qualifications. ISCED 6 should contain only those with a Doctorate. However any form of postgraduate qualification (including masters and PGCE) was coded in the 'higher degree' category for parents of the MCS children, and the actual type of qualification cannot be obtained. All those with first degrees were coded as ISCED 5A, and all those with higher degrees were coded as ISCED 6. This means that ISCED 6 numbers are inflated, and if these data were to be used for international comparison, this would need to be taken into account either by exclusion of the ISCED 6 category or, preferably, through combining ISCED 5A and ISCED 6 levels. As it stands the ISCED 6 category coded for the MCS parents represents all postgraduate education and ISCED 5A represents undergraduate level education.

Table A.34. Observed number of mothers and father's in each ISCED level in the subsample of natural mothers and fathers in the MCS

ISCED levels	Mothers (n)	Fathers (n)
0	11	4
1	91	57
2	829	775
3c	3013	2723
3a	1458	1364
5b	1624	1432
5a	1847	1503
6	636	788

Table A.34 shows the observed numbers of parents classified in each ISCED level. There were very small numbers of parents classified in ISCED levels 0 and 1. Therefore ISCED levels 0, 1 and 2 were combined into a 'no qualifications' category which consists of people without any qualifications.

Appendix P: Different measures of parental education

In this appendix I use different measures of parental education to investigate whether the finding that father's education has a stronger association with child overweight than mother's education is driven by the use of the ISCED measure. In table A.34 I present the results whereby mother's and father's education are measured by the highest NVQ equivalent qualification achieved. The NVQ qualifications, like the ISCED consider both vocational and academic qualifications. As can be seen in table A.35, the results follow much the same pattern as the results presented in the main analysis with father's education having a stronger association than mother's education, and the association persisting after full adjustment of the models.

In table A.36 I consider parental education as measured by the highest academic qualification obtained. In this specification the odds of a child being overweight vary for both mother's and father's education. Whilst it is not immediately obvious from looking at the tables, the results are very similar when using just highest academic qualification obtained. The reference category has been set at a low level of parental education (GCSE grades D-G), as it has in all other tables for consistency purposes, however the pattern is much clearer when the higher degree category is set at the reference category.

Table A.35. Sequential logistic regression models showing mother's and father's education measured by the highest NVQ or equivalent obtained. Results shown in odds ratios.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Mother NVQ 1			Reference Category				
Mother NVQ 2	0.92 (0.11)	0.95 (0.12)	0.94 (0.12)	0.92 (0.11)	0.94 (0.12)	0.97 (0.12)	1.00 (0.13)
Mother NVQ 3	0.81 (0.11)	0.87 (0.12)	0.87 (0.12)	0.82 (0.11)	0.86 (0.12)	0.88 (0.12)	0.89 (0.12)
Mother NVQ 4	0.73 (0.09)*	0.85 (0.11)	0.84 (0.11)	0.77 (0.10)+	0.84 (0.11)	0.84 (0.11)	0.90 (0.12)
Mother NVQ 5	0.61 (0.10)**	0.75 (0.13)+	0.74 (0.13)+	0.64 (0.11)**	0.71 (0.12)+	0.70 (0.12)*	0.78 (0.13)
Mother Overseas	0.74 (0.14)	0.76 (0.15)	0.76 (0.15)	0.69 (0.14)+	0.69 (0.14)+	0.70 (0.14)+	0.79 (0.15)
Mother None	0.97 (0.15)	0.92 (0.14)	0.91 (0.14)	0.87 (0.14)	0.87 (0.14)	0.86 (0.14)	0.93 (0.15)
Father NVQ 1				Reference Category			
Father NVQ 2		0.85 (0.11)	0.85 (0.11)	0.81 (0.10)	0.83 (0.10)	0.84 (0.11)	0.87 (0.12)
Father NVQ 3		0.82 (0.12)	0.82 (0.12)	0.80 (0.12)	0.81 (0.12)	0.85 (0.13)	0.88 (0.14)
Father NVQ 4		0.72 (0.09)*	0.71 (0.09)**	0.67 (0.09)**	0.71 (0.09)**	0.76 (0.10)*	0.79 (0.11)+
Father NVQ 5		0.58 (0.09)***	0.58 (0.09)***	0.53 (0.09)***	0.59 (0.10)**	0.63 (0.10)**	0.70 (0.12)*
Father Overseas		0.86 (0.14)	0.86 (0.14)	0.82 (0.14)	0.82 (0.14)	0.82 (0.14)	0.82 (0.15)
Father None		1.03 (0.16)	1.01 (0.15)	0.98 (0.14)	0.97 (0.14)	0.96 (0.14)	0.99 (0.15)
Father Missing		1.24 (0.26)	1.22 (0.25)	1.24 (0.28)	1.23 (0.28)	1.34 (0.42)	1.17 (0.36)
N	9,703	9,703	9,703	9,703	9,703	9,703	9,835

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 0 = Mothers education,
Model 1 = Model 0 + Fathers education,
Model 2 = Model 1 + Child's gender
Model 3 = Model 2 + child's ethnicity, whether or not mother, father or child have longstanding illness or disability, mothers age, fathers age, region of residence, and number of children in the household
Model 4 = Model 3 + Mother's social class & Father's social class
Model 5 = Model 4 + equivalised time averaged Income
Model 6 = Model 5 + Mothers BMI & Fathers BMI,

Parameter tests of the overall association between mother's education and child overweight

Model 0 F(6,383) =3.30, P=0.0035

Model 1 F(6,383) =0.93, P=0.4704

Parameter tests of the overall association between father's education and child overweight

Model 1 F(7,382)=3.72, P=0.0007
Model 2 F(7,382)=3.93, P=0.0004
Model 3 F(7,382)=4.55, P=0.0001
Model 4 F(7,382)=2.64, P=0.0111
Model 5 F(7,382)=2.22, P=0.0319
Model 6 F(7,382)=1.34, P= 0.2321

Table A.36. Sequential logistic regression models showing the relationship between mother's and father's education measured by the highest academic qualification obtained. Results shown in odds ratios.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Mother Higher deg	0.61	0.76	0.75	0.63	0.69	0.65	0.73
	(0.10)**	(0.13)	(0.13)+	(0.11)**	(0.12)*	(0.12)*	(0.13)+
Mother First deg	0.58	0.70	0.69	0.62	0.66	0.63	0.72
	(0.07)***	(0.09)**	(0.09)**	(0.08)***	(0.09)**	(0.09)***	(0.10)*
Mother Diploma in HE	0.79	0.87	0.88	0.79	0.82	0.81	0.86
	(0.11)+	(0.12)	(0.12)	(0.11)+	(0.11)	(0.12)	(0.12)
Mother A/AS/S levels	0.71	0.78	0.78	0.71	0.73	0.73	0.79
	(0.10)*	(0.11)+	(0.11)+	(0.10)*	(0.10)*	(0.10)*	(0.12)
Mother GCSE A-C	0.89	0.92	0.93	0.88	0.90	0.91	0.97
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.11)
Mother GCSE D-G	Reference Category						
Mother Other	0.83	0.87	0.87	0.76	0.76	0.76	0.84
	(0.19)	(0.20)	(0.20)	(0.17)	(0.17)	(0.18)	(0.20)
Mother none	0.91	0.85	0.85	0.81	0.80	0.80	0.87
	(0.12)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.12)
Mother Missing	1.18	0.93	0.91	0.76	0.79	0.77	0.93
	(0.26)	(0.21)	(0.20)	(0.20)	(0.21)	(0.20)	(0.24)
Father Higher Deg		0.69	0.69	0.63	0.68	0.70	0.82
		(0.11)*	(0.11)*	(0.11)**	(0.12)*	(0.12)*	(0.15)
Father First Deg		0.80	0.81	0.73	0.78	0.81	0.90
		(0.11)+	(0.11)	(0.10)*	(0.11)+	(0.12)	(0.13)
Father Diploma in HE		0.91	0.90	0.85	0.88	0.91	0.98
		(0.13)	(0.13)	(0.12)	(0.13)	(0.13)	(0.15)
Father A/AS/S levels		0.82	0.81	0.77	0.80	0.82	0.90

		(0.13)	(0.13)	(0.13)	(0.14)	(0.14)	(0.16)
Father GCSE A-C		0.98	0.98	0.94	0.96	0.97	1.03
		(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.14)
Father GCSE D-G		Reference Category					
Father Other		0.91	0.89	0.83	0.84	0.84	0.86
		(0.20)	(0.20)	(0.19)	(0.19)	(0.19)	(0.21)
Father none		1.21	1.19	1.17	1.16	1.14	1.14
		(0.16)	(0.16)	(0.15)	(0.15)	(0.15)	(0.15)
Father missing		1.41	1.42	1.37	1.38	1.40	1.35
		(0.21)*	(0.21)*	(0.21)*	(0.21)*	(0.23)*	(0.25)
	9,705	9,705	9,705	9,705	9,705	9,705	9,705

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 0 = Mothers education,
Model 1 = Model 0 + Fathers education,
Model 2 = Model 1 + Child's gender
Model 3 = Model 2 + child's ethnicity, whether or not mother, father or child have longstanding illness or disability, mothers age, fathers age, region of residence, and number of children in the household
Model 4 = Model 3 + Mother's social class & Father's social class
Model 5 = Model 4 + equivalised time averaged Income
Model 6 = Model 5 + Mothers BMI & Fathers BMI

<u>Parameter tests of the overall association between mother's education and child overweight</u>
Model 0 F(8,381)=5.17, p=0.0000
Model 1 F(8, 381)=1.63, P=0.1148
<u>Parameter tests of the overall association between father's education and child overweight</u>
Model 1 F(8, 381)=3.81, P=0.0003
Model 2 F(8,381)=3.79, P=0.0003
Model 3 F(8, 381)=4.39, P=0.0000
Model 4 F(8, 381)=2.98, P=0.0030
Model 5 F(8, 381)=2.65, P=0.0078
Model 6 F(8, 381)=1.26,P=0.2636

Appendix Q: Robustness check: KHB

Karlson Holm & Breen (2012) suggest that directly comparing coefficients between nested logit or probit models is erroneous. The difference in the magnitude of coefficients between nested models may not be the result of confounding or mediation, but may be due to rescaling of the model. They suggest that if the added covariate ‘z’ is correlated to the outcome variable ‘y’, it will change the magnitude of the coefficients on the variable of interest ‘x’ even if the newly added covariate and the variable of interest are not correlated. This is because the ‘z’ covariate will reduce the variance in the error term.

They introduce a novel solution which decomposes the change in the coefficients into a rescaling component and a confounding component, as well as creating a method for formally testing the degree of confounding. The method is to compare the full model with a reduced model that substitutes some z variables with the residuals of the z variables from a regression of the z variables on the variable of interest ‘x’, please refer to Karlson, Holm and Breen (2012) for full explanation and details.

In the table below, table A.37, the reduced model in model 0 refers to the model with only mother’s education in, and the full model refers to the model with father’s education included. From model 1 onwards the reduced model is the model with mothers and fathers education in. In model 1 the full model includes sex and demographic characteristics parental longstanding illness/disability, parental age, region of residence, whether child has longstanding illness/disability, number of children in household, and parental report of child’s ethnicity). In model 2 the full model includes everything in model 1 and equivalised weekly income. In Model 3 the full model includes everything in model 2 and mother’s and father’s social class. In model 4 the full model includes everything in model 3 and mothers and fathers BMI. The difference test formally tests the hypothesis that the effect of confounding, net of rescaling, is zero.

Table A.37 shows the coefficients of the full model specification (including the variables listed below), the reduced model specification (model 0 just mother's education in model, model 1-4 just mothers and fathers education in model) and a t-test of the difference between the full and reduced coefficients.

	OR	Model 0	Model 1	Model 2	Model 3	Model 4
Mother No qualifications	*Reference Category*					
Mother ISCED 3C	Reduced	1.01 (0.11)	1.14 (0.13)	1.14 (0.13)	1.14 (0.13)	1.15 (0.13)
	Full	1.14 (0.13)	1.14 (0.14)	1.19 (0.15)	1.20 (0.15)	1.14 (0.15)
	Diff	0.89 (0.03)**	1.00 (0.07)	0.96 (0.08)	0.95 (0.08)	1.01 (0.14)
Mother ISCED 3A	Reduced	0.86 (0.10)	1.02 (0.13)	1.02 (0.12)	1.02 (0.12)	1.01 (0.13)
	Full	1.01 (0.13)	1.00 (0.13)	1.05 (0.14)	1.06 (0.14)	0.99 (0.14)
	Diff	0.85 (0.04)***	1.02 (0.07)	0.97 (0.08)	0.95 (0.08)	1.03 (0.14)
Mother ISCED 5B	Reduced	0.91 (0.11)	1.11 (0.14)	1.11 (0.14)	1.11 (0.14)	1.12 (0.14)
	Full	1.11 (0.14)	1.05 (0.14)	1.09 (0.15)	1.12 (0.16)	1.05 (0.16)
	Diff	0.82 (0.04)***	1.06 (0.08)	1.02 (0.09)	0.99 (0.09)	1.06 (0.15)
Mother ISCED 5A	Reduced	0.63 (0.08)***	0.84 (0.11)	0.84 (0.11)	0.84 (0.11)	0.83 (0.11)
	Full	0.84 (0.11)	0.78 (0.11)+	0.80 (0.12)	0.83 (0.13)	0.84 (0.13)
	Diff	0.75 (0.05)***	1.07 (0.08)	1.05 (0.09)	1.01 (0.09)	0.98 (0.14)
Mother ISCED 6	Reduced	0.66 (0.10)**	0.91 (0.14)	0.91 (0.14)	0.91 (0.14)	0.91 (0.14)
	Full	0.92 (0.14)	0.81 (0.13)	0.82 (0.14)	0.86 (0.15)	0.86 (0.15)
	Diff	0.72 (0.05)***	1.13 (0.08)+	1.11 (0.11)	1.06 (0.10)	1.05 (0.16)
Mother Missing	Reduced	0.99 (0.16)	0.96 (0.15)	0.96 (0.15)	0.96 (0.15)	0.97 (0.15)
	Full	0.97 (0.15)	0.91 (0.15)	0.93 (0.16)	0.94 (0.16)	0.96 (0.17)
	Diff	1.02 (0.04)	1.06 (0.08)	1.03 (0.08)	1.03 (0.08)	1.01 (0.14)
Father No qualifications	*Reference category*					
Father ISCED 3C	Reduced		0.81 (0.10)+	0.81 (0.10)+	0.81 (0.10)+	0.81 (0.10)+
	Full		0.82 (0.10)	0.85 (0.11)	0.85 (0.11)	0.84 (0.11)
	Diff		0.99 (0.06)	0.96 (0.07)	0.95 (0.07)	0.97 (0.13)
Father ISCED 3A	Reduced		0.73 (0.10)*	0.73 (0.10)*	0.73 (0.10)*	0.73 (0.10)*
	Full		0.74	0.79	0.80	0.80

			(0.10)*	(0.11)+	(0.11)+	(0.11)
	Diff		0.99	0.93	0.92	0.91
			(0.06)	(0.07)	(0.07)	(0.12)
Father ISCED 5B	Reduced		0.74	0.74	0.74	0.74
			(0.10)*	(0.10)*	(0.10)*	(0.10)*
	Full		0.73	0.79	0.81	0.81
			(0.10)*	(0.11)+	(0.11)	(0.12)
	Diff		1.02	0.94	0.92	0.92
			(0.07)	(0.07)	(0.07)	(0.13)
Father ISCED 5A	Reduced		0.64	0.64	0.64	0.64
			(0.09)**	(0.09)**	(0.09)**	(0.09)**
	Full		0.62	0.69	0.71	0.74
			(0.09)***	(0.10)*	(0.11)*	(0.11)*
	Diff		1.04	0.93	0.90	0.87
			(0.07)	(0.08)	(0.08)	(0.12)
Father ISCED 6	Reduced		0.55	0.55	0.55	0.56
			(0.10)***	(0.10)***	(0.10)***	(0.10)***
	Full		0.53	0.59	0.62	0.69
			(0.09)***	(0.10)**	(0.11)**	(0.13)*
	Diff		1.04	0.93	0.89	0.81
			(0.07)	(0.08)	(0.08)	(0.12)
Father Missing	Reduced		1.14	1.15	1.14	1.20
			(0.15)	(0.15)	(0.15)	(0.16)
	Full		1.15	1.17	1.18	1.12
			(0.15)	(0.16)	(0.16)	(0.18)
	Diff		1.00	0.98	0.97	1.07
			(0.06)	(0.08)	(0.08)	(0.16)
N		9,705	9,705	9,705	9,705	9,705

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 0 Reduced Mother's education: Full Father's education
Model 1: Reduced Mother's education, Father's education: Full gender + demographics
Model 2: Reduced Mother's education, Father's education: Full gender + demographics +Class
Model 3: Reduced Mother's education, Father's education: Full gender + demographics +Class Income
Model 4: Reduced Mother's education, Father's education: Full gender + demographics +Class Income + parental BMI

Appendix R: Full list of coefficients included in analysis for chapter 6

Table A.38. Full list of coefficients regarding the influence of parental education on child obesity.

	Model 0a	Model 0b	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Mother ISCED 3C	1.01 (0.11)		1.14 (0.13)	1.15 (0.13)	1.14 (0.14)	1.19 (0.15)	1.20 (0.15)	1.14 (0.14)
Mother ISCED 3A	0.86 (0.10)		1.01 (0.12)	1.03 (0.12)	1.00 (0.13)	1.05 (0.14)	1.06 (0.14)	0.99 (0.14)
Mother ISCED 5B	0.91 (0.10)		1.11 (0.14)	1.13 (0.14)	1.05 (0.13)	1.09 (0.15)	1.12 (0.16)	1.05 (0.15)
Mother ISCED 5A	0.63 (0.08)***		0.84 (0.11)	0.84 (0.11)	0.78 (0.11)+	0.80 (0.12)	0.83 (0.13)	0.84 (0.13)
Mother ISCED 6	0.67 (0.09)**		0.92 (0.14)	0.92 (0.14)	0.81 (0.13)	0.82 (0.14)	0.86 (0.15)	0.86 (0.15)
Mother missing	1.00 (0.15)		0.97 (0.15)	0.98 (0.15)	0.91 (0.14)	0.93 (0.15)	0.94 (0.15)	0.96 (0.16)
Father ISCED 3C		0.82 (0.09)+	0.81 (0.09)+	0.81 (0.10)+	0.82 (0.10)+	0.85 (0.10)	0.85 (0.11)	0.84 (0.11)
Father ISCED 3A		0.73 (0.09)*	0.73 (0.09)*	0.73 (0.09)*	0.74 (0.10)*	0.79 (0.10)+	0.80 (0.10)+	0.80 (0.11)+
Father ISCED 5B		0.73 (0.09)*	0.74 (0.10)*	0.74 (0.10)*	0.73 (0.10)*	0.79 (0.11)+	0.81 (0.11)	0.81 (0.11)
Father ISCED 5A		0.59 (0.08)***	0.64 (0.09)**	0.65 (0.09)**	0.62 (0.09)***	0.69 (0.10)*	0.71 (0.10)*	0.74 (0.11)*
Father ISCED 6		0.49 (0.08)***	0.55 (0.09)***	0.55 (0.09)***	0.53 (0.09)***	0.59 (0.10)**	0.62 (0.11)**	0.69 (0.12)*
Father missing		1.13 (0.14)	1.14 (0.14)	1.14 (0.14)	1.15 (0.15)	1.17 (0.15)	1.18 (0.16)	1.12 (0.17)
Female				1.39 (0.08)***	1.40 (0.08)***	1.40 (0.08)***	1.39 (0.08)***	1.40 (0.09)***

Mixed	1.54	1.51	1.47	1.57
	(0.29)*	(0.29)*	(0.28)*	(0.30)*
Indian	1.45	1.46	1.41	1.53
	(0.31)+	(0.31)+	(0.30)	(0.35)+
Pakistani/Bangladeshi	1.14	1.11	1.06	1.07
	(0.15)	(0.16)	(0.15)	(0.17)
Black/Black British	2.16	2.15	2.02	1.77
	(0.57)**	(0.57)**	(0.54)**	(0.51)*
Other Ethnic group	0.91	0.87	0.83	0.91
	(0.25)	(0.24)	(0.23)	(0.26)
Child illness/disability missing	0.74	0.73	0.70	0.71
	(0.52)	(0.52)	(0.50)	(0.54)
Child Illness/Disability Yes	1.13	1.13	1.12	1.10
	(0.09)	(0.09)	(0.09)	(0.09)
North West	1.04	1.05	1.07	1.17
	(0.15)	(0.15)	(0.15)	(0.17)
Yorkshire + the Humber	1.13	1.14	1.16	1.27
	(0.17)	(0.18)	(0.18)	(0.20)
East Midlands	1.02	1.03	1.05	1.10
	(0.15)	(0.16)	(0.16)	(0.17)
West Midlands	1.01	1.02	1.03	1.09
	(0.15)	(0.16)	(0.16)	(0.18)
East of England	0.91	0.93	0.95	1.02
	(0.14)	(0.14)	(0.15)	(0.16)
London	1.00	1.02	1.06	1.10
	(0.14)	(0.14)	(0.15)	(0.16)
South East	0.90	0.93	0.96	1.03
	(0.12)	(0.13)	(0.13)	(0.14)
South West	0.93	0.94	0.94	1.00
	(0.15)	(0.15)	(0.15)	(0.17)

Wales	1.26	1.27	1.28	1.36
	(0.17)+	(0.17)+	(0.17)+	(0.19)*
Scotland	1.01	1.02	1.03	1.14
	(0.14)	(0.14)	(0.14)	(0.16)
Northern Ireland	1.35	1.36	1.36	1.49
	(0.17)*	(0.18)*	(0.18)*	(0.20)**
Father's age	1.00	1.00	1.00	0.99
	(0.01)	(0.01)	(0.01)	(0.01)
Mother's age	1.02	1.03	1.03	1.02
	(0.01)**	(0.01)**	(0.01)***	(0.01)*
dad Illness/Disability YES	1.00	1.00	0.99	0.91
	(0.08)	(0.08)	(0.07)	(0.07)
dad Illness/Disability Missing	0.94	0.92	0.92	0.92
	(0.08)	(0.08)	(0.08)	(0.09)
mum Illness/Disability YES	1.14	1.13	1.12	1.05
	(0.08)+	(0.08)+	(0.08)	(0.08)
mum Illness/Disability Missing	1.28	1.30	1.33	1.18
	(0.53)	(0.54)	(0.56)	(0.45)
Number of children	0.91	0.91	0.89	0.88
	(0.03)**	(0.03)**	(0.03)**	(0.03)***
NS-SEC 2		0.97	0.93	0.86
		(0.11)	(0.11)	(0.10)
NS-SEC 3		0.85	0.81	0.76
		(0.11)	(0.11)	(0.10)*
NS-SEC 4		0.79	0.74	0.71
		(0.13)	(0.13)+	(0.13)+
NS-SEC 5		1.04	0.97	0.80
		(0.18)	(0.17)	(0.15)

NS-SEC 6						0.92 (0.13)	0.85 (0.12)	0.79 (0.12)
NS-SEC 7						0.99 (0.15)	0.91 (0.15)	0.86 (0.14)
Missing						0.94 (0.18)	0.87 (0.17)	0.85 (0.18)
NS-SEC 2						1.17 (0.11)	1.13 (0.11)	1.11 (0.11)
NS-SEC 3						1.06 (0.16)	1.00 (0.15)	1.01 (0.15)
NS-SEC 4						1.24 (0.13)*	1.17 (0.12)	1.19 (0.13)+
NS-SEC 5						1.07 (0.13)	1.01 (0.12)	0.98 (0.12)
NS-SEC 6						1.32 (0.17)*	1.23 (0.16)	1.18 (0.16)
NS-SEC 7						1.30 (0.16)*	1.21 (0.15)	1.12 (0.13)
Missing						1.25 (0.30)	1.16 (0.28)	1.14 (0.28)
Equivalised income							1.00 (0.00)*	1.00 (0.00)
Father BMI								1.09 (0.01)***
Father BMI missing								1.18 (0.12)+
Mother BMI								1.10 (0.01)***
Mother BMI missing								1.11 (0.11)
N	9,705	9,705	9,705	9,705	9,705	9,705	9,705	9,705

Appendix S: Sensitivity analyses: Multiple Imputation

Multiple imputation is utilised in the main document to fill in the missing data points in chapter 6. There were several choices I made in the analysis, each of which could potentially influence the results. Therefore I present here, the results had I made different choices regarding the analysis. This allows me to check whether the results I have reported are sensitive to the specific combination of choices I have made about how to construct the imputation model, or whether the results are reasonably robust to these different choices. The choices I made in the main analysis was specify prediction equations for some variables, to have 20 complete data sets created (m=20) and to allow the algorithm to run through 20 cycles before created each completed data set. I also chose to include to the MCS survey design and attrition weights in the imputation model, as this was the recommended practise.

In the main text I present the results where no prediction equations are specified. Here I present the results for a smaller number of completed data sets and a smaller number of cycles (table A.39 & table A.40). I also present the results where the MCS survey design and attrition weights were not utilised in the imputation model (table A.41). In all cases the 'svy' commands are used for the post imputation analysis. These 'svy' commands take into account the clustering in the MCS data, the stratification of the MCS sample and apply the 'dovwt2' probability weight to the analyses.

The results of the analyses after creating 5 complete data sets with ten cycles of the algorithm are presented in table A.39. In this analysis the MCS survey design and attrition weight was included in the imputation model. The results show the same pattern as those in the imputation model reported in the main analysis, however the magnitude of the differences in the odds is smaller and the standard errors are larger. This means there is less certainty in the estimates, as might be expected with fewer data sets to pool across. As shown in model 5 in table A.39, the difference in the odds for fathers education are not statistically significant, however in the imputation model in the main analysis the differences in the odds are statistically significant for the highest and lowest educated fathers.

Table A.39. Results for multiple imputation with 5 complete data sets and 10 cycles, model weighted with 'dovwt2' weight.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Mother no qualifications						
Mother ISCED 3C	1.00 (0.11)	1.06 (0.12)	1.08 (0.13)	1.13 (0.14)	1.14 (0.14)	1.07 (0.14)
Mother ISCED 3A	0.86 (0.10)	0.95 (0.11)	0.95 (0.12)	1.01 (0.14)	1.02 (0.14)	0.92 (0.13)
Mother ISCED 5B	0.91 (0.11)	1.02 (0.13)	0.99 (0.13)	1.03 (0.14)	1.06 (0.15)	0.97 (0.14)
Mother ISCED 5A	0.63 (0.08)***	0.77 (0.10)*	0.74 (0.11)*	0.76 (0.12)+	0.79 (0.13)	0.79 (0.13)
Mother ISCED 6	0.67 (0.10)**	0.85 (0.13)	0.76 (0.12)+	0.78 (0.14)	0.82 (0.14)	0.81 (0.15)
Father no qualifications						
Father ISCED 3C		0.83 (0.10)	0.84 (0.11)	0.88 (0.12)	0.89 (0.12)	0.88 (0.12)
Father ISCED 3A		0.75 (0.10)*	0.76 (0.10)*	0.82 (0.12)	0.83 (0.12)	0.85 (0.13)
Father ISCED 5B		0.77 (0.11)+	0.76 (0.11)+	0.84 (0.13)	0.86 (0.14)	0.87 (0.14)
Father ISCED 5A		0.66 (0.10)**	0.64 (0.09)**	0.73 (0.12)+	0.76 (0.13)+	0.79 (0.13)
Father ISCED 6		0.57 (0.09)***	0.55 (0.09)***	0.63 (0.11)**	0.67 (0.12)*	0.76 (0.14)
N	9,705	9,705	9,705	9,705	9,705	9,705

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 0 = Mothers education,
Model 1 = Model 0 + Fathers education,
Model 2 = Model 1 + Child's gender
Model 3 = Model 2 + child's ethnicity, whether or not mother, father or child have longstanding illness or disability, mothers age, fathers age, region of residence, and number of children in the household
Model 4 = Model 3 + Mother's social class & Father's social class
Model 5 = Model 4 + equivalised time averaged Income
Model 6 = Model 5 + Mothers BMI & Fathers BMI

In table A.40 I present the results for the imputation model whereby 5 complete data sets were created, but the algorithm was run through twenty cycles. As can be seen in table A.40, the estimated odds ratios and standard errors change very little with the addition of the extra cycles in the imputation model. In table A.41 the results are presented whereby the imputation model did not include the MCS weights, 5 complete data sets and 10 cycles of the algorithm were run through in this specification. Again, the patterning of the results is very similar to all other specifications. However, the standard errors are smaller than when the survey design weights are included in the model.

Table A.40. Results for multiple imputation with 5 complete data sets and 20 cycles, model weighted with 'dovwt2' weight.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Mother no qualifications						
Mother ISCED 3C	0.99 (0.10)	1.06 (0.11)	1.08 (0.12)	1.14 (0.13)	1.15 (0.14)	1.06 (0.13)
Mother ISCED 3A	0.85 (0.09)	0.93 (0.10)	0.94 (0.11)	1.00 (0.13)	1.02 (0.13)	0.91 (0.12)
Mother ISCED 5B	0.90 (0.10)	1.01 (0.12)	0.98 (0.12)	1.03 (0.14)	1.06 (0.14)	0.96 (0.14)
Mother ISCED 5A	0.63 (0.08)***	0.78 (0.10)+	0.75 (0.10)*	0.77 (0.12)+	0.80 (0.12)	0.80 (0.13)
Mother ISCED 6	0.67 (0.09)**	0.85 (0.13)	0.77 (0.12)+	0.79 (0.14)	0.83 (0.14)	0.82 (0.14)
Father no qualifications						
Father ISCED 3C		0.84 (0.10)	0.85 (0.11)	0.88 (0.11)	0.89 (0.11)	0.89 (0.11)
Father ISCED 3A		0.74 (0.09)*	0.75 (0.09)*	0.81 (0.10)+	0.82 (0.10)	0.83 (0.10)
Father ISCED 5B		0.76 (0.10)*	0.75 (0.10)*	0.83 (0.11)	0.84 (0.11)	0.87 (0.11)
Father ISCED 5A		0.65 (0.11)**	0.62 (0.10)**	0.71 (0.13)+	0.74 (0.14)	0.79 (0.14)
Father ISCED 6		0.56 (0.10)**	0.54 (0.10)**	0.61 (0.13)*	0.64 (0.13)*	0.74 (0.15)
N	9,705	9,705	9,705	9,705	9,705	9,705

*+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001*

Model 0 = Mothers education,
Model 1 = Model 0 + Fathers education,
Model 2 = Model 1 + Child's gender
Model 3 = Model 2 + child's ethnicity, whether or not mother, father or child have longstanding illness or disability, mothers age, fathers age, region of residence, and number of children in the household
Model 4 = Model 3 + Mother's social class & Father's social class
Model 5 = Model 4 + equivalised time averaged Income
Model 6 = Model 5 + Mothers BMI & Fathers BMI

Table A.41. Results for multiple imputation with 5 complete data sets and 10 cycles, without weighting applies to the imputation model.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Mother no qualifications	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
Mother ISCED 3C	0.99 (0.10)	1.07 (0.11)	1.09 (0.12)	1.12 (0.13)	1.13 (0.13)	1.05 (0.13)
Mother ISCED 3A	0.85 (0.09)	0.95 (0.11)	0.95 (0.12)	0.99 (0.13)	1.01 (0.13)	0.90 (0.13)
Mother ISCED 5B	0.90 (0.10)	1.04 (0.12)	1.00 (0.12)	1.02 (0.14)	1.05 (0.14)	0.96 (0.14)
Mother ISCED 5A	0.62 (0.08)***	0.79 (0.10)+	0.75 (0.11)*	0.76 (0.12)+	0.79 (0.12)	0.78 (0.13)
Mother ISCED 6	0.66 (0.10)**	0.87 (0.14)	0.78 (0.13)	0.78 (0.14)	0.82 (0.14)	0.80 (0.14)
Father no qualifications		1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
Father ISCED 3C		0.82 (0.09)+	0.83 (0.09)+	0.86 (0.10)	0.86 (0.10)	0.86 (0.10)
Father ISCED 3A		0.73 (0.09)**	0.74 (0.09)*	0.79 (0.10)+	0.80 (0.10)+	0.81 (0.11)
Father ISCED 5B		0.71 (0.10)*	0.70 (0.10)*	0.76 (0.11)+	0.77 (0.11)+	0.78 (0.11)+
Father ISCED 5A		0.61 (0.08)***	0.59 (0.08)***	0.66 (0.09)**	0.68 (0.10)**	0.73 (0.11)*
Father ISCED 6		0.53 (0.09)***	0.51 (0.09)***	0.57 (0.10)**	0.60 (0.11)**	0.68 (0.13)*
	9,705	9,705	9,705	9,705	9,705	9,705

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Model 0 = Mothers education,
Model 1 = Model 0 + Fathers education,
Model 2 = Model 1 + Child's gender
Model 3 = Model 2 + child's ethnicity, whether or not mother, father or child have longstanding illness or disability, mothers age, fathers age, region of residence, and number of children in the household
Model 4 = Model 3 + Mother's social class & Father's social class
Model 5 = Model 4 + equivalised time averaged Income
Model 6 = Model 5 + Mothers BMI & Fathers BMI